

DOCTORAL THESIS
Rosario Sanchis-Font



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

User eXperience
evaluation
on university
virtual learning
through
sentiment
analysis

Directors:
Prof. M^a José Castro-Bleda
Prof. M^a Begoña Jordá-Albiñana

València. March 2023





UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

User eXperience evaluation on university virtual learning through sentiment analysis

DOCTORAL THESIS

Author:

Rosario Sanchis-Font

Directors:

Prof. M^a José Castro-Bleda

Prof. M^a Begoña Jordá-Albiñana

PhD in Design, Manufacture & Management of Industrial Projects.
Research Centre in Graphic Design & Lab Design.

València. March 2023

Abstract

The use of new technologies and the number of users of university online learning systems have spread around the world in the last decades, showing a further increase with the spread of the Covid-19 pandemic since 2020. Additionally, ISO 9241-210:2019 sets international quality standards for designing human-computer interaction products, services, and systems that meet usability, accessibility, and User eXperience (UX) requirements. Therefore, the concept of UX has become very important as a quality requirement. For several authors, UX is a multidimensional concept that includes the motivations, feelings, and needs of end users. On the other hand, the United Nations' (UN) Sustainable Development Goal (SDG) 4 for 2030 aims to ensure inclusive, equitable, quality education for all globally. In this sense, in order to design interfaces and learning experiences in university environments that respect all quality specifications, it is necessary to evaluate the user experience of these environments automatically and accurately beforehand. Thus, the main objective of this thesis is to identify the most relevant specific characteristics in the user experience of university e-learning environments that allow specific and automatic analysis of the students' feelings in order to lay the foundations for the design of user-centered e-learning platforms. To this end, the study proposes to analyse the needs and feelings of online university students with digital, advanced, and efficient artificial intelligence methods. Therefore, this project investigates the application of machine learning models of sentiment analysis for the evaluation of user experience. These artificial intelligence techniques have been applied to the responses received from more than 2,000 university students surveyed from postgraduate online studies and massive open online courses (MOOCs). The results present the basis of a model that allows ontologically classifying categories or aspects of university online education and knowing the users' polarity of feeling about their e-learning experience in an automatic way. In this way, it has been possible

to find out the students' opinions in an automated way with regard to key categories of digital teaching. In addition, student comments have been classified into several UX e-learning or UXEL dimensions. Also, it has been identified the polarity of sentiment for each dimension. To sum up, this work has generated four major contributions to the scientific community. Firstly, an adaptation of the validated questionnaire UEQ-S integrated and adapted to three e-learning platforms for specific postgraduate courses at the Universitat de València and at the Universidad Rey Juan Carlos; and for MOOCs at the Universitat Politècnica de València. Secondly, this thesis has generated an innovative application of sentiment analysis and machine learning methods through natural language processing for the evaluation of the user experience of university online students. Therefore, this method provides the analysis of learners opinions and classifies them according to their polarity in positive, negative or neutral. And thirdly, this scientific work brings out a proprietary ontology of aspects for the virtual learning experience associated with UX dimensions. This ontology used with sentiment analysis tools, allows classifying the polarity of student opinions (positive, neutral, negative) by key categories of e-learning (VLE, Teacher, Student, Sound, Image, Material, Exercise, Evaluation and Communication) and group the comments in three dimensions UX e-learning or UXEL (VLE, Social Connections, and Learning Resources and Tools). Finally, these contributions will help to evaluate in an automatic and accurate way several university e-learning environments in order to design user-centered virtual learning experiences more personalised and inclusive for all which suit quality standards and meet UN SDG 4 for 2030.

Resumen

El uso de nuevas tecnologías y el número de usuarios de sistemas de enseñanza online universitaria se han extendido alrededor del mundo en las últimas décadas, mostrando un mayor incremento con la propagación de la pandemia Covid-19 desde 2020. Adicionalmente, la normativa ISO 9241-210:2019 establece los estándares internacionales de calidad para diseñar productos, servicios y sistemas de interacción persona-ordenador que cumplan con requisitos de usabilidad, accesibilidad y de experiencia de usuario (User eXperience - UX). Por tanto, el concepto de UX ha cobrado mucha importancia como requisito de calidad. Para diversos autores, la UX es un concepto multidimensional que incluye las motivaciones, sentimientos y necesidades de los usuarios finales. Por otra parte, el Objetivo de Desarrollo Sostenible (ODS) 4 de la Organización de las Naciones Unidas (ONU) para el 2030 persigue asegurar a nivel global una educación inclusiva, igualitaria, para todos y de calidad. En este sentido, con el fin de diseñar interfaces y experiencias de aprendizaje en entornos universitarios que respeten todas especificaciones de calidad se requiere evaluar previamente la experiencia de usuario de estos entornos de manera automática y precisa. Por tanto, el objetivo principal de esta tesis es identificar las características concretas más relevantes en la experiencia de usuario de entornos de e-learning universitarios que permitan analizar específica y automáticamente el sentimiento de los estudiantes con el fin de asentar las bases para diseñar plataformas de aprendizaje virtual centradas en los usuarios. Con esta finalidad, el estudio plantea analizar las necesidades y sentimientos de los estudiantes on-line universitarios con métodos digitales, avanzados y eficientes de inteligencia artificial. Por ello, el presente proyecto investiga la aplicación de modelos de aprendizaje automático (*machine learning*) de análisis de sentimiento para la evaluación de la experiencia de usuario.

Estas técnicas de inteligencia artificial se han aplicado sobre las respuestas recibidas de entre los más de 2.000 estudiantes universitarios encuestados procedentes de estudios de posgrado online y de cursos en línea masivos y abiertos (MOOCs). Los resultados presentan las bases de un modelo que permite clasificar ontológicamente categorías o aspectos de la educación en línea universitaria y conocer la polaridad del sentimiento de los usuarios respecto a su experiencia e-learning de manera automática. De este modo, se han podido conocer las opiniones de los estudiantes de manera automatizada con respecto a categorías claves de la enseñanza digital. Además, los comentarios de los estudiantes se han clasificado en distintas dimensiones UX e-learning o UXEL. Así mismo, se ha identificado la polaridad del sentimiento para cada dimensión. En resumen, este trabajo ha generado cuatro importantes contribuciones a la comunidad científica. En primer lugar, una adaptación del cuestionario validado UEQ-S integrado y adaptado a tres plataformas de e-learning para cursos específicos de posgrado en la Universitat de València y en la Universidad Rey Juan Carlos; y para MOOCs en la Universitat Politècnica de València. En segundo lugar, esta tesis ha generado una aplicación innovadora de métodos de análisis de sentimiento y aprendizaje automático mediante el procesamiento del lenguaje natural para la evaluación de la experiencia de usuario de estudiantes universitarios online. Así, este método proporciona el análisis de las opiniones de los alumnos y las clasifica según su polaridad en positivas, negativas o neutras. Y en tercer lugar, este trabajo científico aporta una ontología propia de aspectos para la experiencia de aprendizaje virtual asociada a dimensiones UX. Esta ontología utilizada con herramientas de análisis de sentimiento, permite clasificar la polaridad de las opiniones de los alumnos (positiva, neutra, negativa) por categorías clave del e-learning (VLE, Profesor, Alumno, Sonido, Imagen, Material, Ejercicio, Evaluación y Comunicación) y agrupar los comentarios en tres dimensiones UX e-learning o UXEL (VLE, Conexiones Sociales, y Recursos y Herramientas de Aprendizaje). Por último, estas

contribuciones ayudarán a evaluar de forma automática y precisa diversos entornos universitarios de e-learning con el fin de diseñar experiencias de aprendizaje virtual centradas en el usuario, más personalizadas e inclusivas para todos y todas, que se adapten a los estándares de calidad y cumplan con el ODS 4 de la ONU para el 2030.

Resum

L'ús de noves tecnologies i el nombre d'usuaris de sistemes d'ensenyament en línia universitària s'han estés al voltant del món en les últimes dècades, mostrant un major increment amb la propagació de la pandèmia COVID-19 des de 2020. Addicionalment, la normativa ISO 9241-210:2019 estableix els estàndards internacionals de qualitat per a dissenyar productes, serveis i sistemes d'interacció persona-ordenador que complisquen amb requisits d'usabilitat, accessibilitat i d'experiència d'usuari (*User eXperience* - UX). Per tant, el concepte de UX ha cobrat molta importància com a requisit de qualitat. Per a diversos autors, la UX és un concepte multidimensional que inclou les motivacions, sentiments i necessitats dels usuaris finals. D'altra banda, l'Objectiu de Desenvolupament Sostenible (ODS) 4 de l'Organització de les Nacions Unides (ONU) per al 2030 persegueix assegurar a nivell global una educació inclusiva, igualitària, per a tots i de qualitat. En aquest sentit, amb la finalitat de dissenyar interfícies i experiències d'aprenentatge en entorns universitaris que respecten totes especificacions de qualitat es requereix avaluar prèviament l'experiència d'usuari d'aquests entorns de manera automàtica i precisa. Per tant, l'objectiu principal d'aquesta tesi és identificar les característiques concretes més rellevants en l'experiència d'usuari d'entorns d'*e-learning* universitaris que permeten analitzar específica i automàticament el sentiment dels estudiants amb la finalitat d'assentar les bases per a dissenyar plataformes d'aprenentatge virtual centrades en els usuaris. Amb aquesta finalitat, l'estudi planteja analitzar les necessitats i sentiments dels estudiants en línia universitaris amb mètodes digitals, avançats i eficients d'intel·ligència artificial. Per això, el present projecte investiga l'aplicació de models d'aprenentatge automàtic (*machine learning*) d'anàlisi de sentiment per a l'avaluació de l'experiència d'usuari. Aquestes tècniques d'intel·ligència artificial s'han aplicat sobre les respostes rebudes d'entre els més de 2.000 estudiants universitaris enquestats procedents d'estudis de postgrau en línia i de cursos massius en

línia en obert (MOOCs). Els resultats presenten les bases d'un model que permet classificar ontològicament categories o aspectes de l'educació en línia universitària i conèixer la polaritat del sentiment dels usuaris respecte a la seua experiència *e-learning* automàticament. D'aquesta manera, s'han pogut conèixer les opinions dels estudiants de manera automatitzada respecte a categories claus de l'ensenyament digital. A més, s'han classificat els comentaris dels estudiants en diferents dimensions UX *e-learning* o UXEL. D'aquesta manera, s'han identificat la polaritat del sentiment per a cada dimensió. En resum, aquest treball ha generat quatre importants contribucions a la comunitat científica. En primer lloc, una adaptació del qüestionari validat UEQ-S integrat i adaptat a tres plataformes d'*e-learning* per a cursos específics de postgrau a la Universitat de València i en la Universidad Rey Juan Carlos; i per a MOOCs a la Universitat Politècnica de València. En segon lloc, aquesta tesi ha generat una aplicació innovadora de mètodes d'anàlisi de sentiment i aprenentatge automàtic mitjançant el processament del llenguatge natural per a l'avaluació de l'experiència d'usuari d'estudiants universitaris en línia. Així, aquest mètode proporciona l'anàlisi de les opinions dels alumnes i les classifica segons la seua polaritat en positives, negatives o neutres. I en tercer lloc, aquest treball científic aporta una ontologia pròpia d'aspectes per a l'experiència d'aprenentatge virtual associada a dimensions UX. Aquesta ontologia utilitzada amb eines d'anàlisi de sentiment, permet classificar la polaritat de les opinions dels alumnes (positiva, neutra, negativa) per categories clau de l'*e-learning* (VLE, Professor, Alumne, So, Imatge, Material, Exercici, Avaluació i Comunicació) i agrupar els comentaris en tres dimensions UX *e-learning* o UXEL (VLE, Connexions Socials, i Recursos i Eines d'Aprenentatge). Finalment, aquestes contribucions ajudaran a avaluar de manera automàtica i precisa diversos entorns universitaris d'*e-learning* amb la finalitat de dissenyar experiències d'aprenentatge virtual centrades en l'usuari, més personalitzades i inclusives per a tots i totes, que s'adapten als estàndards de qualitat i complisquen amb l'ODS 4 de l'ONU per al 2030.

Share your knowledge. It is a way to achieve immortality.

Dalai Lama XIV

Agradecimientos

Son muchas las personas a las que les estoy enteramente agradecida pues han contribuido directa o indirectamente, en mayor o en menor medida, en la realización de esta tesis doctoral. Deseo darle especial mención a aquellas personas que forman o han formado parte de mi vida personal, académica y profesional por su relevancia y motivos que detallo a continuación.

Doy infinitas gracias a mis dos directoras de tesis, la Prof. María José Castro y la Prof. Begoña Jordá. Me siento muy afortunada y privilegiada por contar con la dirección de dos mujeres científicas y catedráticas UPV que han podido trasladarme su experiencia consolidada y buena guía hasta el final de esta tesis.

Gracias María José por darle inteligencia, perspectiva, cuerpo, contenido y método a esta tesis, por tus buenos consejos, por tu gran dedicación y por tus valiosas aportaciones además de tu gran amistad; gracias a ti ha sido posible y más fácil hacer esta tesis.

Gracias Begoña, por darle valor a esta tesis, por tu motivación, tu optimismo y por creer en mi desde el principio, por tu apoyo constante, esfuerzo, consejos y sabiduría que ha hecho posible que esta tesis se publique.

Gracias a la Dra. Patricia Díaz-Gimeno, mi mujer, por ser esa luz maravillosa que me ha dado la fuerza para conseguirlo, por su amor, por su dedicación y ayudarme a encontrar y darle sentido a esta tesis y por contagiarme con su excelencia, pasión y vocación científica. Gracias, además, por alentarme, darme el espacio, tiempo y valor para poder escribir los artículos y la tesis con éxito.

Gracias a Nico, mi hijo, por ser mi otra luz, y ser testigo de esta tesis desde su gestación, por su alegría, sonrisa y paciencia constante que me han dado la vitalidad para continuar con la tesis y finalizarla.

Gracias a mi padre y a mi madre por darme la vida y valorar la importancia de la educación y del saber.

Gracias a todos los autores con los que he colaborado en las cinco publicaciones que aquí se recogen, además de la Prof. María José Castro-Bleda y Prof. Begoña Jordá-Albiñana. Estos co-autores son la Dra. Jimena González-Del-Río, la Dra. Olga Ampuero-Canellas, el Dr. Carlos Turró-Ribalta, el Dr. Ignacio Despujol-Zabala, el Dr. José-Ángel González-Barba, el Dr. Ferran Plá, el Dr. Lluís-F. Hurtado y José Luis López Cuerva. También gracias a la Dra. Teresa Magal-Royo por su ayuda en una de las publicaciones.

Enormes gracias a la Fundación IVI, y en concreto a su director el Dr. Nicolás Garrido-Puchalt y a María Martínez de IVI-RMA Global Education, a Medigene Press SL., y a su CEO Loreto Crespo. A todos ellos estoy muy agradecida por su disponibilidad, colaboración y las facilidades mostradas para realizar las estancias de investigación que pude desarrollar en sus centros y que han permitido obtener datos claves para producir esta tesis doctoral.

Y GRACIAS en mayúsculas al Dr. Nacho Despujol y al Dr. Carlos Turró, responsables de la plataforma MOOCs UPV, por su generosidad y esfuerzo en hacer posible la integración del cuestionario UEQ-S en el formulario de evaluación de los usuarios de los MOOCs UPV, además de su dedicación en la obtención y tratamiento de los datos periódicamente a lo largo de cuatro años.

Muchas gracias a Nacho Medina, Fundador y CTO de Zetta Genomics, por su tiempo y darme la oportunidad de desarrollar mi estancia de investigación internacional en Inglaterra con ellos.

Miles de gracias a los más de 2.000 participantes, estudiantes universitarios y usuarios de cursos de posgrados online de la Universitat de València, Universidad Rey Juan Carlos, y de los MOOCs de la Universitat Politècnica de València. Ellos son los que han dado significado a esta tesis al convertirse en la población universitaria de estudio. En concreto doy gracias a los más de 500 usuarios que han participado activamente

respondiendo con sus opiniones, y que posteriormente hemos podido porcesar y analizar con técnicas de inteligencia artificial.

Gracias a los estudiantes a los cuales he dirigido experimentalmente en sus TFGs en la ETSINF UPV, por sus aportaciones indirectas en esta tesis: Salehodine Ben Hassen Jemni, José Luis Lopez Cuerva y Pravin Venkatraman.

Gracias a los miembros evaluadores y del tribunal de esta tesis por contribuir a que se publique con proyección internacional y con determinados niveles de excelencia.

Gracias a la UPV, a su Escuela de Doctorado, a servicio de Biblioteca, a la Escuela Técnica Superior de Ingeniera del Diseño, la Escuela Técnica Superior de Ingeniería Informática, al grupo de investigación ELiRF y al Instituto VRAIN, al Departamento de Ingeniería Gráfica, a la Comisión académica del Programa de doctorado en Diseño, Fabricación y Gestión de Proyectos Industriales, en especial a su director, el Prof. Miguel Ardid. Gracias Miguel por tu predisposición, atención y facilitar el proceso en la comisión y trámites necesarios. Y gracias al equipo rectoral UPV por su buen gobierno orientado a impulsar trabajos científicos con una I+D, innovación y transferencia de calidad.

Gracias a toda mi familia, la de origen y la elegida, a mis amigas y amigos por escuchar mis alegrías y penas y apoyarme en este proceso de investigación predoctoral y redacción de la tesis de alguna manera, haciendo este camino más agradable y distendido. En especial, a mis padres políticos Inma y Ángel por su gran ayuda y cariño, a mis sobrinos, Miguel, Arnau y Marc, a mis tíxs, primxs, a mis hermanas políticas Inma y Teresa por su interés y apoyo. A mis amigxs, en especial a Marta, M^a José y sus Pablos, Rosa F., Belén, Carol, Samer, Carmen, Sara R. y Jon, Paco y Ángeles, Alberto y Valeria, Mire y Vicen, Ángela y Álvaro, Raquel, Fabiola, Sara S., Cris, Claudia y José, Beita, Isa, Amparo, M^a José M., Mónica y Augusto, Ana D., Debora... y muchísimas más personas que recojo en mis recuerdos y que me han animado en este camino predoctoral. Y a los compis, mamás y papás, amigxs del cole de Nico por lo buenos momentos en familia que han ayudado a desconectar. También

gracias a aquellas personas que con su pequeña ayuda me facilitaron que pudiera avanzar con esta tesis.

Gracias a Taiga, mi gata, por acompañarme con su ronroneo pegada a mi y al portátil a cualquier hora del día mientras escribía esta tesis y sus artículos. Y a Olga por su preciada ayuda en las tareas domésticas que me han permitido dedicarle tiempo de calidad a la tesis.

Gracias a las dificultades encontradas en estos años, provocadas por personas del ámbito académico, profesional y personal, que voluntaria o involuntariamente me han puesto en situaciones dolorosas. Enfrentarme a estos momentos difíciles ha sido una oportunidad para crecer rápidamente espiritualmente y desarrollarme como ser humano. Gracias a estas circunstancias he podido encontrar nuevos recursos y habilidades personales y profesionales que hoy me hacen sentirme más valiosa y alineada con mi misión en esta vida. En especial, quiero darles las gracias con profundo respeto y amor a mi hermano, Miguel, a mi madre, Rosario, y a mi padre, Miguel.

Grandes gracias a las personas que me han ayudado con su profesionalidad a sobrellevar la dureza de los dos últimos años por temas personales, y que me ha permitido encontrar la tranquilidad para concentrarme mejor en el desarrollo de esta tesis. A mis grandes amigas la Dra. Marian R., Sara R., Rosa B. y Asún. Y también enormes gracias a Esther, Violeta, Merche, Gisela y Jesús, entre otros.

Gracias a mis compañerxs de trabajo de la Fundación CPI por su apoyo moral y escucha. En especial a mi compi Lupe, Inma, Carlos, Euse, Adrián, Irene, Miguel, Juanvi, Fernando y Salva.

Gracias a la Delegación de Estudiantes de Doctorado UPV, en concreto al equipo de doctorandos que me acompañaron en 2018 durante mi representación como delegada de alumnos de doctorado de la UPV, muchos de ellos ya doctores. Gracias a Ana Albalat, Vanessa Grazón, Enrique Serra, Narce D.G. Guzmán, Sandra Martínez, Daniel Catalá y Amparo Bonet por crear equipo conmigo en la delegación, colaborar con las propuestas y movilizarlas juntos en la UPV.

Y gracias a las personas que me ayudaron a iniciar este proyecto doctoral en la UPV. Gracias a la Dra. Helena Rodríguez-Hernández por ponerme en contacto con el departamento y al Dr. José Luis Martín Montesinos por su generosidad, buenos consejos y ayuda académica. A ambos gracias por estar en el lugar, momento y modo adecuado en 2013.

Gracias a todas las personas que han hecho y harán posible que este proyecto científico vea la luz y pueda brillar durante décadas. Y también quiero dar las gracias de antemano a todas aquellas personas que me acompañarán en un futuro en generar y visibilizar mi producción científica, estimularme y fomentar mi curiosidad por la ciencia.

Y finalmente, gracias a ti por leer esta tesis doctoral.

Nota de la autora

La **tesis doctoral** “User eXperience evaluation on university virtual learning through sentiment analysis” realizada por Rosario Sanchis-Font se enmarca en el **Programa de doctorado en Diseño, Fabricación y Gestión de Proyectos Industriales** de la Universitat Politècnica de València (UPV).

La tesis se ha desarrollado por **compendio de artículos** presentando una compilación de cinco artículos publicados en revistas científicas y en actas de congresos de impacto, siendo aceptada y autorizada por la Comisión Académica del Programa de Doctorado (CAPD) al que se adscribe.

Es una **tesis de contenido interdisciplinar** donde han intervenido, tanto en la dirección de la misma como en la colaboración para la redacción artículos, principalmente personal docente investigador (PDI) UPV del departamento de **Ingeniería Gráfica** de la Escuela Técnica Superior de Ingeniería del Diseño y del Departamento de **Sistemas Informáticos y Computación** (DSIC) de la Escuela Técnica Superior de Ingeniería Informática. La tesis hibrida conocimiento científico de ingeniería del diseño interactivo con herramientas de inteligencia artificial de aprendizaje automático y procesamiento del lenguaje natural para ofrecer resultados de investigación originales e innovadores en ambas disciplinas.

Para llevar a cabo el proceso de investigación y producir los resultados de esta tesis doctoral se han **realizado tres estancias de investigación**, dos nacionales y una internacional, y una **colaboración** con el equipo de los MOOCs UPV del **Área de Sistemas de Información y Comunicaciones** (ASIC) de la UPV.

Las **estancias nacionales** de investigación se han desarrollado en la Fundación IVI y en Medigene Press S.L. para la evaluación de la experiencia e-learning de usuarios de posgrados biomédicos universitarios de la Universitat de València y Universidad Rey Juan Carlos.

La **estancia internacional** se ha desarrollado en Inglaterra, en la spin-off Zetta Genomics LTD, empresa surgida y creada por equipo investigador del Computational Biology Lab de Cambridge University. La estancia se ha desarrollado con el propósito de conocer y evaluar la experiencia de usuario interactiva de una plataforma web genómica para la comunidad biomédica de Reino Unido. Esta estancia internacional ha contado con una financiación pública ERASMUS+.

Finalmente, en el proceso de desarrollo de esta tesis se ha realizado un **proyecto de colaboración** con el equipo del ASIC de la UPV. En concreto, junto con los responsables de la plataforma e-learning MOOCs UPV, el Dr. Carlos Turró y el Dr. Nacho Despujol, se ha llevado a cabo los pasos necesarios para obtener los datos para la evaluación de la experiencia de usuario de los estudiantes MOOCs de la UPV.

Index

1. INTRODUCTION	1
1.1. BACKGROUND	3
1.1.1. E-learning	3
1.1.2. User eXperience (UX)	3
1.1.3. User-Centered Design (UCD).....	5
1.1.4. UX evaluation	6
1.1.5. UX and UCD in e-learning	7
1.1.6. Natural Language Processing and Sentiment Analysis	8
1.1.7. Sentiment analysis in e-learning.....	12
1.2. OBJECTIVES	13
1.2.1. General objectives	13
1.2.2. Specific objectives	13
1.3. THESIS STRUCTURE	14
1.3.1. Publication 1	14
1.3.2. Publication 2	14
1.3.3. Publication 3	15
1.3.4. Publication 4	16
1.3.5. Publication 5	16
2. PUBLICATIONS	19
2.1. Publication 1 IMPROVING THE VIRTUAL LEARNING EXPERIENCE: USER CENTERED DESIGN IN E-LEARNING	21
2.2. Publication 2 INTEGRACIÓN DEL “USER EXPERIENCE QUESTIONNAIRE SHORT” EN MOOCS UPV	35

2.3. Publication 3 APPLYING SENTIMENT ANALYSIS WITH CROSS-DOMAIN MODELS TO EVALUATE USER EXPERIENCE IN VIRTUAL LEARNING ENVIRONMENTS	51
2.4. Publication 4 CROSS-DOMAIN POLARITY MODELS TO EVALUATE USER EXPERIENCE IN E-LEARNING	65
2.5. Publication 5 E-LEARNING UNIVERSITY EVALUATION THROUGH SENTIMENT ANALYSIS CENTERED ON USER EXPERIENCE DIMENSIONS	83
3. DISCUSSION	105
4. CONCLUSIONS / CONCLUSIONES	111
4.1. General conclusions.....	113
4.2. Achievement of Goals	114
4.3. Contributions made	116
4.4. Further lines of research	119
4.5. Conclusiones en español	120
ABBREVIATIONS AND ACRONYMS.....	127
REFERENCES (Introduction and Discussion).....	129

1. INTRODUCTION

1.1. BACKGROUND

1.1.1. E-learning

In the XXIst century, online learning in higher education has become more popular due to its accesibility, flexibility and personalized approach (Al-Fraihat, 2019). E-learning allows students to learn at their own pace, time, and place. That has been increasing in last years when the COVID-19 pandemic started in 2020. This fact fostered a significant increase in the use of online or hybrid teaching formats and the use of Information and Communication Technologies (ICT) in education (Ntshwarang, 2021).

On the other hand, the United Nations' 2030 Agenda for Sustainable Development includes Goal 4 to ensure inclusive, equitable, and quality education for all, including university education. UNESCO's report on this goal emphasizes the need for inclusive and quality ICT tools for learning (2021).

1.1.2. User eXperience (UX)

Therefore, research on e-learning for higher education is focusing on human-computer interaction (HCI) in order to improve the success of university student training and the quality of online learning (Bates, 2005; Zaharias, 2009; Ali, 2017). Factors evaluated in interactive learning include teachers, e-learning systems, and educational content (Alebeisat, 2022).

HCI tools developers, agents and industry require to focus their interactive systems on end-users to design and provide quality systems upon the international standards requirements ISO. These interactive systems are the “combination of hardware, software and/or services that receives input from, and communicates output to, users” (ISO 9241-210:2019). This international standard is related to ergonomics of human system-interaction and human-centered design for interactive systems.

Also, the ISO 9241-210:2019 standard provides specifications for the quality of HCI systems and includes a framework for interaction that takes into account usability, accessibility, and user experience based on user-centered design.

ISO describes user experience (UX) as all the emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and achievements that occur before, during and after the use of interactive systems. So, UX enhances human interaction within the hardware or software components. In all this context, UX in e-learning has gained importance and several authors have pointed out its definition. For Zaharias (2012) UX is a multidimensional concept focused on human needs and aspects of beauty, fun, pleasure and personal growth that are experienced by human interaction with the computer. Other authors, such as Rauschenberger et al. (2013) define UX as a concept that encompasses both pragmatic (clarity, efficiency and dependence) and hedonic (stimulation and novelty) qualities. Along these lines, Hassenzahl (2006) highlights the importance of the pragmatic aspect of interactive products and the hedonic aspect for the design of experiences, understanding the UX concept as a consequence of the user's internal state (predispositions, expectations, needs, motivation, mood, etc), the characteristics of the designed system (complexity, purpose, usability, functionality, etc.), and the context.

Learning methods, services and products have evolved towards a digital stage, where all educational agents (students, teachers, administrators, developers and stakeholders) interact on virtual environments every day, not only for educational aims but also in personal and global management of educational data.

From the end-users point of view, they all access to the digital media no matter their age, physical abilities, cultural backgrounds, place or time of use, navigate, study, teach, do exercises, publish, evaluate, participate, socialize and behave differently within the same virtual learning environment. Therefore, it is required to design e-learning systems and

interfaces, which consider users' skills, motivations, feelings and needs in order to succeed in the learning process. The UX includes a multidimensional concept and focusses more in human needs and the aspects of beauty, fun, pleasure, and personal growth rather than the value of the product or instrument used (Kujala, 2011; Zaharias & Mehlenbacher, 2012).

1.1.3. User-Centered Design (UCD)

UX performance demands a User-Centered Design (UCD) to develop systems that meet users' needs. UCD is a flexible iterative design methodology, a philosophy and a discipline that understands and analyses how people use computers, objects and systems to propose more usable, accessible, beautiful and easy-to-use interfaces. The aim of UCD in technology is to make good interfaces for delivering pleasant experiences to users when interacting with computers (Norman, 1986; Adelson, 2010).

Several UCD methodologies have been carried out by different researchers highlighting the importance of user involvement in the design process. On this approach users are an active agent in the system design considering users' needs and preferences result of their participation (Damodaran, 1996; Courage & Baxter, 2005; Brown & Katz, 2011). All of the previous authors show different design methodologies, which involve the end-user from the beginning of the design process and being users the center of the process for interaction design. This is also known as co-creation, collaborative or participatory design (J. Gulliksen et al., 2003).

These studies help to develop interactive products which are more efficient, easy, beautiful and engage to end-users. So, one of the aims in Human-Computer Interaction (HCI) research is to know which qualities are involved in UX.

1.1.4. UX evaluation

From this perspective, there are several aspects to consider when understanding UX to develop a HCI system. On the approach to measure the experiential qualities in HCI systems some authors have researched on finding the qualities and dimensional aspects related to delivery UX questionnaires (Hassenzahl et al. 2003; van Schaik & Ling, 2012a; Rauschenberger et al. 2013; Law et al. 2014). Rauschenberger (2013) included in the User Experience Questionnaire (UEQ) 6 scales with 26 items in total covering the attractiveness (general impression towards the product); the classical usability aspects: efficiency (possibility to use the product fast and efficient and user interface looks organized); perspicuity (easy to understand how to use and to get familiar with the product); dependability (users feeling in control and security and predictability when interacting with the product); and user experience aspects: stimulation (users interest and excitement when using the product and feeling of motivation to use the product again) originality (innovative and creative design of the product and how the product grab users attention).

In the interest of understanding the user's opinion from this perspective, Spallazo (2021) brings together the 129 UX evaluation methods for interactive systems. The most remarkable methods to evaluate UX for several authors (Diaz-Oreiro, 2021) are AtrakkDiff, UEQ or meCUE. Although these UX evaluation tools are very widespread by the experts, their application for online education is very limited, as they are laborious methods to be processed and not very specific.

The use of these questionnaires to evaluate user experience has helped to measure the qualities for designing virtual interfaces which include the motivations, feelings, and needs of end users.

1.1.5. UX and UCD in e-learning

So, in order to carry out the learning process successfully through online university courses UX requirements need to be considered when designing and redesigning hardware and software applications. In this way, in the last years, UX has been considered when designing Virtual Learning Environments (VLEs) (Zaharias et al. 2012). VLEs include a wide range of technology-enabled learning environments, such as Learning Management Systems (LMSs), computer games or Virtual Worlds.

In the new millennium, many studies have been carried out to understand human experience in virtual environments. There is a small field of research of UX applied to virtual learning environments where we can find its basis on the studies of usability features when designing educational software and e-learning systems (Squires,1999; Zaharias, 2004 and 2011). Usability refers to e-learning interfaces which are effective, efficient and satisfy end-users. In this sense, e-learning studies from UX perspective complete the concept of usability by including new aspects such as learners' motivation and engagement.

The methodology of UCD to improve e-learning environments and its interfaces considers the user experience in all the stages of the iterative process: understand and analyse user needs and context of use, specifying user requirements, ideate, prototype and evaluate (Gena, 2006; Garreta-Domingo and Mor, 2007).

De Lera and others have applied UCD research on the e-learning environments pursuing an improvement of the learner's interaction and entertainment (De Lera et al., 2013). The authors include the emotional aspects with UCD methodology applied to educational learning environments. They carried out a research project on Open University of Catalonia (UOC) virtual platform and found out the need of including the emotional and more "human" dimension in the design of interfaces to achieve more engaging and enjoyable learning environments, which

consequently will enhance the global user experience (GUX) of the UOC learners.

In order to evaluate the UX in e-learning, authors such as Zaharias (2012), Mtebe (2015) or Ovesleová (2016) have established some UX categories or criteria applied to the field of online education. The authors clarify the guidelines to ensure a user-centered design of interactive learning platforms and a positive and quality experience, but without developing automated methods.

A deep understanding of the UX of these educational platforms will enhance the design of environments that meet the functional, aesthetic, and emotional characteristics and needs required by users. UX evaluation is in development in most applications and, also, in VLEs.

1.1.6. Natural Language Processing and Sentiment Analysis

Some work has been done in eCommerce, using natural language processing to improve their UX, for instance, to search products in a more intelligent way, using artificial intelligent tools such as sentiment analysis to extract insights from the reviews made by the customers on the product or identifying trends and trying to answer best to the customers' concerns. Several conferences have been launched around these ideas, such as the Workshop on Economics and Natural Language Processing (NLP) (<https://julielab.de/econlp/2019/>) or the First International Workshop on e-Commerce and NLP (<https://www.aclweb.org/portal/content/first-international-workshop-e-commerce-and-nlp>)

Sentiment analysis is the process of using NLP and machine learning techniques to analyze and determine the emotional tone or attitude expressed in a piece of text. The text could be a text from social media, a product review, a news article, or any other form of textual data. The goal of sentiment analysis is to identify

whether the overall sentiment expressed in the text is positive, negative, or neutral. This is typically done by analyzing the words, phrases, and context of the text to determine the emotional tone of the text. Sentiment analysis can be used in various fields, including marketing, customer service, and politics, to gauge public opinion, measure customer satisfaction, and identify trends and patterns in data. Sentiment analysis is a field of research related to computational linguistics, Natural Language Processing, and text mining (Mejova, 2009) and it is one of the most active areas in Natural Language Processing since the early 2000s.

So, machine learning models of sentiment analysis will allow to classify the opinion found in the text under two opposing sentiments, based on their Positive (P) or Negative (N) polarity. If the text does not have any polarity, is classified as Neutral (NEU).

Processing the natural language on this opinion analysis requires of a corpus. Corpus is a collection of linguistic data, either compiled from written texts or transcribed from recorded speech (Khurana, 2022). Another important concept in artificial intelligence is the ontology of the corpus. For Fensel (2001) *“ontologies are developed to provide a machine-processable semantics of information sources that can be communicated between different agents (software and humans)”*. An ontology provides a vocabulary of terms and relations with which to model the domain. Therefore, ontology catalogs the variables required for some set of computation and establishes the relationships between them.

Pang et al. (2002) and Turney PD (2002) addressed the importance of “sentiment classification” for a big number of tasks such as “message filtering, recommender systems or business intelligence applications”. A decade after, until our days, the popularity of sentiment analysis has been increasing and Deep Learning has consolidated as a well-established alternative to the previous machine learning systems. Thus, Deep Learning is

the state of the art in sentiment analysis (Socher et al., 2013; Kim, 2014; Baziotis, 2017; González, Hurtado and Plá, 2018; González et al, 2020).

Other sentiment analysis approaches were addressed by manually generating polarity lexicons (Liu et al., 2005; Wilson et al., 2005). However, the efforts required to develop these resources and the good performance of machine learning systems on this task made the research community to move towards data driven approaches. A survey of the most widely used machine learning approaches for the sentiment analysis problem can be found in Liu (2012).

Recently, the predominant systems to perform sentiment analysis are neural network-based approaches (Zhang et al., 2018). The most popular models are Convolutional Neural Networks (CNN) Kim (2014), Long Short-Term Memories (LSTM) (Hochreiter and Schmidhuber, 1997), and combinations of CNN and LSTM (Sadr et al., 2019). Moreover, the enrichment of these architectures by using attention mechanisms (Bahdanau et al. 2015) and Transformers (Vaswani et al., 2017) are lately used.

The interest on sentiment analysis has increased along with the popularity of social networks and the user interactions on them. The most studied social network for sentiment analysis tasks is Twitter, where the users are allowed to broadcast opinions about any topic by using only 280 characters and media content.

Several workshops are organized in order to address the sentiment analysis task in Twitter, providing corpora and resources to the participants for training and evaluating their systems. The most known workshops are the International Workshop on Semantic Evaluation (SemEval) and the Workshop on Semantic Analysis at SEPLN (TASS and IberLEF) for English and Spanish language, respectively.

For the last task of English sentiment analysis presented at SemEval (Rosenthal et al. 2017), most of the participating teams

proposed neural network models mainly based on LSTM and CNN, being the two best systems based on these approaches along with pre-trained word embeddings on big collections of tweets. Concretely, the winner team proposed a two-layer bidirectional LSTM with attention mechanisms (Baziotis et al. 2017), while the second ranked team addressed the task by using a combination of LSTM and CNN (Cliche M, 2017).

For the Spanish sentiment analysis task of TASS 2019 (Diaz-Galiano et al, 2019) and TASS 2020 (González et al, 2020) the predominant presence of deep learning components was also observable, where almost all the systems proposed by the participants made use of them. It is worthy to note the great interest on the Transformer model (Vaswaniet et al., 2017) being used mainly with the aim of fine-tuning pretrained contextual representations of words (Devlin et al., 2019; González et al., 2021).

In addition to these kinds of systems, a large number of commercial products and frameworks have also proliferated to facilitate the development and deployment of sentiment analysis systems based on machine learning, such as GoogleCloud (Cloud Natural Language API, 2019, <https://cloud.google.com/natural-language>), IBM Watson Natural Language Understanding, (2019 <https://www.ibm.com/cloud/watson-natural-language-understanding>), Microsoft Text Analytics (Microsoft Azure: Text Analytics API (2019). <https://azure.microsoft.com/es-es/services/cognitive-services/text-analytics/>) and MeaningCloud (MeaningCloud: Demo de Analítica de Textos (2019). <https://www.meaningcloud.com/es/demos/demo-analitica-textos>). These kinds of products allow us to perform text analytics such as sentiment analysis, in a broad variety of domains and languages in an easy way, obtaining also competitive results.

1.1.7. Sentiment analysis in e-learning

The application of sentiment analysis tools on UX opinions will provide a better and more accurate understanding of human needs in the interaction with VLEs.

On the other hand, Natural Language Processing technologies allow to analyze students' opinions from text and sentiment analysis. Thus, Clarizia et al. (2018) propose to use data mining to analyze the sentiment present through text comments among students in order to allow the teacher to accommodate the online learners' moods. The authors aim to outline a model that successfully responds to the analysis task, rather than the results of the learning experience, by obtaining an overall analysis of user sentiment. Depending on the level at which one wishes to treat the text, one can extract the polarity of the whole document, Moraes et al. (2013), the polarity of each sentence or the polarity of each aspect appearing in the text.

Currently, there are several commercial NLP tools on the market that analyze sentiment on texts automatically. Thus, MeaningCloud, GoogleCloud Natural Language or Microsoft Azure Text Analytics stand out for their widespread use. In this area, we found studies of sentiment analysis of user comments on their experience with commercially available techniques, Zulkifli et al. (2019). These authors analyze the polarity of English customer reviews from Amazon, Yelp, and IMDb on social networks using three tools: Python NLTK Text Classification, Myopia and MeaningCloud. The study showed that MeaningCloud is the technique with the highest accuracy at 82.1%. MeaningCloud is a commercial text analytics NLP application programming interface (API) and tool, released in 2015 as an evolution of an earlier product called Textalytics. Both tools have been validated by several authors in the study of their performance (Joshi et al. 2018; Singh et al., 2018; Zulkifli et al., 2019).

Other preliminary studies on the application of sentiment analysis techniques to analyze the perceptions of distance learners have

also been published (Mac Kim et al., 2010; Magayon et al., 2021). Both works focus on general aspects of the experience without addressing specific characteristics of the online learning process from a user experience perspective. But the problem of detecting specific areas, aspects and features of improvement of the e-learning user experience in an automatic way remains unresolved.

1.2. OBJECTIVES

1.2.1. General objectives

Identify and lay the foundations of the most relevant user experience characteristics of university e-learning environments that allow automatic analysis of user sentiment.

1.2.2. Specific objectives

1.2.2.1. Select the characteristics to be studied in our university population related to user experience in e-learning environments.

1.2.2.2. To survey the university population users of the e-learning platforms in reference to the selected characteristics.

1.2.2.3. To analyse the feeling of university users' experience with different machine learning models. This analysis aims to choose the most appropriate models that best fit the reality we want to investigate.

1.2.2.4. Categorise the most relevant user experience characteristics in university e-learning environments obtained with the most appropriate machine learning models of natural language processing and sentiment analysis.

1.3. THESIS STRUCTURE

This doctoral thesis is a compilation of five indexed scientific publications that seek to respond to the general objective and specific objectives previously indicated in this scientific work.

Thus, this thesis by compendium of articles gathers in chapter 2 the five mentioned publications. Each of the articles presents its bibliography at the end of each one. Subsequently, the discussion of this doctoral thesis is presented in chapter 3, and its conclusions in chapter 4. Conclusions are written in both languages, English and Spanish. Then, it follows a list of abbreviations and acronyms to facilitate the reading of the thesis.

And finally, the book ends with a section with the references of the chapters, except for the references of the articles, which they are already presented at the end of each publication.

1.3.1. Publication 1

IMPROVING THE VIRTUAL LEARNING EXPERIENCE: USER CENTERED DESIGN IN E-LEARNING. R. Sanchis-Font, B. Jordá-Albiñana, J. González-Del-Río, O. Ampuero-Canellas.

International conference INTED. 11th annual International Technology, Education and Development Conference. Valencia. March, 2017.

DOI: [10.21125/inted.2017.2356](https://doi.org/10.21125/inted.2017.2356)

- Article published in IATED Digital Library, which also it is included in the Web of Science (Conference Proceedings Citation Index).

1.3.2. Publication 2

INTEGRACIÓN DEL “USER EXPERIENCE QUESTIONNAIRE SHORT” EN MOOCS UPV.

Sanchis Font, R., Castro Bleda, M. J., Turró Ribalta, C., & Despujol Zabala, I. (2018, september). In IN-RED 2018. IV Congreso Nacional de Innovación Educativa y Docencia en Red (pp. 393-404). Editorial Universitat Politècnica de València.

DOI: [10.4995/INRED2018.2018.8840](https://doi.org/10.4995/INRED2018.2018.8840)

- Article published in book proceedings of IN-RED Conference (2018) por la editorial Universitat Politècnica de València y en el Repositorio Institucional de la Universitat Politècnica de València (RiuNet).
- Article indexed in Dialnet, the bibliographic portal for Hispanic scientific literature, and also indexed in Semantic Scholar.

1.3.3. Publication 3

APPLYING SENTIMENT ANALYSIS WITH CROSS-DOMAIN MODELS TO EVALUATE USER EXPERIENCE IN VIRTUAL LEARNING ENVIRONMENTS

Sanchis-Font, R., Castro-Bleda, M.J., González, JÁ. (2019). In: Rojas, I., Joya, G., Catala, A. (eds) Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science(), vol 11506. Springer, Cham.

DOI: https://doi.org/10.1007/978-3-030-20521-8_50

- Article published as a chapter in Lecture Notes in Computer Science book series (LNCS, volume 11506) which it is included in the Web of Science.
- SCImago Journal Rank (SJR) index in 2019 of LNCS book series: 0.427 / Q2 in Computer Sciences (miscellaneous).

1.3.4. Publication 4

CROSS-DOMAIN POLARITY MODELS TO EVALUATE USER EXPERIENCE IN E-LEARNING. Sanchis-Font, R., Castro-Bleda, M.J., González, J. et al. Neural Process Letters (2020). Springer.

DOI: <https://doi.org/10.1007/s11063-020-10260-5>

- Article published in the international journal Neural Process Letters, which it is included in the Web of Science and indexed in Journal Citations Report (JCR).
- JCR index in 2021 for Neural Process Letters: 2.565 / Q3 in Computer Sciences and Artificial Intelligence.
- SJR index in 2020 for Neural Process Letters: 0.463 / Q2 in Artificial Intelligence, Computer Communications and Networks.

1.3.5. Publication 5

E-LEARNING UNIVERSITY EVALUATION THROUGH SENTIMENT ANALYSIS CENTERED ON USER EXPERIENCE DIMENSIONS. Sanchis-Font, R., Castro-Bleda, M., Jorda-Albiñana, M., Lopez-Cuerva, L. (0). DYNA, DYNA-ACCELERADO (0). [5 pp]. (2022)

FONT, R. S., BLEDA, M. J. C., ALBIÑANA, M. B. J., & CUERVA, L. L. (2023). EVALUACIÓN DEL APRENDIZAJE EN LÍNEA UNIVERSITARIO MEDIANTE ANÁLISIS DE SENTIMIENTO CENTRADO EN DIMENSIONES DE EXPERIENCIA DE USUARIO. DYNA, 98(2), 147-153.

DOI: <https://doi.org/10.6036/10603>

- Article published in the international engineering journal DYNA, which it is included in the Web of Science and indexed in Journal Citations Report (JCR).

- JCR DYNA index in 2021: 2.070 / Q3 in Engineering, Multidisciplinary.
- SJR DYNA index en 2021 (Spain): 0.16 / Q4 in Engineering, (miscellaneous).

2. PUBLICATIONS

2.1. Publication 1 | IMPROVING THE VIRTUAL LEARNING EXPERIENCE: USER CENTERED DESIGN IN E-LEARNING

R. Sanchis-Font, B. Jordá-Albiñana, J. González-Del-Río, O. Ampuero-Canellas.

Congreso Internacional INTED. 11th annual International Technology, Education and Development Conference. Valencia. Marzo 2017.

DOI: [10.21125/inted.2017.2356](https://doi.org/10.21125/inted.2017.2356)

Publication 1 presents a review of different methodologies and features for designing user experience-oriented e-learning that integrates aspects of user-centred design. The findings will help to improve the factors to take into account for designing the experience of e-learning for both students and teachers in a virtual learning interfaces in education within a user centered design process.

IMPROVING THE VIRTUAL LEARNING EXPERIENCE: USER-CENTERED DESIGN IN E-LEARNING

**R. Sanchis-Font, B. Jordá-Albiñana,
J. González-Del-Río, O. Ampuero-Canellas**
*Departamento de Ingeniería Gráfica.
Universitat Politècnica de València. (SPAIN)*

Abstract

Late developments and evolution of virtual learning make users to demand more usable, accessible, friendly, engaged and affective interfaces in order to achieve successfully the process of learning. User eXperience (UX) design highlights the human experiential aspect when designing virtual learning environments and focusses in the different users as co-creators to build effective and sustainable e-learning interfaces. From this perspective, User-Centered Design (UCD) methodology integrates UX to develop e-learning interfaces for all learners (students and teachers). The article presents a methodology for an effective virtual learning oriented to UX and UCD issues involved. The findings will help to improve the factors to take into account for designing the experience of e-learning for both students and teachers in a virtual learning interfaces in education.

Keywords: e-learning, User-Centered Design (UCD), User eXperience (UX) interfaces, interface design.

Introduction

Learning methods, services and products have evolved towards a digital stage, where all educational agents (students, teachers, administrators, developers and stakeholders) interact on virtual environments every day, not only for educational aims but also in personal and global management of educational data.

From the end-users point of view, they all access to the digital media no matter their age, physical abilities, cultural backgrounds, place or time of use, navigate, study, teach, do exercises, publish, evaluate, participate, socialize and behave differently within the same virtual learning environment. Therefore, it is required to design e-learning systems and interfaces, which consider users' skills, motivations, feelings and needs in order to succeed in the learning process. The user experience (UX) includes a multidimensional concept and focusses more in human needs and the aspects of beauty, fun, pleasure, and personal growth rather than the value of the product or instrument used (Zaharias and Mehlenbacher, 2012) [1], (Kujala, 2011) [2].

In this perspective user experience in Human Computer Interaction (HCI) is describe by the ISO 9241-210:2010 normative as:

"...person's perceptions and responses resulting from the use and/or anticipated use of a product, system or service. User experience includes all the users' emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviours and accomplishments that occur before, during and after use." (ISO, 2010) [3].

There is a small field of research of UX applied to virtual learning environments where we can find its basis on the studies of usability features when designing educational software and e-learning systems (Squires,1999) [4], (Zaharias 2004, 2011) [5] [6]. Usability refers to e-learning interfaces which are effective, efficient and satisfy end-users. In late e-learning studies from UX perspective the concept of usability is completed by new aspects such as learners' motivation and engagement.

UX performance demands a User-Centered Design (UCD) to develop systems that meet users' needs. UCD is a flexible iterative design methodology, a philosophy and a discipline that understands and analyses how people use computers, objects and systems to propose more usable, accessible, beautiful and easy-to-use interfaces. The aim of UCD in technology is to make good interfaces

for delivering pleasant experiences to users when interacting with computers (Norman, 1986) [7] (Adelson, 2010) [8].

Several UCD methodologies has been carried out by different researchers highlighting the importance of user involvement in the design process. On this approach users are an active agent in the system design considering users´ needs and preferences result of their participation, (Damodaran, 1996) [9], (Courage & Baxter, 2005) [10], (Brown & Katz, 2011) [11]. All of them show different design methodologies, which involves the end-user from the beginning of the design process and being users the center of the process for interaction design. This is also known as co-creation, collaborative or participatory design (J. Gulliksen et Al.,2003) [12].

The methodology of UCD to improve e-learning environments and its interfaces considers the user experience in all the stages of the iterative process: understand and analyse user needs and context of use, specifying user requirements, ideate, prototype and evaluate (Gena, 2006) [13], (Garreta-Domingo and Mor, 2007) [14].

De Lera and others have applied UCD research on the e-learning environments pursuing an improvement of the learner´s interaction and entertainment (De Lera,et Al.,2013) [15]. The authors include the emotional aspects with UCD methodology applied to educational learning environments. They have carried out a research on Open University of Catalonia (UOC) virtual platform and found out the need of including the emotional and more “human” dimension in the design of interfaces to achieve more engaging and enjoyable learning environments, which consequently will enhance the global user experience (GUX) of the UOC learners.

Therefore, it is required to design UX with UCD methodology applied to e-learning systems and interfaces and integrate users´ skills, motivations, feelings and needs in order to succeed in the online learning process.

User experience aspects for developing virtual environments

In the new millennium, many studies have been carried out to understand human experience in virtual environments. These studies help to develop interactive products which are more efficient, easy, beautiful and engage to end-users. So, one of aims is to know which are the qualities involved in UX.

Some authors describe UX because of user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.). (Hassenzahl & Tractinsky, 2006) [16].

From this perspective, there are several aspects to consider when understanding UX to develop a HCI system. On the approach to measure the experiential qualities in HCI systems. Law reviewed the UX literature and selected the 12 most used constructs in UX measurements (see Table 1). These 12 variables at the same time had multi-dimensional aspects. For instance, the construct of "flow" is psychometrically measured by 9 dimensions (e.g., concentration, control, loss of self-consciousness, etc.) (van Schaik & Ling, 2012a) [17]., (Law et Al. 2014) [18]. The Law's questionnaire made over a group of people show a list of UX qualities to be measured and some of them where perceived as measurable and other ones as non-measurable experiential qualities.

Zaharias based their study on educational worlds on UX aspects (Zaharias et Al 2011) [6] using the model proposed by Hassenzahl (Hassenzahl et Al. 2003) [19]. The tool, called *AttrakDiff*, measures the UX in interactive products through a questionnaire centered in attributes from users' perception. This instrument delivers 23 interactive experience items presented in a bipolar scale of qualities and grouped in four main constructs. These constructs are Pragmatic Quality (PQ), related to usability; Hedonic Quality Stimulation (HQS) evaluates user personal growth and the need to

improve personal skills and knowledge; Hedonic Quality Identification (HQI), which focuses on the human need to be perceived by others in a particular way; and Attraction (ATT) which evaluates the global appeal of an interactive system or product.

On the other hand, Rauschenberger created a new questionnaire, which is named User Experience Questionnaire (UEQ) and available on-line, to measure UX in interactive products for optimizing them (Rauschenberger et al. 2013) [20]. They included in the questionnaire, 6 scales with 26 items in total covering the attractiveness (general impression towards the product); the classical usability aspects: efficiency (possibility to use the product fast and efficient and user interface looks organized); perspicuity (easy to understand how to use and to get familiar with the product); dependability (users feeling in control and security and predictability when interacting with the product); and user experience aspects: stimulation (users interest and excitement when using the product and feeling of motivation to use the product again) originality (innovative and creative design of the product and how the product grab users attention).

Table 1: Measure factors in UX in interactive products Questionnaires

Extracted from Rauschenberger et Al. (2013) [20],
Law et Al. (2014) [18] and Hassenzahl et Al. (2003) [19].

Rauschenberger et Al. (2013) [20]	Law et Al (2014) [18]	Hassenzahl et Al. (2003)* [19]
1) Attractiveness.	1)Flow	<i>Pragmatic Quality (PQ):</i>
2) Efficiency	2) Aesthetic/beauty	1)Humaneness
3) Perspicuity.	3) Emotion	2)Simplicity
4) Dependability.	4) Enjoyment	3)Practicality
5) Stimulation.	5) Affect	4)Directness
6) Novelty.	6) Arousal/valence	5)Predictability
	7) Hedonic quality	6)Clearness
	8) Intrinsic motivation	7)Management
	9) Presence	<i>Hedonic Quality Stimulation (HQS):</i>
	10) Engagement	8) Originality
	11) Attractiveness	9)Creativity
	12) Satisfaction	10)Courage
		11) Innovation
		12) Excitement
		13) Challenge
		14) Newness
		<i>Hedonic Quality Identification (HQI)</i>
		15) Integrator
		16) Professionalism
		17) Classiness
		18) Value
		19) Inclusion
		20) Bringing user closer to people
		21) Presentation
		<i>Attraction (ATT)</i>
		22)Beauty
		23)Goodness

(*) These attributes are originally presented by Hassenzahl et al (2003) [19] in opposite pairs of anchors (e.g. Technical-Human, Complicated-

Simple, ...). In this table attributes are presented as nouns summarizing in one word the main concept to be assessed in interactive UX.

As we can observe, UX includes a multidimensional concept and focusses in human needs and the aspects of beauty, fun, pleasure, and personal growth rather than the value of the product or instrument used, which improves or worsens along the time of use (Kujala, 2011) [2].

Zaharias findings have described four main factors affecting UX in Learning Managements Systems (LMS): Pragmatic Quality, Authentic Learning, Motivation and Engagement and Autonomy and Relatedness. (Zaharias et al, 2016) [21].

Therefore, we find different dimensions to be considered for improving human experience in virtual environments, and most of them are related to affective and experiential qualities, which enrich user interaction within virtual learning systems.

User-centered design methodology for designing e-learning interfaces

UCD methodology has been developed by different researchers highlighting the importance of the person, as the user, in the creation and design, not only objects also environments and/or systems for learning purposes in digital scenarios.

Rapanta and Cantoni considered the learner's perspective as a requirement of quality for designing online learning environments. They understand the e-learning environment as an experience to be co-constructed with the learners. Empathising with the learners is clue to anticipate the learning experience and get closer to them to simplify the design of the on-line experience, (Rapanta & Cantoni, 2014) [22].

So, UCD process is a tool to guarantee the quality of the final product. In Table 2 is shown different stages to perform UCD in

interactive learning systems according several studies on interaction systems design.

Table 2: Stages in UCD process applied to interactive learning systems

Gabbard et Al. (1999) [23]	Gulliksen et Al. (2003) [12]	Garreta-Domingo & Mor (2007) [14]
<ul style="list-style-type: none"> -User task analysis -Expert guidelines-based evaluation. -Formative user-centered evaluation. -Summative comparative evaluations. 	<ul style="list-style-type: none"> -Analyse requirements and user needs. -Design for usability by prototyping. -Evaluate use in context. -Feed-back. Plan the next iteration. 	<ul style="list-style-type: none"> - User requirements. - Iterative design and evaluation of prototypes. - Pilot test groups.
Wallach & Scholz, (2012) [24]	Zaina & Alvaro (2015) [25]	ISO 9241-210 [3]
<ul style="list-style-type: none"> -Scope. -Analyse. -Design. -Validate. -Deliver. 	<ul style="list-style-type: none"> -Research. -Ideation. -Validate. -Evaluating. 	<ul style="list-style-type: none"> -Understanding and specifying the context of use. -Specifying the user requirements. -Producing design solutions. -Evaluating the design.

UCD methodologies presented on Table 2 include the iteration as an intrinsic activity to be performed in all the stages. Iterative design is based in the repetition of the activity within the cycle from different perspectives by moving back and forward into the different stages

with the aim of reaching the best solution. This means that the main phases are not linear in time, but UCD process is carried out in several cycles to improve the results of the designed system.

The authors in Table 2 present similar approaches for designing user-centered interactive interfaces. Gabbard UCD methodology applies to virtual environments to ensure usability and a cost-effective strategy for evaluating and improving user interaction. (Gabbard et Al. ,1999) [23]. With the same focus, improving usability in software development, Gulliksen describes the principles and tools and techniques associated for designing systems with a user-centered approach. Nevertheless, Gulliksen present the real limitations found to put into practice the UCD methodology, partly because some organizations and designers still do not recognize the benefits of involving users in the development process. (Gulliksen et Al., 2003) [12]

In the field of e-learning, we find the applied work of Garreta-Domingo and Mor (2007) [14] to develop a virtual classroom at the University (UOC) with UCD methodology. They focus on the learner experience on the interface, by considering not only usability, but also accessibility for all user needs. They integrate in the design process two end-users: students and professors. The results showed opposites reactions and perceptions from both users towards the same interface re-design. A change of architecture interface of the virtual classroom was considered as a radical change by the students, meanwhile professors perceived it as a small change. Their findings also noticed the benefits of personalising the learner interface, which enhanced the learner sense of control and the learning process. Garreta-Domingo underline the need to develop the Learner-Centered Design (LCD) as a specific methodology that guarantees the “learning” and the good learner experience. (Garreta-Domingo and Mor, 2007) [14]. De Lera also followed this aim by researching UOC user experience and designing a Personal Learning Environments (PLE) that enable learners to have the power on their own learning.(De Lera et al, 2012) [26]. Under this perspective, Zaharias highlighted the PLE as

a new market trend required to design the new generation of LMS. (Zaharias et al, 2016) [21].

Conclusions

Virtual educational research shows the need of fostering good, attractive, engaging, enjoying and emotional experiences to all interactive users when using virtual learning environments. On this pursuit, UX research describes how to measure learners experience to create holistic experiences for all end-users (students and teachers). UCD methodologies bring us a tested method to succeed designing e-learning systems centered on the learners (LCD) and improve the virtual learning environments by enhancing the learner experience.

Further research and applications

Further research on this work is motivated by the benefits to educational agents of developing e-learning platforms with user-centered design methodology. From these findings, there is a field to research for developing and improving methodologies on Learner-Centered Design.

The applications of UX and LCD research in e-Learning fields are applicable to any organisation (public and private) that offers e-Learning to internal and/or external learners to develop personalised learning environments for improving the learning experience and increasing the retention of their users.

ACKNOWLEDGEMENTS

Many thanks to PhD. Teresa Magal-Royo for her professional support and great input made on the present work.

REFERENCES

- [1] P. Zaharias; B. Bmehlenbacher. "Exploring User Experience (UX) in virtual learning environments". *International Journal of Human-Computer Studies*, vol. 70, no 7, pp. 475-477, 2012.
- [2] S. Kujala, V. Roto, K. Väänänen-Vainio-Mattila, E. Karapanos, & A. Sinnelä. "UX Curve: A method for evaluating long-term user experience". *Interacting with Computers*, vol. 23, no 5, pp. 473-483, 2011.
- [3] International Standardization Organization (ISO) Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO). Switzerland, 2009.
- [4] D. Squires. "Usability and educational software design: special issue of interacting with computers". *Interacting with Computers*, vol. 11, no 5, pp. 463-466, 1999.
- [5] P. Zaharias. (2004). "Usability and e-learning: the road towards integration". *eLearn Magazine*, vol. 2004, no 6, 2004.
- [6] P. Zaharias, M. Belk, P. Germanakos & G. Samaras. "User Experience in Educational Virtual Worlds". CHI 2011, 2011.
- [7] D. A. Norman & S. W. Draper (Eds.) *User centered system design: New perspectives on human-computer interaction*. Hillsdale, NJ, Lawrence Erlbaum Associates. 1986
- [8] B. Adelson. "The emotional design of everyday things: The 2006 Benjamin Franklin Medal in computer and cognitive science presented to Donald A. Norman". *Journal of the Franklin Institute*, vol. 347, no 4, pp. 681-687. 2010
- [9] L. Damodaran. "User involvement in the systems design process-a practical guide for users", *Behaviour & information technology*, vol. 15, no 6, pp. 363-377, 1996.
- [10] C. Courage & K. Baxter. *Understanding your users: a practical guide to user requirements: methods, tools, and techniques*. Gulf Professional Publishing, 2005.
- [11] T. Brown & B. Katz "Change by design". *Journal of product innovation management*, vol. 28, no 3, pp. 381-383, 2011.
- [12] J. Gulliksen, B. Göransson, I. Boivie, S. Blomkvist, J. Persson & A. Cajander. "Key principles for user-centred systems design". *Behaviour and Information Technology*, 22 vol., no 6, pp. 397-409, 2003.
- [13] C. Gená. "A user-centered approach for adaptive systems evaluation". In 5th Workshop on User-Centred Design and Evaluation of Adaptive Systems, pp. 20-23, 2006.
- [14] M. Garreta-Domingo & E. Mor, E. "User centered design in e-learning environments: from usability to learner experience". *learning*, vol. 11, no 2, p. 16, 2007.

- [15] E. de Lera, M. Almirall, L. Valverde & M. Gisbert. "Improving user experience in e-learning, the case of the open university of catalonia". In *International Conference of Design, User Experience, and Usability*, pp. 180-188, Springer Berlin Heidelberg, 2013.
- [16] M. Hassenzahl, N. Tractinsky. "User experience-a research agenda", *Behaviour & information technology*, vol. 25, no 2, pp. 91-97, 2006.
- [17] P. Van Schaik, J. Ling. "An experimental analysis of experiential and cognitive variables in web navigation", *Human-Computer Interaction*, vol. 27, no 3, pp. 199-234, 2012
- [18] E.L.C. LAW, P. Van Schaik, V. Roto. "Attitudes towards user experience (UX) measurement", *International Journal of Human-Computer Studies*, vol. 72, no 6, pp. 526-541, 2014
- [19] M. Hassenzahl; M. Burmester, F. Koller. "AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality". *Mensch & Computer*. pp. 187-196, 2003
- [20] M. Rauschenberger, M. Schrepp, M. P. Cota, S. Olschner, & J. Thomaschewski, "Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (ueq). example: spanish language version". *IJIMAI*, vol. 2, no 1, pp. 39-45, 2013.
- [21] P. Zaharias & C. Pappas. "Quality Management of Learning Management Systems: A User Experience Perspective". *Current Issues in Emerging eLearning*, vol. 3, no 1, p. 5, 2016.
- [22] C. Rapanta, L. Cantoni. "Being in the users' shoes: Anticipating experience while designing online courses". *British Journal of Educational Technology*, vol. 45, no 5, pp. 765-777, 2014
- [23] J. L. Gabbard, D. Hix & J. E. Swan. "User-centered design and evaluation of virtual environments". *IEEE computer Graphics and Applications*, vol. 19, no 6, pp. 51-59. 1999.
- [24] D. Wallach & S.C. Scholz. "User-centered design: why and how to put users first in software development". In *Software for people*, pp. 11-38. Springer Berlin Heidelberg. 2012.
- [25] L. A. Zaina & A. Álvaro. "A design methodology for user-centered innovation in the software development area". *Journal of Systems and Software*, no 110, pp. 155-177, 2015.
- [26] E. de Lera, M. Almirall, C. Fernández & M. Gisbert. "Designing innovative open spaces for learning: the case of the Open University of Catalunya". *Procedia-Social and Behavioral Sciences*, no 55, pp. 1000-1008. 2012.

2.2. Publication 2 | INTEGRACIÓN DEL “USER EXPERIENCE QUESTIONNAIRE SHORT” EN MOOCS UPV

Sanchis Font, R., Castro Bleda, M. J., Turró Ribalta, C., & Despujol Zabala, I. (2018, September). In IN-RED 2018. IV Congreso Nacional de Innovación Educativa y Docencia en Red (pp. 393-404). Editorial Universitat Politècnica de València.

DOI: [10.4995/INRED2018.2018.8840](https://doi.org/10.4995/INRED2018.2018.8840)

Publication 2 presents the adapted integration of the User Experience Questionnaire Short (UEQ-S) (Schrepp et al., 2017) into the Massive Online Open Courses (MOOCs) of the Universitat Politècnica de València (UPV). Thus, the integration of the UEQ-S in the student evaluation questionnaire will allow us to obtain new data to investigate the interaction with the platform in order to enhance and improve the future of the learning experience in UPV MOOCs.

Integración del “User Experience Questionnaire Short” en MOOCs UPV*

Rosario Sanchis-Font (1), María José Castro-Bleda (2), Carlos Turró-Ribalta e Ignacio Despujol-Zabala (3)

- (1) Escuela Doctorado, Universitat Politècnica de València, rosanfon@doctor.upv.es
- (2) Dpto. Sistemas Informáticos y Computación, Universitat Politècnica de València, mcastro@dsic.upv.es
- (3) Área de Sistemas de la Información y las Comunicaciones, Universitat Politècnica de València, turro@cc.upv.es, ndespujol@asic.upv.es

Abstract

The assessment of the learning experience of Massive Open Online Courses (MOOCs) at Universitat Politècnica de València (UPV) is currently performed through a questionnaire which only evaluates the pedagogical methodology and the functionality of the platforms edX and UPVx, without registering the pleasure of the learning experience. This paper shows the integration of User eXperience (UX) perspective in this questionnaire. User perception on usability and pleasure of using the e-learning platform will be measured in order to achieve a deeper knowledge of users preferences with the ultimate goal to upgrade the UPV MOOCs. To this end, the short version of the validated UX questionnaire of Thomaschewski (2017) “User Experience Questionnaire Short” (UEQ-S) has been adapted to Spanish for UPV MOOC users in order to retrieve answers about joy of use, attractiveness, efficiency and usability of UPV MOOC platforms. In conclusion, the integration of the adapted UEQ-S within the current UPV MOOC questionnaire will provide us new data for further research on the interaction within the platform. The aim is to enhance and improve the future learning experience in UPV MOOCs in order to position them as the best global massive open online courses and UPV as the top world-wide reference in MOOCs.

Keywords: *User experience, UX, MOOC, usability, pleasure of interaction, human computer interaction, interactive experience, e-learning perception, interface design evaluation, on-line platforms.*

Resumen

Actualmente la evaluación de la experiencia de los Massive Open Online Course (MOOCs) de la Universitat Politècnica de València (UPV) se realiza a través de una encuesta que valora únicamente la metodología pedagógica del curso y la funcionalidad de las plataformas edX y UPVx, sin registrar el placer de la experiencia de aprendizaje del usuario. Con esta propuesta se evaluará la formación on-line por primera vez en la UPV con ítems que valoren la experiencia de la interacción o User eXperience (UX). Esto significa conocer la percepción del usuario sobre la usabilidad y placer de uso de la plataforma con el fin de mejorar los UPV MOOCs. Para ello, se ha integrado en la encuesta sobre los MOOC UPV el modelo breve de cuestionario de

Thomaschewski (2017) "User Experience Questionnaire Short" (UEQ-S). Esta encuesta se ha adaptado al español y al usuario de la encuesta MOOC UPV, con el objeto de obtener valores sobre el placer de uso, atractivo, eficiencia, estimulación y usabilidad de la plataforma de e-learning. Así, la integración del UEQ-S en el cuestionario MOOC UPV nos permitirá obtener nuevos datos para investigar sobre la interacción con la plataforma con el fin de potenciar y mejorar el futuro de la experiencia de aprendizaje en los MOOCs UPV y posicionar a la UPV como referente en formación on-line abierta masiva.

Keywords: *experiencia usuario, MOOC, usabilidad, placer de interacción, plataformas on-line, percepción e-learning, UX, user experience, cuestionario aprendizaje interactivo, diseño de interfaces.*

1. Introducción

La incipiente demanda de la formación on-line por un público masivo y global ha originado la producción de Massive Open Online Courses (MOOCs) en universidades de todo el mundo. De esta manera, desde 2012 los cursos abiertos a un público ilimitado a través de plataformas en internet como edX, Udacity, FutureLearn o Coursera están experimentando un crecimiento exponencial (Espada y col. 2014). En este ámbito de educación abierta masiva on-line, la Universitat Politècnica de València (UPV) actualmente ofrece los MOOCs a través de las plataformas edX y UPVx. Atendiendo a la clasificación MOOC (2014) (Espada y col. 2014), los MOOCs UPV son "xMOOCs", puesto que se generan en el contexto de la universidad y siguen una estructura y metodología similar a otros cursos universitarios, siendo además impartidos por miembros de la propia UPV. Actualmente la UPV se posiciona como la quinta universidad a nivel mundial en número de cursos realizados en edX (la plataforma de MOOC impulsada por Harvard y el MIT), con más de 60 cursos y más de 200.000 alumnos (Noticia UPV: Top 5 mundial en MOOC La Universitat Politècnica de València y la plataforma edX, impulsada por MIT y Harvard, renuevan su colaboración 2018). En 2018, la UPV se ha convertido en la primera universidad de habla hispana en superar el millón de inscripciones en edX (Noticia UPV: La UPV, primera universidad de habla hispana en superar el millón de inscripciones en edX, la plataforma de MOOC impulsada por Harvard y el MIT 2018).

Los MOOCs UPV son cursos on-line abiertos a toda la población mundial, y fundamentalmente de habla hispana, donde cientos de miles de estudiantes acceden a la plataforma de aprendizaje con diversidad de habilidades, de edad, de culturas, de procedencia y de lugares físicos de acceso, en diversos momentos y con diferencias en tiempo y en comportamiento de navegación (Sanchis-Font y col. 2017). Por ello, se requiere en estos entornos interactivos un diseño de interfaces de aprendizaje que incluyan las experiencias, motivaciones, sentimientos y necesidades de todos los usuarios con el fin de llevar a cabo el proceso de aprendizaje con éxito a través de los cursos on-line. Para ello, es prioritario conocer la experiencia de usuario o "User eXperience" (UX) de los usuarios de entornos interactivos. El concepto UX es multidimensional y centrado en las necesidades humanas y aspectos de belleza,

diversión, placer y crecimiento personal que se experimentan en la interacción humana con la computadora (Zaharias y Mehlenbacher 2012).

El concepto UX es una ampliación del concepto de usabilidad, que se define como la cualidad de facilidad y satisfacción de uso en un contexto determinado (Bevan 2001). La disciplina UX y sus aplicaciones miden no sólo las cualidades de uso, sino que integran todas las cualidades experienciales (emociones, creencias, comportamientos, ...) en sistemas interactivos entre ordenador-persona, además de las características de usabilidad. Desde la perspectiva de la interacción humana con la computadora, la normativa ISO 9241-210:2010 (2009) describe el UX como todas las emociones, creencias, preferencias, percepciones, respuestas físicas y psicológicas, comportamientos y logros que ocurren antes, durante y después del uso.

Con el fin de mejorar esta experiencia digital de aprendizaje masivo, los productores de MOOCs y universidades utilizan herramientas de evaluación de la experiencia de sus usuarios pero desde la perspectiva de la usabilidad (Espada y col. 2014). Sin embargo, es cada vez más necesario evaluar el aspecto de placer y diversión en el aprendizaje en línea en estas plataformas MOOC, a las cuales acceden usuarios diversos y desde diferentes dispositivos tecnológicos. En este proyecto de investigación se presenta la integración de un cuestionario UX, el “User Experience Questionnaire Short” (UEQS) (Schrepp, Hinderks y Thomaschewski 2017) en la actual encuesta que realiza la UPV en sus MOOCs con el fin último de potenciar el placer de la experiencia de interacción y aprendizaje en las plataformas edX y UPVx.

2. Objetivos

El objetivo general de este trabajo es la integración del cuestionario validado “User Experience Questionnaire Short” (UEQ-S) (Schrepp, Hinderks y Thomaschewski 2017) para obtener una rápida evaluación de la experiencia del usuario en la interacción con las plataformas MOOC UPV de los usuarios durante 6 meses aproximadamente (desde abril a septiembre 2018). Con la versión breve del cuestionario se pretende medir el impacto en el usuario del entorno MOOC UPV, valorando así la estética y funcionalidad de la plataforma desde las 6 escalas y comparando los valores con respecto a otros productos interactivos. En concreto, los objetivos específicos son:

1. Integrar en la plataforma en encuestas UPV “Limesurvey” (<https://www.limesurvey.org/es/>), el cuestionario UEQ-S en el actual modelo de cuestionario de evaluación del usuario MOOC UPV con el fin que nos permita recabar los datos sobre la percepción del usuario de las plataformas MOOC UPV de manera fácil y rápida desde las siguientes escalas con los valores del UEQ-S: atracción, claridad, eficiencia, fiabilidad, motivación y novedad.
2. Obtener el registro de datos del máximo de usuarios de todos los cursos MOOC UPV desde abril hasta septiembre 2018, con el fin de poder generar gráficas y comparativas por tipos de usuarios (ej. hombres, mujeres, etc), tipos de cursos, ediciones de un mismo curso y comparativas con respecto a otros productos interactivos del mercado con los valores del UEQ (informe benchmark) (Laugwitz, Held y Schrepp 2008).

3. Testar la plataforma MOOC UPV en base al diseño de la experiencia del usuario (UX) mediante el cuestionario validado científicamente UEQ-S.

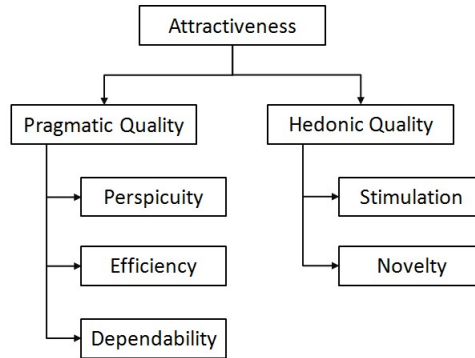


Fig. 1: Estructura de las 6 escalas valoradas en el UEQ-S: atracción, claridad, eficiencia, fiabilidad, motivación y novedad (Rauschenberger y col. 2013).

3. Desarrollo de la innovación

3.1 ¿Qué es UEQ-S?

El cuestionario UEQ-S (Schrepp, Hinderks y Thomaschewski 2017) es la versión breve del cuestionario “User Experience Questionnaire” (UEQ) (Rauschenberger y col. 2013), que evalúa el impacto y percepción del usuario sobre las propiedades de un producto interactivo. Ambos cuestionarios evalúan las cualidades de “eXperiencia del Usuario”, disciplina también conocida como “User eXperience” o UX, la cual valora no solo la usabilidad del producto interactivo, sino que además se centra en su atractivo y el placer de uso percibido por el usuario.

Con esta finalidad, el cuestionario UEQ-S mide dos grandes grupos de cualidades sobre la experiencia de uso del producto, plataforma o entorno interactivo (pragmático y hedónico) organizadas en 6 escalas: atracción, claridad, eficiencia, fiabilidad, motivación y novedad (véase Figura 1).

3.2 ¿Qué ítems valora UEQ-S?

Ya hemos comentado que el cuestionario UEQ-S es la versión reducida del cuestionario UEQ (Rauschenberger y col. 2013). El cuestionario original UEQ registra 26 ítems para evaluar esas mismas 6 escalas. En la versión reducida UEQ-S los ítems 1 a 4 evalúan la cualidad pragmática (usabilidad) y los ítems 5 a 8 miden la cualidad hedónica (placer de uso) como ilustra la Figura 2, versión inglesa, y la Figura 3, versión española. Cada uno de los ítems se valora con una puntuación de 1 a 7 según la escala likert (Likert 1932).

Cualidad Prágmatica	1	obstructive	○ ○ ○ ○ ○ ○ ○ ○	supportive
	2	complicated	○ ○ ○ ○ ○ ○ ○ ○	easy
	3	inefficient	○ ○ ○ ○ ○ ○ ○ ○	efficient
	4	confusing	○ ○ ○ ○ ○ ○ ○ ○	clear
Cualidad Hedónica	5	boring	○ ○ ○ ○ ○ ○ ○ ○	exciting
	6	not interesting	○ ○ ○ ○ ○ ○ ○ ○	interesting
	7	conventional	○ ○ ○ ○ ○ ○ ○ ○	inventive
	8	usual	○ ○ ○ ○ ○ ○ ○ ○	leading edge

Fig. 2: Ítems y cualidades de la versión inglesa del cuestionario UEQ-S (Schrepp, Hinderks y Thomaschewski 2017) con los ítems destacados en dos colores según la cualidad que valoran (azul: cualidad pragmática; naranja: cualidad hedónica).

Cualidad Prágmatica	1	obstruivo	○ ○ ○ ○ ○ ○ ○ ○	impulsor de apoyo
	2	complicado	○ ○ ○ ○ ○ ○ ○ ○	fácil
	3	ineficiente	○ ○ ○ ○ ○ ○ ○ ○	eficiente
	4	confuso	○ ○ ○ ○ ○ ○ ○ ○	claro
Cualidad Hedónica	5	aburrido	○ ○ ○ ○ ○ ○ ○ ○	emocionante
	6	no interesante	○ ○ ○ ○ ○ ○ ○ ○	interesante
	7	convencional	○ ○ ○ ○ ○ ○ ○ ○	original
	8	convencional	○ ○ ○ ○ ○ ○ ○ ○	novedoso

Fig. 3: Ítems y cualidades en español del cuestionario UEQ-S, extraídos de la versión española del UEQ (Rauschenberger y col. 2013), con los ítems destacados en dos colores según la cualidad que valoran (azul: cualidad pragmática; naranja: cualidad hedónica).

En 2017, R. Sanchis-Font realizó la adaptación de determinados ítems a un castellano más comprensible para el usuario hispanohablante y de entornos interactivos de educación superior on-line de universidades españolas. Esta adaptación se realiza para el cuestionario UEQ original de 26 ítems para la investigación en UX del entorno de los másteres oficiales on-line de la Universitat de València desarrollados desde Fundación IVI. Así, el UEQ se adapta en lenguaje con la previsión de poder aplicarlo en los productos interactivos on-line de otras universidades españolas. De estos ítems algunos pertenecen al cuestionario UEQ-S, tal y como se observa en la Figura 4.

Las herramientas de medición y valores que provee el cuestionario UEQ-S son las mismas que las del cuestionario largo UEQ (Laugwitz, Held y Schrepp 2008). En este sentido, los autores facilitan herramientas de análisis validadas para realizar una evaluación benchmark que integra los valores evaluados de 246 productos interactivos con un total de 9905 participantes (Rauschenberger y col. 2013).

	Ítem opuesto (izquierda)	Valores de la escala likert	Ítem opuesto (derecha)
1	entorpecedor	1 2 3 4 5 6 7	ofrece ayuda
2	complicado	1 2 3 4 5 6 7	fácil
3	ineficiente	1 2 3 4 5 6 7	eficiente
4	claro	1 2 3 4 5 6 7	confuso
5	aburrido	1 2 3 4 5 6 7	emocionante
6	interesante	1 2 3 4 5 6 7	nada interesante
7	innovador	1 2 3 4 5 6 7	convencional
8	tradicional	1 2 3 4 5 6 7	novedoso

Fig. 4: Adaptación propia del cuestionario UEQ-S para usuarios hispanohablantes de universidades españolas. Los ítems de pares opuestos están ordenados de izquierda a derecha según la escala likert ascendente y con los valores asignados por los autores del UEQ-S.

3.3 Actual cuestionario sobre la experiencia en los MOOCs UPV

La educación abierta de la UPV en formato MOOC realiza una encuesta a sus alumnos/as al finalizar el curso para conocer un poco más sobre su experiencia en la realización del curso y poder mejorar la experiencia de próximas ediciones. El cuestionario que ofrece los MOOC UPV son 12 preguntas sobre el curso, 5 sobre la plataforma y una sobre los módulos.

En las Figuras 5 a 7 se muestran las pantallas de visualización desde el ordenador de un usuario realizando el cuestionario MOOC UPV a fecha de 23 enero 2018 (Encuesta MOOC UPV al curso edX Liderazgo para mandos intermedios 2018). Estas cuatro pantallas impresas pertenecen al cuestionario del curso MOOC UPV “Liderazgo para mandos intermedios” realizado a través de la plataforma edX.

Fig. 5: Cuestionario MOOC UPV (Pantalla 1).

EDX: Liderazgo para mandos intermedios

0% 100%

Cuestionario

Satisfacción con el curso
En esta sección le preguntamos su opinión sobre el contenido del curso en 12 preguntas.

* ¿Por qué ha inscrito en el curso?

	Completamente en desacuerdo	En desacuerdo	Neutral	De acuerdo	Completamente de acuerdo
Para completar mis estudios académicos	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Para desarrollar una habilidad útil en mi trabajo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Para completar mi currículum vitae	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Por el prestigio de la Universidad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Por el placer de aprender	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

* ¿Has completado el curso?
Seleccione una de las siguientes opciones

¿Cómo was el nivel del curso en relación a tus conocimientos previos?
Seleccione una de las siguientes opciones

¿Cómo valoras el sistema de aprendizaje utilizado?
Seleccione una de las siguientes opciones

La frecuencia de publicación y el ritmo de aprendizaje del curso me han resultado:
Seleccione una de las siguientes opciones

La duración de los vídeos me ha parecido:
Seleccione una de las siguientes opciones

El contenido de los vídeos me ha parecido:
Seleccione una de las siguientes opciones

Las endormas y actividades me han parecido:
Seleccione una de las siguientes opciones

Las opciones de resolución de dudas me han parecido:
Seleccione una de las siguientes opciones

En general con el curso:
Seleccione una de las siguientes opciones

Me expectativas con el curso han quedado satisfechas:
 1 2 3 4 5 Sin respuesta

A la semana le ha dedicado al curso:
Seleccione una de las siguientes opciones

Fig. 6: Cuestionario MOOC UPV (Pantalla 2).

EDX: Liderazgo para mandos intermedios

0% 100%

Castellano ▼

Satisfacción con la plataforma
En esta sección te preguntamos sobre tu experiencia con la plataforma en 5 preguntas

• **Creo que la velocidad de respuesta de la plataforma ha sido:**
Seleccione una de las siguientes opciones

Por favor escoja... ▼

• **¿Has tenido algún problema con la plataforma?**
Seleccione una de las siguientes opciones

Por favor escoja... ▼

• **Para mi utilizar la plataforma (cambiar de unidad, acceder a los exámenes, cambiar de lección o actividad, ver los videos, acceder al perfil, etc) ha sido:**
Seleccione una de las siguientes opciones

Por favor escoja... ▼

• **Los navegadores que he utilizado son:**
Marque las opciones que correspondan

Chrome
 Firefox
 Internet Explorer
 Safari en Mac
 Safari en IPAD
 Iphone
 Navegador Android
 Opera
 Otros:

¿Qué mejorarias o añadirías a la plataforma?

Fig. 7: Cuestionario MOOC UPV (Pantalla 3).

3.4 Integración del cuestionario UX en las plataformas MOOC UPV

Para la integración del cuestionario UX en la sección “sobre la plataforma” del cuestionario MOOC UPV se han adoptado las siguientes propuestas:

1. Mantener las dos últimas preguntas ya existentes: sobre los navegadores y comentarios de mejora.
2. Integrar los 8 pares-ítem del UEQ-S en la sección de preguntas sobre la plataforma, la cual se encuentran en la pantalla 3 (Figura 7).
3. Después valorar con detenimiento aquellas preguntas que el actual cuestionario de MOOC UPV presenta y las cuales encontramos de manera similar desde el cuestionario UEQ-S, se propone eliminar las primeras 3 preguntas de la sección de la plataforma (pantalla 3, Figura 7). Las preguntas que podríamos evaluar desde el cuestionario UEQ-S son:

Creo que la velocidad de respuesta de la plataforma ha sido...

(Esta pregunta estaría incluida en el par “ineficiente-eficiente”)

¿Has tenido algún problema con la plataforma?

(Esta pregunta estaría incluida en los pares de ítems en color azul -escalas pragmáticas-)

Para mí utilizar la plataforma (cambiar de unidad, acceder a los exámenes, cambiar de lección o actividad, ver los vídeos, acceder al perfil, etc) ha sido. . .

(Esta pregunta estaría incluida en los pares de ítems en color azul -escalas pragmáticas-)

4. En el caso de que “edX Analytics” nos impida conocer los dispositivos de acceso al curso MOOC UPV por el usuario, se incluirá la siguiente pregunta:

¿Con qué dispositivos accedes al curso? (puedes seleccionar más de 1): PC / Tablet / Móvil / Otros

5. El cuestionario UEQ-S se ha adaptado para el usuario MOOC UPV del siguiente modo (véase Figura 8). La Figura 9 muestra la interfaz gráfica del cuestionario:

- a) Semánticamente para que todos los ítems sean adjetivos y de género femenino (acorde con el sustantivo calificado: “la plataforma es...”).
- b) Semánticamente los ítems se diferencien más entre sí, y se ajusten más al aspecto valorado por el usuario MOOC UPV de habla hispana (por ejemplo, el ítem “impulsor de apoyo” se modificó en la versión larga para actuales investigaciones de plataformas de estudios de posgrado a “ofrece ayuda”, en la propuesta para MOOCs UPV se ha resumido en una frase adjetivo, “de apoyo”).
- c) Para cada ítem semánticamente positivo se incluyen ejemplos para facilitar al usuario su comprensión. Estos ejemplos son más genéricos por dos motivos, uno de ellos es para no influir demasiado en la decisión del usuario en el momento de la respuesta. Otra de las razones atiende a que cada usuario es diferente y su percepción también. El ejemplo genérico permitirá que, dentro esa diversidad de usuarios, que cada encuestado pueda ver recogida todas sus experiencias interactivas y asignarles un valor percibido. Por eso también los ejemplos se presentan con verbos impersonales.

No obstante, se propone dejar el cuestionario UX lo más genérico posible en esta primera fase de recogida de datos (de abril a septiembre 2018) ya que es susceptible de ajuste para una segunda recogida. Tras los resultados preliminares se prevé poder obtener información sobre aspectos valorados positivamente y negativamente de la plataforma MOOC UPV y poder centrar la investigación en aquellos aspectos de interés según tipo de usuario (por ejemplo: usuario hombre y usuario mujer), de manera que podamos detallar ejemplos e incluso incluir nuevas preguntas sobre su experiencia interactiva.

- d) Además, estéticamente para no confundir al usuario en el método de presentación de las respuestas se presentarán todos los ítems semánticos con connotaciones positivas a la derecha y todos los negativos a la izquierda. Este cambio implica que algunos pares de ítems inviertan el orden de visualización asignado por los autores del UEQ-S.

Por favor, selecciona la opción más próxima a tu experiencia en la plataforma del curso MOOC UPV.			
<i>"La plataforma MOOC UPV es..."</i>			
1	<i>entorpecedora</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>1 2 3 4 5 6 7</p>	<p>de apoyo</p> <p>Por ejemplo: la plataforma ofrece funciones y contenidos que nos permiten tener el control de navegación y de interacción, la plataforma permite la opción de poder adaptarla a nuestras necesidades y tipo de uso, ...</p>
2	<i>complicada</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>1 2 3 4 5 6 7</p>	<p>fácil</p> <p>Por ejemplo: Es fácil aprender el funcionamiento de navegación de la plataforma, fácil conocer el acceso a los contenidos y como interactuar en ella, ...</p>
3	<i>ineficiente</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>1 2 3 4 5 6 7</p>	<p>eficiente</p> <p>Por ejemplo: la velocidad de respuesta de la plataforma ha sido rápida, las pantallas muestran la información y los accesos de manera organizada, ...</p>
4	<i>confusa</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>7 6 5 4 3 2 1</p>	<p>clara</p> <p>Por ejemplo: La navegación por la plataforma está muy clara y es bastante transparente, es fácil reconocer los accesos y complementos en las</p>
5	<i>aburrida</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>1 2 3 4 5 6 7</p>	<p>entretenida</p> <p>Por ejemplo: ha sido entretenido y ha sido un disfrute navegar e interactuar en la plataforma, etc.</p>
6	<i>nada interesante</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>1 2 3 4 5 6 7</p>	<p>interesante</p> <p>Por ejemplo: las opciones de interacción y de navegación han mantenido constantemente el interés y las ganas de continuar, ...</p>
7	<i>convencional</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>7 6 5 4 3 2 1</p>	<p>innovadora</p> <p>Por ejemplo: la plataforma presenta un aspecto y ofrece una navegación diferente al resto de plataformas e-learning, ...</p>
8	<i>tradicional</i>	<p>○ ○ ○ ○ ○ ○ ○</p> <p>1 2 3 4 5 6 7</p>	<p>novedosa</p> <p>Por ejemplo: la plataforma presenta un aspecto y ofrece una navegación de última generación, ...</p>

Fig. 8: Cuestionario UEQ-S adaptado para MOOC UPV.

EDX: Buscar en Internet

0% 100%

Castellano ▼

Satisfacción con la plataforma
En esta sección te preguntamos sobre tu experiencia con la plataforma en 5 preguntas

**Por favor, selecciona la opción más próxima a tu experiencia en la plataforma utilizada para seguir el curso:
"La plataforma MOOC es..."**

	1	2	3	4	5	6	7
Entorpecedora (1) a De apoyo (7) La plataforma es de apoyo si ofrece funciones y contenidos que nos permiten tener el control de navegación y de interacción, la plataforma permite la opción de poder adaptarla a nuestras necesidades y tipo de uso,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Complicada (1) a Fácil (7) La plataforma es fácil si es fácil aprender el funcionamiento de navegación, fácil conocer el acceso a los contenidos, fácil interactuar con ella, ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ineficiente (1) a Eficiente (7) La plataforma es eficiente si su velocidad de respuesta es rápida, las pantallas muestran la información y los accesos de manera organizada, ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Confusa (1) a Clara (7) La plataforma es clara si la navegación por la plataforma es clara y transparente, es fácil reconocer los accesos y complementos en las diferentes pantallas, ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aburrida (1) a Entretenida (7) La plataforma es entretenida si ha sido entretenido y un placer navegar e interactuar con ella, ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	1	2	3	4	5	6	7
<p>Nada interesante (1) a Interesante (7) La plataforma es interesante si las herramientas de interacción y de navegación han mantenido el interés y las ganas de continuar, ...</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>Convencional (1) a Innovadora (7) La plataforma es innovadora si presenta un aspecto innovador y ofrece una navegación diferente al resto de plataformas e-learning, ...</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<p>Tradicional (1) a Novedosa (7) La plataforma es novedosa si presenta un aspecto y ofrece una navegación de última generación, ...</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

? Para cada opción tienes que seleccionar un número de 1 a 7 que describa tu experiencia entre dos descripciones opuestas. Por ejemplo, para el ítem 1 selecciona 1 si la plataforma te ha parecido "entorpecedora" y 7 si te ha parecido "de apoyo", o cualquier valor para indicar el punto entre estos dos calificadores en el que crees que está la plataforma (por ejemplo, si seleccionas 4 sería neutro).

*** Los navegadores que he utilizado son:**
Marque las opciones que correspondan

Chrome
 Firefox
 Internet Explorer
 Safari en Mac
 Safari en IPAD
 Iphone
 Navegador Android
 Opera
 Otro:

¿Con qué dispositivos accedes al curso? (puedes seleccionar más de uno)
Marque las opciones que correspondan

PC
 Tablet
 Móvil
 Otro:

¿Qué mejorarías o añadirías a la plataforma?

Fig. 9: Interfaz del cuestionario UEQ-S integrado en el cuestionario MOOC UPV extraído en Abril 2018 del curso en edX "Buscar en Internet".

4 Conclusiones

Por primera vez la UPV introduce un cuestionario validado científicamente en la evaluación de los MOOC por los estudiantes, y que además evalúa la experiencia de la interacción de los usuarios desde la perspectiva UX. La integración del cuestionario

corto de experiencia del usuario UEQ-S en los MOOC UPV nos permitirá conocer tanto la usabilidad como el placer de uso que experimentan la multitud y diversidad de usuarios de cursos abiertos masivos on-line de la UPV. Los datos obtenidos en las encuestas nos facilitará información sobre aspectos de preferencia y comportamiento de la interacción de los usuarios con el fin último de potenciar experiencias placenteras de aprendizaje online. Los resultados analizados de este cuestionario integrado en los MOOC UPV nos ayudarán a establecer las pautas para desarrollar un diseño de interfaz de experiencias positivas de aprendizaje masivas, en abierto, online y adaptado a todos los usuarios. De esta manera se pretende que todo usuario global de MOOC escoja siempre a la UPV como su primera opción para su formación en abierto y en línea porque ofrece experiencias de aprendizaje online placenteras, atractivas y motivadoras.

Referencias bibliográficas

(2009). Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO). Switzerland.

Bevan, Nielg (2001). "International standards for HCI and usability". En: International Journal of Human-Computer Studies 55.4, págs. 533-552. doi: 10.1006/ijhc.2001.0483.

Encuesta MOOC UPV al curso edX Liderazgo para mandos intermedios (2018). Accedido Ene 23, 2018. <https://encuesta.upv.es/index.php/374466?token=625590&lang=es>.

Espada, J. P. y col. (2014). "Method for analysing the user experience in MOOC platforms". En: 2014 International Symposium on Computers in Education (SIIE). IEEE, págs. 157-162. doi: 10.1109/siie.2014.7017722.

Laugwitz, Bettina, Theo Held y Martin Schrepp (2008). "Construction and Evaluation of a User Experience Questionnaire". En: Lecture Notes in Computer Science. Springer Berlin Heidelberg, págs. 63-76. doi: 10.1007/978-3-540-89350-9_6.

Likert, R. (1932). "A technique for the measurement of attitudes". En: Archives of Psychology 22.140.

Noticia UPV: La UPV, primera universidad de habla hispana en superar el millón de inscripciones en edX, la plataforma de MOOC impulsada por Harvard y el MIT (2018). Accedido May 24, 2018. <https://www.upv.es/noticias-upv/noticia10053-cursos-masivos-es.html>.

Noticia UPV: Top 5 mundial en MOOC La Universitat Politècnica de València y la plataforma edX, impulsada por MIT y Harvard, renuevan su colaboración (2018). Accedido Feb 2, 2018. <https://www.upv.es/noticias-upv/noticia-9788top-5-mundial-e-es.html>.

Rauschenberger, Maria y col. (2013). "Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ).

Example: Spanish Language Version”. En: *International Journal of Interactive Multimedia and Artificial Intelligence* 2.1, págs. 39-45. doi: 10.9781/ijimai. 2013.215.

Sanchis-Font, R. y col. (2017). “Improving the virtual learning experience: Usercentered design in e-learning”. En: *INTED2017 Proceedings. 11th International Technology, Education and Development Conference*. Valencia, Spain: IATED, págs. 9902-9907. doi: 10.21125/inted.2017.2356.

Schrepp, Martin, Andreas Hinderks y Jorg Thomaschewski (2017). “Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S)”. En: *International Journal of Interactive Multimedia and Artificial Intelligence* 4.6, págs. 103-108. doi: 10.9781/ijimai.2017.09.001.

Zaharias, Panagiotis y Brad Mehlenbacher (2012). “Editorial: Exploring User Experience (UX) in Virtual Learning Environments”. En: *Int. J. Hum.-Comput. Stud.* 70.7, págs. 475-477. doi: 10.1016/j.ijhcs.2012.05.001.

2.3. Publication 3 | APPLYING SENTIMENT ANALYSIS WITH CROSS-DOMAIN MODELS TO EVALUATE USER EXPERIENCE IN VIRTUAL LEARNING ENVIRONMENTS

APPLYING SENTIMENT ANALYSIS WITH CROSS-DOMAIN MODELS TO EVALUATE USER EXPERIENCE IN VIRTUAL LEARNING ENVIRONMENTS




Sanchis-Font, R., Castro-Bleda, M.J., González, JÁ. (2019). In: Rojas, I., Joya, G., Catala, A. (eds) *Advances in Computational Intelligence. IWANN 2019. Lecture Notes in Computer Science (LNCS)*, vol 11506. Springer, Cham.

DOI: https://doi.org/10.1007/978-3-030-20521-8_50

In Publication 3 we test the application of machine learning sentiment analysis tools to the free text comments of 133 users of graduate courses and MOOCs (37 in English and 96 in Spanish) in order to obtain their polarity (positive, negative or neutral). To do so, we use cross-domain models trained with a corpus from different domains, models trained with a corpus from different domains (Twitter posts for each language) and models trained with general domains.

The results will allow us to advance the research in the following work published in greater depth to apply sentiment analysis techniques in order to assess the experience of online university students more accurately.

Applying Sentiment Analysis with Cross-Domain Models to Evaluate User eXperience in Virtual Learning Environments

Rosario Sanchis-Font , Maria Jose Castro-Bleda ^(✉) ,
and José-Ángel González 

Universitat Politècnica de València, Valencia, Spain
rosanfon@doctor.upv.es, mcastro@dsic.upv.es, jgonba2@inf.upv.es

Abstract. Virtual Learning Environments are growing in importance as fast as e-learning is becoming highly demanded by universities and students all over the world. This paper investigates how to automatically evaluate User eXperience in this domain. Two Learning Management Systems have been evaluated, one system is an ad-hoc system called “Conecto” (in Spanish and English languages), and the other one is an open-source Moodle personalized system (in Spanish). We have applied machine learning tools to all the comments given by a total of 133 users (37 English speakers and 96 Spanish speakers) to obtain their polarity (positive, negative, or neutral) using cross-domain models trained with a corpus of a different domain (tweets for each language) and general models for the language. The obtained results are very promising and they give an insight to keep going the research of applying sentiment analysis tools on User eXperience evaluation. This is a pioneering idea to provide a better and accurate understanding on human needs in the interaction with Virtual Learning Environments. The ultimate goal is to develop further tools of automatic feed-back of user perception for designing Virtual Learning Environments centered in user’s emotions, beliefs, preferences, perceptions, responses, behaviors and accomplishments that occur before, during and after the interaction.

Keywords: Machine learning · Sentiment analysis · Polarity · User eXperience · Virtual Learning Environments · Learning Management Systems

1 Introduction

Human Computer Interaction (HCI) tools developers, agents and industry require to focus their interactive systems on end-users in order to design and provide quality systems upon the international standards requirements ISO.

Partially supported by the Spanish MINECO and FEDER funds under project TIN2017-85854-C4-2-R. Work of J. A. González is financed under grant PAID-01-17.

These interactive systems are the “combination of hardware, software and/or services that receives input from, and communicates output to, users” (ISO 9241-20: 2010) [1]. This international standard is related to ergonomics of human system-interaction and human-centered design for interactive systems. It provides requirements and recommendations for human-centered design principles and activities throughout the life cycle of computer-based interactive systems. It is intended to be used by those managing design processes, and is concerned with ways in which both hardware and software components of interactive systems can enhance human-system interaction.

Therefore “User eXperience” (UX) enhances human interaction within the hardware or software components, being the UX concept multidimensional and centered in human needs. This UX concept goes beyond usability, interaction experience and design by involving two main qualities: traditional HCI usability and accessibility balanced with hedonic and affective design [19]. In this perspective, in [6], UX is described as a consequence of a user’s internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organizational/social setting, meaningfulness of the activity, voluntariness of use, etc.). Therefore, these authors conclude that UX is considering three perspectives: emotion and affect of the user, technology and the hedonic instrument and the experiential aspect. As a result, UX includes a multidimensional concept and focuses in human needs and the aspects of beauty, fun, pleasure, and personal growth rather than the value of the product or instrument used [6], which improves or worsens along the time of use [9].

UX has to be considered when designing and redesigning hardware and software applications. In this way, in the last years, UX has been taken into account when designing Virtual Learning Environments (VLEs) [19]. VLEs includes a wide range of technology-enabled learning environments, such as Learning Management Systems (LMSs), computer games or Virtual Worlds.

In order to evaluate UX in VLEs, we have used the validated User Experience Questionnaire (UEQ) [16], addressed to 559 users of biomedical postgraduate studies. Two LMSs have been evaluated using this adapted UEQ: one LMS is an ad-hoc system called “Conecto” (in Spanish and English languages), and the other one is an open-source Moodle personalized system (in Spanish).

We have applied machine learning tools to all user’s comments using a cross-domain model trained with tweets for each language [3] and a general system for text analytics (MeaningCloud [11]). The application of sentiment analysis tools on UX comments will provide a better and accurate understanding on human needs in the interaction with VLE for postgraduate and biomedical online learning. The ultimate goal is to develop further tools of automatic feed-back of user perception for designing user-centered VLE valued by users for its usability, quality and pleasure of use.

This paper is organized as follows. Next Section gives a brief view about the state of the art of the work presented here. Section 3 describes the data collection and preprocessing from the questionnaires. The used models for sentiment analysis are described in Sect. 5. Section 6 presents the experimental results and their analysis. Finally, the conclusions are drawn in the last Section.

2 State of the Art

Sentiment analysis is one of the most active areas in Natural Language Processing since the early 2000s. Concretely, since Pang et al. [14], who addressed the importance of “sentiment classification” for a large number of tasks such as *message filtering, recommender systems or business intelligence applications*. A decade after, until our days, the popularity of the sentiment analysis has been increasing and Deep Learning has consolidated as a well-established alternative to the previous machine learning systems. Thus, Deep Learning is the state of the art in sentiment analysis [2, 4, 8, 18]. Our approach uses state-of-the-art models, neural networks trained with tweets in English and Spanish, as described in Sect. 5.1.

In addition, a large number of commercial products and frameworks have also proliferated to facilitate the development and deployment of sentiment analysis systems based on machine learning, such as Google Cloud [5], IBM Watson [7], Microsoft Text Analytics [13] and MeaningCloud [12]. This kind of products allow us to perform text analytics such as sentiment analysis, in a broad variety of domains and languages in an easy way, obtaining also competitive results. For this reason, besides our neural network models, MeaningCloud models will be used in our work as explained in Sect. 5.2.

But, though the promising results of natural language processing and, in particular, of sentiment analysis, generally speaking, UX evaluation is immature in most applications and, especially, in VLEs. Some work has been done in eCommerce, using natural language processing to improve their UX, for instance, to search products in a more intelligent way, using sentiment analysis to extract insights from the reviews made by the customers on the product or identifying trends and trying to answer best to the customers’ concerns. Several new conferences have recently been launched around these ideas (see, for example, <https://julielab.de/econlp/2019/> or <https://www.aclweb.org/portal/content/first-international-workshop-e-commerce-and-nlp>).

Another research line covered in this paper is the use of cross-domain polarity classification approach, that is, the texts to be classified belong to a different domain from those used in the training phase. Most work have been done within the classic approach, the so-called single-domain polarity classification, which classifies texts in the same domain to which the texts used in the training phase belong to. Due to the lack of training data (only comments from 133 users), we used cross-domain models, those trained with another domain (tweets) and those trained with general data.

32. What kind of user are you?

Student

Teacher/Tutor

Administrator

Other

33. Have you previously used other e-learning platforms?

Yes, I have used previously other e-learning platforms.

No, I have not used previously other e-learning platforms

NR/NK

34. Please, let us know any comments about your experience on the environment of IVI e-learning Master:

By doing this questionnaire you accept the use of your data for scientific purposes. This data is completely confidential and only it will be used for current and future papers, reports and studies that might be produced after processing the information by Fundación IVI and UPV. Please, click on the blue button «Enviar» to send the questionnaire. Many thanks for your contribution and tell us your experience.

Enviar cuestionario

Fig. 1. “Other comments” box from UX questionnaire delivered to English speaker users on Conecto LMS of IVI Foundation Biomedical International Master (edition 2017–18).

3 Experimental Data

The validated User Experience Questionnaire (UEQ) [16] was used in order to automatically evaluate UX in our VLEs. This questionnaire is a list of closed-ended questions, but we added questions concerning to sociodemographics data (age, sex, etc.) and an open field “Other comments” (see Fig. 1 for a screenshot of the questionnaire in one VLE). It is a text entry box to express any comments related to UX in the course, which is an opportunity to get new and more precise information about their experience, not only by close-ended questions.

Two LMSs have been evaluated: one system is an ad-hoc system called “Conecto” (in Spanish and English language)¹, and the other one is an open-source Moodle personalized system (in Spanish)². We have collected data in different editions: 2016–17, 2017–18, January 2018, and April 2018, and at the middle and final term of each course. The UEQ was addressed to 559 users. Only 133 users (37 English speakers and 96 Spanish speakers) filled the “Other comments” box.

We have performed experiments at three different semantic levels of decreasing complexity:

1. *Observation.* We measured the polarity of the whole observation. Each entry is composed by one or more sentences. There were 96 Spanish comments and 37 English observations.

¹ https://postgrado.adeituv.es/es/cursos/salud-7/assisted-reproduction/datos_generales.htm.

² <https://medicinagenomica.com/eugmygo/>.

2. *Sentence*. As an observation from one user can be composed by one or more sentences, we automatically split each observation into sentences, being one sentence the text between points. We got 151 Spanish sentences and 61 English sentences.
3. *Meaningful unit*. Also, we parsed complex statements in a semi-supervised way in order to obtain meaningful units, that is, one sentence or part of one sentence which has own meaning. We obtained in this way 227 Spanish units and 70 English units.

The observation can be composed of more than one sentence, and it is very usual to mix positive and negative comments about different concepts in different sentences, so many comments are tagged as neutral (see Table 1 and some examples in Table 2 to illustrate this idea). This fact hides the intention of the user, which is tagged as neutral when she or he is not, that is the reason we automatically split the original observations into sentences and measuring the polarity of each sentence. Finally, we desired in this study to explore the idea of detecting polarity with cross-domain models in less complex structures, splitting the sentences in meaningful units which are usually only positive or negative.

After this process, all of these units (whole observations, automatic sentences and meaningful units) were manually tagged according its polarity (positive, negative, or neutral). Positive and negative sentiment units were annotated, being tagged as neutral those units without presence of any emotion or feelings (i.e., “No applicable.”) or when the unit provided the same amount of positive and negative feelings (i.e., “Some of the modules were very interesting and valuable but some of them confusing as too genetic details involved.”). Two human taggers did both the parsing of complex statements in meaningful units and the annotation of each unit as positive, negative or neutral. See the total number of units and the class distributions in Table 1. As it can be observed, there are more positive than negative samples. The neutral category decreased from the whole comment (a complex statement) to the meaning unit (usually, with polarity or, less frequently, with lack of sentiment). All these samples were used as test set, and they were automatically labeled by using the proposed models (neural networks and MeaningCloud models). Following, Table 2 gives some examples of tagged observations, sentences and meaningful units.

4 Evaluation Metrics

In order to evaluate the systems with the gold standard, different evaluation metrics were used. Concretely, as defined below, we used Accuracy (Acc , Eq. 1) and macro F_1 (MF_1 , Eq. 3) to reduce the impact of corpus imbalance in the evaluation. Moreover, the F_1 per class (being c the *positive*, *negative*, and *neutral* class in F_1^c , Eq. 2) is shown to observe the behavior of our systems at class level.

$$Acc = \frac{\sum_{c \in C} \sum_{x \in \Omega_c} [f(x) = c]}{|\Omega|} \quad (1)$$

Table 1. Different units extracted from the “Other comments” box, from 133 users (37 English speakers and 96 Spanish speakers).

Type of unit	Language	VLE	Total	Positive	Negative	Neutral
Observation	Spanish	Conecto	31	12 (39%)	8 (26%)	11 (35%)
		Moodle	65	35 (54%)	13 (20%)	17 (26%)
		Total	96	47 (49%)	21 (22%)	28 (29%)
	English	Conecto	37	18 (49%)	8 (22%)	11 (30%)
Sentence	Spanish	Conecto	51	21 (41%)	23 (45%)	7 (14%)
		Moodle	100	49 (49%)	32 (32%)	19 (19%)
		Total	151	70 (46%)	55 (37%)	26 (17%)
	English	Conecto	61	31 (51%)	18 (29%)	12 (20%)
Meaningful unit	Spanish	Conecto	63	31 (49%)	31 (49%)	1 (2%)
		Moodle	164	97 (59%)	59 (36%)	8 (5%)
		Total	227	128 (56%)	90 (40%)	9 (4%)
	English	Conecto	70	37 (53%)	25 (36%)	8 (11%)

Table 2. Examples of tagged observations, sentences and meaningful units, with its polarity.

Unit	Example	Polarity
Observation	<i>Excellent opportunity to learn with our busy routine concerns regarding very low volume of speakers as a very quiet room required even a fan disturbs the volume.</i>	Neutral
Sentence	<i>Excellent opportunity to learn with our busy routine concerns regarding very low volume of speakers as a very quiet room required even a fan disturbs the volume.</i>	Neutral
Meaningful unit	<i>Excellent opportunity to learn with our rou-</i>	Positive
Meaningful unit	<i>tine concerns regarding very low volume of speakers as a very quiet room required even a fan disturbs the volume.</i>	Negative
Observation	<i>Well-organized and structured course. Great study material (articles) but not enough time to read them all. Keep up the good work.</i>	Neutral
Sentence	<i>Well-organized and structured course.</i>	Positive
Sentence	<i>Great study material (articles) but not enough time to read them all.</i>	Neutral
Sentence	<i>Keep up the good work.</i>	Positive
Meaningful unit	<i>Well-organized and structured course.</i>	Positive
Meaningful unit	<i>Great study material (articles)</i>	Positive
Meaningful unit	<i>but not enough time to read them all.</i>	Negative
Meaningful unit	<i>Keep up the good work.</i>	Positive

$$F_1^c = \frac{2 \cdot P_c \cdot R_c}{P_c + R_c} \quad (2)$$

$$MF_1 = \frac{1}{|C|} \sum_{c \in C} F_1^c \quad (3)$$

where Ω is the set of samples, Ω_c are the samples of class c in Ω , $y(x)$ is the prediction of the model f for a given sample x , C is the set of classes, $[\cdot]$ denotes the Iverson bracket, and P_c and R_c are the prediction and recall measure of each class:

$$P_c = \frac{\sum_{x \in \Omega_c} [f(x) = c]}{\sum_{x \in \Omega} [y(x) = c]} \quad R_c = \frac{\sum_{x \in \Omega_c} [f(x) = c]}{|\Omega_c|} \quad (4)$$

5 Polarity Models

Cross-domain models for both Spanish and English are used to address the problem of sentiment analysis on VLEs. On the one hand, Convolutional Neural Networks (CNN) were used to train models for sentiment analysis tasks on Twitter, both in Spanish and English, proposed in national and international competitions [10, 17]. On the other hand, we used the sentiment analysis module provided by the product “Software as a Service” MeaningCloud [12], which acts as a general domain polarity classifier both for English and Spanish.

5.1 Convolutional Neural Networks Models

To determine the polarity of the students’ opinions, we used polarity models based on the use of word embeddings and deep learning. Unfortunately, due to the lack of training data, it was not possible to learn robust models specifically for the task described in this paper. Instead, we used models trained, by our research group, for similar tasks related to the social network Twitter [10, 17] both for English and Spanish. Table 3 shows some details of the corpora used to train the models.

Specifically, our trained models are based on the use of Convolutional Neural Networks. This architecture is inspired by the work described in [8], which has obtained competitive results in text classification tasks such as sentiment analysis or irony detection. Each opinion is represented as a 50×300 matrix where each word of the opinion - up to a maximum of 50 - is represented as a 300-dimensional embedding. Zero padding at the start of the matrix was used for opinions with less than 50 words. We applied several one-dimensional (the width of the filter is constant and equal to the dimension of the embeddings) convolutions with different height filters in order to extract the sequential structure of the text. Subsequently, we applied Global Max Pooling to the feature maps in order to extract the most salient features for each region size. The final decision is carried out by a softmax fully-connected layer. Table 4 shows the performance of the models for the test set of the two tasks [3, 15], along with their 95% confidence

Table 3. Characteristics of the corpora used to train the CNN models (both for English and Spanish).

Task	Set	Total	Positive	Negative	Neutral	None
SemEval 17 (English)	Train	39656	15705 (40%)	6203 (15%)	17748 (45%)	N/A
	Test	12284	2375 (19%)	3972 (32%)	5937 (49%)	N/A
	Total	51940	18080 (35%)	10175 (19%)	23685 (46%)	N/A
TASS 17 (Spanish)	Train	1008	318 (32%)	418 (41%)	133 (13%)	139 (14%)
	Test	1899	642 (34%)	767 (40%)	216 (11%)	274 (15%)
	Total	2907	960 (33%)	1185 (41%)	349 (12%)	413 (14%)

Table 4. Performance of the CNN (in grey) and MeaningCloud (in white) systems on SemEval 2017 Task 4 (English) and TASS 17 (Spanish) for the test set.

Task	Acc		F ₁	F ₁ ^{pos}	F ₁ ^{neg}	F ₁ ^{neu}	F ₁ ^{none}					
English	0.72±0.00	0.52±0.01	0.73	0.52	0.73	0.47	0.75	0.57	0.70	0.53	N/A	
Spanish	0.62±0.02	0.54±0.02	0.47	0.44	0.67	0.64	0.71	0.60	0.08	0.14	0.42	0.37

intervals. Note that for TASS 2017, the distinction between the classes Neutral (*with both positive and negative feelings*) and None (*lack of sentiment*) is made during training and test. However, when the model trained for TASS 2017 is applied to our UX evaluation task, both classes are considered equal. That is, given a test opinion x , $\operatorname{argmax}_y p(y|x) \in \{Neutral, None\} \rightarrow y = Neutral$.

5.2 MeaningCloud Models

MeaningCloud is a Software As A Service product [12] that provides a large number of tools, easy to use and to deploy, for text processing, analytics and text/audio mining, with the aim of facilitating the resolution of natural language processing problems to developers. It includes tools for summarization, topic extraction, language identification and sentiment analysis.

We have used the sentiment analysis module. This module allowed us to use a classifier, trained in a general domain with texts in multiple languages, to determine the global polarity of user opinions on VLEs. Concretely, we use the field *score tag* in the response of the MeaningCloud API, that indicates the global polarity of the text in 6 different levels: strong positive, positive, neutral, negative, strong negative and without sentiment (None). To carry out our experiments, we collapsed the strong sentiments, i.e., strong positive and strong negative are considered as Positive and Negative, respectively. Moreover, the neutral/none classes are fused in only one class (Neutral). Table 4 shows the performance of the MeaningCloud system for the test set of the two tasks used to train the CNN models in order to compare both systems. As it could be expected, performance is much better with the specific trained CNN models than when using general language models provided by MeaningCloud for a specific task.

In addition, the module is also capable of performing sentiment analysis at segment and aspect level. Thus, it is possible to detect words that express polarity and relate them with the objects of such polarity. That is interesting to capture relevant aspects that influence on the students and were not considered during the questionnaire, or to analyze which aspects of a course tend to be more negative or positive for the students.

On the other hand, the module has a series of additional capabilities such as the detection of subjectivity, the “agreement” and the irony, as well as disambiguation utilities of entities to enrich the sentiment analysis at the level of aspects. However, all these extended features are out of the scope of this work and we have planned to approach them in future work.

6 Experimental Results

We applied these two systems to the proposed task which consists of determining the polarity (positive, negative, or neutral) of each unit. One unit, as explained in Sect. 3, can be one observation, one sentence or one meaning unit.

The results obtained with the two systems for each type of segmentation can be seen in Tables 5, 6 and 7, along with their confidence intervals. It is possible to observe that in all cases, the *neutral* class is the worst detected. That means that it is more difficult to detect the absence of feelings (or the same amount of mixed positive and negative feelings) than detect isolated positive or negative feelings in this kind of comments. Also, *pos* is the best classified class almost in all cases, i.e., positive feelings are the easiest to detect.

It is possible to compare the behavior of the systems in Spanish and English (only on the Conecto system). It is striking that F_1^{pos} is better (or equal) in English than in Spanish, while F_1^{neg} is always worse in English than in Spanish. It seems more difficult to classify the negative class in English than in Spanish, whereas it is easier to classify positive and neutral classes.

In general, every metric is higher when detecting feelings for manually extracted semantic units. This is especially noticeable in the English language. The difference is not so marked in the case of whole comments and automatically extracted sentences.

CNN models have a slightly better behavior on the English language than MeaningCloud models (only in absolute values, their 95% confidence intervals overlap). This may be due to the fact that there are much more training samples for English than for Spanish (see Table 3) and therefore the trained model is able to generalize better. On the contrary, for Spanish, the results in terms of *Acc* and F_1 are better with MeaningCloud models. In this case, CNN models have only been trained with 1000 samples (versus 40k samples for English) and the generalization is much worse than using the general models trained for Spanish in the MeaningCloud system.

Table 5. Experiments with the whole comments. The unit is the whole comment, composed by one or more sentences. First column of each evaluation metric (in grey) is obtained with our approach based on CNN, the second column (in white) is obtained with MeaningCloud models.

Lang.	VLE	Tot.	Acc		MF ₁	F ₁ ^{pos}	F ₁ ^{neg}	F ₁ ^{neu}				
Spanish	Conecto	65	0.58±0.12	0.65±0.12	0.44	0.53	0.71	0.80	0.61	0.57	0.00	0.23
	Moodle	31	0.48±0.18	0.58±0.17	0.43	0.55	0.69	0.69	0.59	0.71	0.00	0.25
	Total	96	0.55±0.10	0.62±0.10	0.43	0.55	0.70	0.77	0.60	0.63	0.00	0.24
English	Conecto	37	0.62±0.16	0.57±0.16	0.52	0.44	0.77	0.74	0.50	0.36	0.29	0.24

Table 6. Experiments with automatically extracted sentences from the whole comments. The unit is one sentence, automatically extracted from the whole comment. First column of each evaluation metric (in grey) is obtained with our approach based on CNN, the second column (in white) is obtained with MeaningCloud models.

Lang.	VLE	Tot.	Acc		MF ₁	F ₁ ^{pos}	F ₁ ^{neg}	F ₁ ^{neu}				
Spanish	Conecto	100	0.56±0.10	0.62±0.10	0.45	0.54	0.66	0.76	0.62	0.58	0.07	0.29
	Moodle	51	0.57±0.14	0.61±0.13	0.45	0.51	0.75	0.75	0.59	0.65	0.00	0.12
	Total	151	0.56±0.08	0.62±0.08	0.45	0.53	0.69	0.76	0.61	0.60	0.04	0.24
English	Conecto	61	0.61±0.12	0.57±0.12	0.49	0.45	0.79	0.78	0.50	0.33	0.18	0.23

Table 7. Experiments with semi-supervised extracted sentences from the whole comments. The unit is one sentence which has own meaning. First column of each evaluation metric (in grey) is obtained with our approach based on CNN, the second column (in white) is obtained with MeaningCloud models.

Lang.	VLE	Tot.	Acc		MF ₁	F ₁ ^{pos}	F ₁ ^{neg}	F ₁ ^{neu}				
Spanish	Conecto	164	0.65±0.07	0.66±0.07	0.50	0.53	0.76	0.82	0.63	0.60	0.12	0.16
	Moodle	63	0.67±0.12	0.65±0.12	0.49	0.52	0.81	0.77	0.67	0.65	0.00	0.13
	Total	227	0.65±0.06	0.66±0.06	0.50	0.53	0.77	0.81	0.64	0.62	0.09	0.16
English	Conecto	70	0.66±0.11	0.66±0.11	0.53	0.53	0.81	0.83	0.59	0.51	0.20	0.26

7 Conclusions and Future Work

In this paper, we have presented a sentiment analysis task to observations written in natural language extracted from questionnaires of postgraduate biomedical online learning students. As stated in the introduction, the application of sentiment analysis tools on UX comments will provide a better and accurate understanding on human needs in the interaction with VLEs. Two Learning Management Systems have been evaluated, both in Spanish and English, applying cross-domain polarity models trained with a corpus of a different domain (tweets for each language) and general models for the language. The obtained

results are very promising and they give an insight to keep going the research of applying sentiment analysis tools on User eXperience evaluation.

The ultimate goal is to develop further tools of automatic feed-back of user perception for designing virtual learning environments valued by users for its usability, quality and pleasure of use. For this, as a future work we will address automatic aspect detection (*pleasure of use, pleasure of learning, learning platforms, video, slides, usability, etc.*) and we will analyze the aspect polarity to capture relevant aspects that influence on the students and, possibly, were not considered during the questionnaire, or to analyze which aspects of a course tend to be more negative or positive for the students. Finally, we are now working with questionnaires on Massive Open Online Course (MOOC) to collect large amounts of data in order to train models for the task of sentiment analysis at global and aspect level on VLEs. A transfer learning approach from models trained with data of other domains could also be applied in order to have more robust models for the task.

References

1. Ergonomics of Human System Interaction-Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO), Switzerland (2009)
2. Baziotis, C., Pelekis, N., Doulkeridis, C.: Datastories at SemEval-2017 task 4: deep LSTM with attention for message-level and topic-based sentiment analysis. In: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), pp. 747–754 (2017)
3. González, J.Á., Plà, F., Hurtado, L.F.: ELiRF-UPV at SemEval-2017 task 4: sentiment analysis using deep learning. In: SemEval@ACL (2017)
4. González, J.Á., Hurtado, L.F., Pla, F.: Análisis de sentimientos en Twitter Basado en Aprendizaje Profundo. In: TASS 2018: Workshop on Semantic Analysis at SEPLN (TASS 2018). CEUR Workshop Proceedings, vol. 2172 (2018)
5. GoogleCloud: Cloud Natural Language API (2019). <https://cloud.google.com/natural-language/>
6. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. Behav. Inf. Technol. **25**(2), 91–97 (2006). <https://doi.org/10.1080/01449290500330331>
7. IBM: Natural Language Understanding (2019). <https://www.ibm.com/watson/services/natural-language-understanding/>
8. Kim, Y.: Convolutional neural networks for sentence classification. arXiv preprint [arXiv:1408.5882](https://arxiv.org/abs/1408.5882) (2014)
9. Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., Sinnelä, A.: UX Curve: a method for evaluating long-term user experience. Interact. Comput. **23**(5), 473–483 (2011)
10. Martínez-Cámara, E., Díaz-Galiano, M., García-Cumbreras, M., García-Vega, M., Villena-Román, J.: Overview of TASS 2017. In: Proceedings of TASS 2017: Workshop on Semantic Analysis at SEPLN (TASS 2017). CEUR Workshop Proceedings, vol. 1896 (2017)
11. MeaningCloud: Demo de Analítica de Textos (2019). <https://www.meaningcloud.com/es/demos/demo-analitica-textos>

12. MeaningCloud: MeaningCloud: Servicios web de analítica y minería de textos (2019). <https://www.meaningcloud.com/>
13. MicrosoftAzure: Text Analytics API (2019). <https://azure.microsoft.com/es-es/services/cognitive-services/text-analytics/>
14. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing, vol. 10, pp. 79–86 (2002)
15. Pla, F., Hurtado, L.F.: Spanish sentiment analysis in Twitter at the TASS workshop. *Lang. Resour. Eval.* **52**(2), 645–672 (2018). <https://doi.org/10.1007/s10579-017-9394-7>
16. Rauschenberger, M., Schrepp, M., Cota, M.P., Olschner, S., Thomaschewski, J.: Efficient measurement of the user experience of interactive products. How to use the User Experience Questionnaire (UEQ). Example: Spanish language version. *Int. J. Interact. Multimedia Artif. Intell.* **2**(1), 39–45 (2013). <https://doi.org/10.9781/ijimai.2013.215>
17. Rosenthal, S., Farra, N., Nakov, P.: Semeval-2017 task 4: sentiment analysis in Twitter. In: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), pp. 502–518 (2017)
18. Socher, R., et al.: Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 1631–1642 (2013)
19. Zaharias, P., Mehlenbacher, B.: Editorial: Exploring User Experience (UX) in virtual learning environments. *Int. J. Hum.-Comput. Stud.* **70**(7), 475–477 (2012). <https://doi.org/10.1016/j.ijhcs.2012.05.001>

2.4. Publication 4 | CROSS-DOMAIN POLARITY MODELS TO EVALUATE USER EXPERIENCE IN E-LEARNING

Sanchis-Font, R., Castro-Bleda, M. J., González, J. Á., Pla, F., & Hurtado, L. F. (2021). Cross-Domain polarity models to evaluate user eXperience in E-learning. *Neural processing letters*, 53(5), 3199-3215.

Publication 4 investigates the automatic evaluation of the user experience of e-learning platforms with a larger number of sentiment analysis techniques and a robuster corpus. This study evaluates the opinions of 583 users (107 English speakers and 476 Spanish speakers) and their positive, negative or neutral polarity by applying six different models (3 deep neural network models and 3 commercial sentiment analysis tools). The results allow us to advance in UX e-learning research with artificial intelligence techniques, which will be the starting point of publication 5 with the aim of analysing the sentiment of characteristic elements in e-learning.

Cross-Domain Polarity Models to Evaluate User eXperience in E-learning

Rosario Sanchis-Font¹  · Maria Jose Castro-Bleda¹  · José-Ángel González¹  · Ferran Pla¹  · Lluís-F. Hurtado¹ 

letters, 53(5), 3199-3215. DOI: <https://doi.org/10.1007/s11063-020-10260-5>

Abstract

Virtual learning environments are growing in importance as fast as e-learning is becoming highly demanded by universities and students all over the world. This paper investigates how to automatically evaluate User eXperience in this domain using sentiment analysis techniques. For this purpose, a corpus with the opinions given by a total of 583 users (107 English speakers and 476 Spanish speakers) about three learning management systems in different courses has been built. All the collected opinions were manually labeled with polarity information (positive, negative or neutral) by three human annotators, both at the whole opinion and sentence levels. We have applied our state-of-the-art sentiment analysis models, trained with a corpus of a different semantic domain (a Twitter corpus), to study the use of cross-domain models for this task. Cross-domain models based on deep neural networks (convolutional neural networks, transformer encoders and attentional BLSTM models) have been tested. In order to contrast our results, three commercial systems for the same task (MeaningCloud, Microsoft Text Analytics and Google Cloud) were also tested. The obtained results are very promising and they give an insight to keep going the research of applying sentiment analysis tools on User eXperience evaluation. This is a pioneering idea to provide a better and accurate understanding on human needs in the interaction with virtual learning environments and a step towards the development of automatic tools that capture the feed-back of user perception for designing virtual learning environments centered in user's emotions, beliefs, preferences, perceptions, responses, behaviors and accomplishments that occur before, during and after the interaction.

Keywords Machine learning · Artificial neural networks · Sentiment analysis · User experience · Virtual learning environments · Learning management systems

Partially supported by the Spanish MINECO and FEDER funds under Project TIN2017-85854-C4-2-R. Work of J.A. González is financed under Grant PAID-01-17.

1 Introduction

Human computer interaction (HCI) tools developers, agents and industry require to focus their interactive systems on end-users in order to design and provide quality systems upon the international standards requirements ISO. These interactive systems are the “combination of hardware, software and/or services that receives input from, and communicates output to, users” (ISO 9241-210:2019) [18]. This international standard is related to ergonomics of human system-interaction and human-centered design for interactive systems. It provides requirements and recommendations for human-centered design principles and activities throughout the life cycle of computer-based interactive systems. It is intended to be used by those managing design processes, and is concerned with ways in which both hardware and software components of interactive systems can enhance human–system interaction.

Therefore “User eXperience” (UX) enhances human interaction within the hardware or software components, being the UX concept multidimensional and centered in human needs. This UX concept goes beyond usability, interaction experience and design by involving two main qualities: traditional HCI usability and accessibility balanced with hedonic and affective design [42]. In this perspective, in [14], UX is described as a consequence of a user’s internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.), and the context (or the environment) within which the interaction occurs (e.g. organizational/social setting, meaningfulness of the activity, voluntariness of use, etc.). Therefore, these authors conclude that UX is considering three perspectives: emotion and affect of the user, technology and the hedonic instrument and the experiential aspect. As a result, UX includes a multidimensional concept and focuses in human needs and the aspects of beauty, fun, pleasure, and personal growth rather than the value of the product or instrument used [14], which improves or worsens along the time of use [21].

UX has to be considered when designing and redesigning hardware and software applications. In this way, in the last years, UX has been taken into account when designing Virtual Learning Environments (VLEs) [42]. VLEs includes a wide range of technology-enabled learning environments, such as Learning Management Systems (LMSs), computer games or Virtual Worlds.

Traditionally, evaluation of UX in VLEs (or in any other product or service), has always been addressed by conventional questionnaires. In this regard, we used the validated User Experience Questionnaire (UEQ) [32], conducted on students of biomedical postgraduate studies and Massive Open Online Courses (MOOCs) students. Three LMSs have been evaluated using this adapted UEQ: an ad-hoc system called “Conecto” (in Spanish and English languages), an open-source Moodle personalized system (in Spanish), and an edX platform (both in Spanish and English languages).

In this paper, instead of evaluating UX in this traditional way, we have addressed the problem in a novel way applying machine learning tools to the users’ opinions, expressed freely in natural language. Preliminary work was done in [35]. Deep learning cross-domain models (Convolutional Neural Networks, Transformer Encoders and Attentional BLSTM models) trained with tweets and different general systems for text analytics (such as MeaningCloud [27,28], Google Cloud [13], and Microsoft Text Analytics [29]) are used to this end. The application of sentiment analysis tools on UX opinions will provide a better and accurate understanding on human needs in the interaction with VLEs. The ultimate goal of this work is to develop further tools of automatic feed-back of user perception for designing user-centered VLEs valued by users for its usability, quality and pleasure of use.

This paper is organized as follows. Next section gives a brief overview about the state of the art of the work presented here. Section 3 describes the data collection from the questionnaires and the labeling process. The evaluation metrics are introduced in Sect. 4. The used models for sentiment analysis are described in Sects. 5 and 6 presents our proposal. Section 7 presents the experimental results and their analysis. Finally, the conclusions and future directions are drawn in the last section.

2 State of the Art

Sentiment analysis is one of the most active areas in Natural Language Processing since the early 2000s. The pioneering works in this field [30,39] pointed out the importance of “sentiment classification” for a large number of tasks such as *message filtering*, *recommender systems* or *business intelligence applications*. Other sentiment analysis approaches were addressed by manually generating polarity lexicons [23,41]. However, the efforts required to develop these resources and the good performance of machine learning systems on this task made the research community to move towards data driven approaches. A survey of the most widely used machine learning approaches for the sentiment analysis problem can be found in [22].

Recently, the predominant systems to perform sentiment analysis are neural network based approaches [43]. The most popular models are Convolutional Neural Networks (CNN) [19], Long Short Term Memories (LSTM) [15], and combinations of CNN and LSTM [34]. Moreover, the enrichment of this architectures by using attention mechanisms [2] and Transformers [40] are lately used.

The interest on sentiment analysis has increased along with the popularity of the social networks and the user interactions on them. The most studied social network for sentiment analysis tasks is Twitter, where the users are allowed to broadcast opinions about any topic by using only 280 characters and media content.

Several workshops are organized in order to address the sentiment analysis task in Twitter, providing corpora and resources to the participants for training and evaluating their systems. The most known workshops are the International Workshop on Semantic Evaluation (SemEval) and the Workshop on Semantic Analysis at SEPLN (TASS) for English and Spanish language, respectively.

For the last task of English sentiment analysis presented at SemEval [33], most of the participating teams proposed neural network models mainly based on LSTM and CNN, being the two best systems based on these approaches along with pre-trained word embeddings on big collections of tweets. Concretely, the winner team proposed a two layer bidirectional LSTM with attention mechanisms [3], while the second ranked team addressed the task by using a combination of LSTM and CNN [4].

For the Spanish sentiment analysis task of TASS 2019 [7], the predominant presence of deep learning components was also observable, where almost all the systems proposed by the participants made use of them. It is worthy to note the great interest on the Transformer model [40], being used mainly with the aim of fine-tuning pre-trained contextual representations of words [6].

Our team proposed a system focused on encoding pre-trained skipgram word embeddings by using a Transformer encoder to carry out the classification [12]. It turned out to be the best system, being the first ranked system. The system of the second ranked team used a logistic

regression classifier on top of different representations, such as word embeddings and bag of characters, by focusing on a novel way of data augmentation [24].

In addition to these kinds of systems, a large number of commercial products and frameworks have also proliferated to facilitate the development and deployment of sentiment analysis systems based on machine learning, such as Google Cloud [13], IBM Watson [17], Microsoft Text Analytics [29], MeaningCloud [28] or Stanford Core NLP [25]. These products allow us to perform text analytics such as sentiment analysis, in a broad variety of domains and languages in an easy way, obtaining also competitive results. For this reason, besides our neural network models, other commercial models will be used in our work as explained in Sect. 5.2.

But, though the promising results on several tasks of natural language processing and, in particular, on sentiment analysis, generally speaking, UX evaluation is immature in most applications and, especially, in VLEs.

Some work has been done in eCommerce, using natural language processing to improve their UX. For instance, to search products in a more intelligent way, using sentiment analysis to extract insights from the reviews made by the customers on the product or identifying trends and trying to answer best to the customers' concerns. Several new conferences have recently been launched around these ideas, such as the Workshop on Economics and Natural Language Processing¹ or the First International Workshop on e-Commerce and NLP.²

Another research line covered in this paper is the use of cross-domain polarity classification approach, that is, the texts to be classified belong to a different domain from those used in the training phase. Most work has been done within the classic approach, the so-called single-domain polarity classification, which classifies texts in the same domain to which the texts used in the training phase belong to. Due to the lack of training data (only opinions from 583 users), we used cross-domain models, those trained with another domain (Twitter) and those trained with general data.

3 Experimental Data

The validated User Experience Questionnaire (UEQ) [32] was used in order to automatically evaluate UX in our VLEs. This questionnaire is a list of close-ended questions, but we added questions concerning to sociodemographic data (age, sex, etc.) and an open field "Other comments" (see Fig. 1 for a screenshot of the questionnaire in one VLE). It is a text entry box to express any opinion or comment related to UX in the course, which is an opportunity to get new and more precise information about their experience, not only by close-ended questions.

Three LMSs have been evaluated using this adapted UEQ: "Conecto",³ which is an ad-hoc system (for Spanish and English users), an open-source Moodle personalized system (for Spanish users),⁴ and an edX platform (both for Spanish and English languages).⁵ We have collected data in different editions of the courses, obtaining an answer to the "Other comments" box from 583 users (107 English speakers and 476 Spanish speakers).

¹ <https://julielab.de/econlp/2019/>.

² <https://www.aclweb.org/portal/content/first-international-workshop-e-commerce-and-nlp>.

³ https://postgrado.adeituv.es/es/cursos/salud-7/assisted-reproduction/datos_generales.htm.

⁴ <https://medicinagenomica.com/eugmygo/>.

⁵ <https://www.upvx.es/>.

Table 1 Examples of tagged observations and sentences, with their polarity

Unit	Example	Polarity
Observation	Overall, this e-learning master environment is very friendly	Positive
Sentence	Overall, this e-learning master environment is very friendly	Positive
Observation	It was good experience to some extent. However, I hope it concentrates more on practical aspect in the future	Neutral
Sentence	It was good experience to some extent	Positive
Sentence	However, I hope it concentrates more on practical aspect in the future	Negative
Observation	Well-organized and structured course. Great study material (articles) but not enough time to read them all. Keep up the good work	Neutral
Sentence	Well-organized and structured course	Positive
Sentence	Great study material (articles) but not enough time to read them all	Neutral
Sentence	Keep up the good work	Positive

(e.g., “Some of the modules were very interesting and valuable but some of them confusing as too genetic details involved.”). Three human annotators (as in [38]) did the annotation of each sentence as positive, negative or neutral.

Secondly, as an observation is a sequence of sentences, and, following Socher’s work [38], based on the structure of the discourse of the observations, observation labeling was carried out from the polarity level of the sentences which compose the observation. The core idea is that the polarity of an observation will be automatically set as positive if it is composed of positive sentences; similarly, it will be set as negative if every sentences is negative; and finally, it will be tagged as neutral if it is composed of positive and negative and/or neutral sentences.

In order to evaluate the inter-annotator agreement we used the following measures: Krippendorff’s alpha (α) [20], Cohen’s kappa (κ) [5] and Scott’s pi (π) [37], both for the observation and sentence levels of the labeling. The obtained results suggest a high correlation among the labeling work of the three annotators at both levels, concretely, $\alpha = \kappa = \pi = 0.88$ for whole Spanish observations, $\alpha = \kappa = \pi = 0.90$ for Spanish sentences, $\alpha = \kappa = \pi = 0.84$ for whole English observations and $\alpha = \kappa = \pi = 0.90$ for English sentences.

These results seem to suggest that the sentiment is more detectable at sentence level than at observation level, where several opinions with different polarity are more likely to happen, therefore, observations are more difficult to label.

The total number of units and the class distributions can be seen in Table 2. As it can be observed, there are more positive than negative samples. The neutral category decreased from the whole comment (a complex statement) to the sentence (usually, with polarity or, less frequently, with lack of sentiment). All samples were used as test set, and they were automatically labeled by using the proposed models (the neural networks systems developed in this work and other general models from commercial tools) and compared with the ground truth label.

Table 2 Observations and sentences extracted from the “Other comments” box

Unit	Language	Total	Positive	Negative	Neutral
Observation	Spanish	476	338 (71%)	85 (18%)	53 (11%)
	English	107	56 (52%)	30 (28%)	21 (20%)
Sentence	Spanish	587	404 (69%)	142 (24%)	41 (7%)
	English	184	90 (49%)	80 (43%)	14 (8%)

4 Evaluation Metrics

Different evaluation metrics were used in order to test the systems. Concretely, as defined below, accuracy (Acc , Eq. 1) and macro F_1 (MF_1 , Eq. 3) were used to reduce the impact of corpus imbalance in the evaluation. Moreover, the F_1 per class c (*positive*, *negative*, or *neutral* class) as defined in Eq. 2 was computed to observe the behavior of our systems at class level.

$$Acc = \frac{\sum_{c \in C} \sum_{x \in \Omega_c} [f(x) = c]}{|\Omega|} \quad (1)$$

$$F_1^c = \frac{2 \cdot P_c \cdot R_c}{P_c + R_c} \quad (2)$$

$$MF_1 = \frac{1}{|C|} \sum_{c \in C} F_1^c \quad (3)$$

Ω is the set of samples, Ω_c are the samples of class c in Ω , $y(x)$ is the prediction of the model f for a given sample x , C is the set of classes, $[\cdot]$ denotes the Iverson bracket, and P_c and R_c are the precision and recall measure of each class, defined as follows:

$$P_c = \frac{\sum_{x \in \Omega_c} [f(x) = c]}{\sum_{x \in \Omega} [y(x) = c]} \quad R_c = \frac{\sum_{x \in \Omega_c} [f(x) = c]}{|\Omega_c|} \quad (4)$$

Moreover, the macro-precision (MP) and macro-recall (MR) are also considered in order to compare the results of our supervised systems with those officially published at SemEval and TASS workshops.

$$MP = \frac{1}{|C|} \sum_{c \in C} P_c \quad MR = \frac{1}{|C|} \sum_{c \in C} R_c \quad (5)$$

5 Cross-Domain Polarity Models

Cross-domain models for both Spanish and English were used to address the problem of sentiment analysis on VLEs. On the one hand, Deep Neural Networks such as Convolutional Neural Networks (CNN), Attentional Bidirectional Long Short Term Memory, and Transformer Encoders (TE) models were used to train models for sentiment analysis tasks on Twitter, both in Spanish and English, proposed in international competitions [7,26,33]. On the other hand, we used the sentiment analysis module provided by several commercial “Software as a service” text analytics products: MeaningCloud [28], Microsoft Text Analytics [29] and Google Cloud [13], which act as general domain polarity classifiers both for the English and the Spanish languages.

Table 3 Characteristics of the corpora used to train the deep neural network models (both for English and Spanish)

Task	Set	Total	Positive	Negative	Neutral	None
SemEval (English)	Train	39,656	15,705 (40%)	6203 (15%)	17,748 (45%)	N/A
	Test	12,284	2375 (19%)	3972 (32%)	5937 (49%)	N/A
	Total	51,940	18,080 (35%)	10,175 (19%)	23,685 (46%)	N/A
TASS (Spanish)	Train	11,755	4211 (36%)	4250 (36%)	1321 (11%)	1973 (17%)
	Test	1507	547 (36%)	543 (36%)	155 (11%)	262 (17%)
	Total	13,262	4758 (36%)	4793 (36%)	1476 (11%)	2235 (17%)

5.1 Deep Neural Networks

To determine the polarity of the students' opinions, we used several polarity supervised models based on the use of word embeddings and deep learning. Unfortunately, due to the lack of training data, it was not possible to learn robust models specifically for the task described in this paper.

Instead, we used models trained by our research group, for similar tasks related to the social network Twitter [7,26,33] both for English and Spanish. The English corpus (including the partitions for training, development and testing purposes) is provided in the Subtask A of Task 4 from SemEval 2017 [33] intended to detect the overall sentiment of a tweet. The Spanish corpus is a combination of two TASS editions (2017 [26] and 2019 [7]) with the aim of increasing the corpus size and taking into account several Spanish variants (including Spain, Mexico, Costa Rica and Uruguay). Due to the masters are opened to international students, several Spanish variants can be used by them, therefore, it is interesting to consider some of these variants during the training phase of the supervised models. In the Spanish case, partitions for training, development and testing were built following a 80%–10%–10% proportion. Table 3 shows some details of the corpora used to train the models.

As input for CNN, TE and BLSTM models, each opinion is represented as a $N \times d$ matrix where each word of the opinion—up to a maximum of N —is represented as a d -dimensional embedding. Depending on the language, different word embeddings are used. For Spanish, 300-dimensional skip gram word embeddings were learned from 87 millions of tweets [16], whereas for English, 400-dimensional skip gram word embeddings from [8] were used.

5.1.1 Convolutional Neural Networks

This architecture is inspired by the work described in [19], which obtained competitive results in text classification tasks such as sentiment analysis or irony detection.

We applied several one-dimensional (the width of the filter is constant and equal to the dimension of the embeddings) convolutions with different height filters, in order to extract the sequential structure of the text. Concretely, heights from 1 to 3 with 256 filters for each height are used. Subsequently, we applied Global Max Pooling to the feature maps in order to extract the most salient features for each region size.

The final decision is carried out by a softmax fully-connected layer. Table 4 show the performance of the models for the test set of the two tasks [11,31].

Note that for TASS, the distinction between the classes Neutral (*with both positive and negative feelings*) and None (*lack of sentiment*) is made during training and test. However,

Table 4 Performance of the supervised deep learning systems on SemEval 2017 Task 4 (English) and TASS (Spanish) for the test set

	CNN				TE				BLSTM			
	Acc	MP	MR	MF ₁	Acc	MP	MR	MF ₁	Acc	MP	MR	MF ₁
Spanish	66.29	55.63	55.07	54.05	66.82	58.40	57.93	57.63	65.49	55.40	52.81	52.71
English	63.88	63.27	62.24	62.59	63.35	62.24	63.97	62.74	64.54	63.43	65.30	64.19

when the model trained for TASS is applied to our UX evaluation task, both classes are merged. That is, given a test opinion x , $\operatorname{argmax}_y p(y|x) \in \{Neutral, None\} \rightarrow y = Neutral$. This is also true for all the other Deep Learning systems presented in this section.

5.1.2 Transformer Encoder

Our system is based on the Transformer [40] model. Initially proposed for machine translation, the Transformer model dispenses with convolution and recurrences to learn long-range relationships. Instead of this kind of mechanisms, it relies on multi-head self-attention, where multiple attentions among the terms of a sequence are computed in parallel to take into account different relationships among them.

On top of the tweet representations, $N_x = 1$ transformer encoders are applied, which relies on multi-head scaled dot-product attention with $h = 8$ different heads and $d_k = d_q = d_v = 64$ attention dimensionality. To do this we used an architecture similar to the one described in [40], including the layer normalization [1] and the residual connections. Due to a vector representation is required to train classifiers on top of these encoders, a global average pooling mechanism was applied to the output of the encoder, and it is used as input to a feed-forward neural network, with only one hidden layer, whose output layer computes a probability distribution over the classes of the task.

5.1.3 Attentional Bidirectional Long Short Term Memory

The system is based on Bidirectional Long Short Term Memory (BLSTM) [15,36] with attention mechanisms. On top of the tweet representations, one 256-dimensional BLSTM is applied and, a context vector is computed from the outputs of the BLSTM network following [2]. In this way, the context vector is a weighted sum of the BLSTM outputs, where the weight associated to each output is computed by means of a feed-forward neural network which is jointly trained with all the other components of the system. Then, the context vector is used as input to a feed-forward neural network with one 256 dimensional hidden layer, whose output layer computes a probability distribution over the classes of the task. Again, the Neutral and None classes are merged when the model is applied on the Spanish UX task, due to the distinction between them in the TASS corpus.

Table 4 show the performance of the models for the test set of the SemEval 2017 and TASS 2019 tasks [11,31]. Again, the Neutral and None classes are merged when the model is applied on the Spanish UX task, due to the distinction between them in the TASS corpus.

5.2 Commercial Systems

5.2.1 MeaningCloud

MeaningCloud is a Software as a Service product [28] that provides a large number of tools, easy to use and to deploy, for text processing, analytics and text/audio mining, with the aim of facilitating the resolution of natural language processing problems to developers. It includes tools for summarization, topic extraction, language identification and sentiment analysis and it supports several languages.

We have used the sentiment analysis module in this work. This module allowed us to use a classifier, trained in a general domain with texts in multiple languages, to determine the global polarity of user opinions on VLEs. Concretely, we use the field *score tag* in the response of the MeaningCloud API, that indicates the global polarity of the text in 6 different levels: strong positive, positive, neutral, negative, strong negative and without sentiment (None). To carry out our experiments, we collapsed the strong sentiments, i.e., strong positive and strong negative are considered as Positive and Negative, respectively. Moreover, the Neutral/None classes are merged in only one class (Neutral).

5.2.2 Google Cloud

Natural Language API of Google Cloud [13] allows to perform several kinds of analysis such as syntactic parsing, entity or sentiment analysis, in general domain texts and also for several languages. It is based on machine learning with the aim of analyzing the structure and the meaning of documents. For sentiment analysis, it computes two values for each document, *score* and *magnitude*. The overall sentiment of the document is computed by the $score \in [-1, 1]$, where negative values refer to Negative sentiment, positive values refer to Positive sentiment and values closer to 0 could suggest the absence of sentiment (None) or the neutralization of positive and negative sentiments (Neutral). In order to distinguish between None and Neutral, the system computes the $magnitude \in [0, \text{inf}]$ which quantifies the sentiment content in the document, so that documents with *score* closer to 0 will be Neutral if its *magnitude* indicates the presence of sentiment ($magnitude > 0$) or None if there is not ($magnitude = 0$). In our case, we only used the *score* value due to in the UX corpus, the Neutral class indicates both situations.

In order to carry out the experimentation and a fair comparison with the supervised systems based on Deep Neural Networks, we fixed a threshold ϵ to a reasonable value of $\epsilon = 0.15$, so that if $-\epsilon \leq score \leq \epsilon$, then the polarity is Neutral; if $score < -\epsilon$, then it is Negative; and, otherwise, the polarity is Positive.

5.2.3 Microsoft Text Analytics

Text Analytics API of Microsoft Azure [29] provides automatic tools to evaluate opinions and topics, such as sentence extraction, entity recognition and sentiment analysis. The system is based on machine learning and it uses different features such as n -grams, word embeddings and part-of-speech tags, being compatible with several languages. The sentiment analysis module computes a $score \in [0, 1]$, being 0 the most Negative value and 1 the most Positive, where values closer to 0.5 suggests Neutral polarity, due to the objectivity or the neutralization of positive and negative sentiments.

In this case, the classification rule is the same as the previously commented for the Google system, but with $\epsilon = 0.05$ and moving the origin from 0 to 0.5; i.e., if $0.5 - \epsilon \leq score \leq$

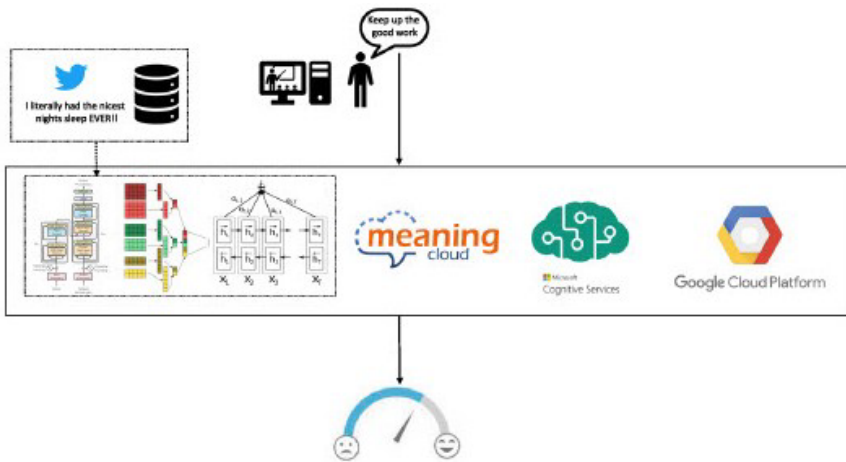


Fig. 2 A scheme of our proposal for evaluating UX about E-learning platforms using general and cross-domain sentiment analysis systems

$0.5 + \epsilon$, then the polarity is Neutral; if $score < 0.5 - \epsilon$, then it is Negative; and, otherwise, the polarity is Positive.

6 Our Proposal

In this paper, we address the UX evaluation problem in a novel way. To do this, we propose to apply machine learning tools to the users' opinions, expressed freely in natural language and additionally, due to the lack of training data in the UX domain, we propose to use general-purpose and cross-domain systems. A scheme of our proposal is illustrated in Fig. 2. We used six different systems to analyze the polarity of the students' opinions about the learning platform. On the one hand, we have considered three general-purpose commercial systems (MeaningCloud [28], Google Cloud [13], and Microsoft Text Analytics [29]). On the other hand, we have used three deep learning cross-domain models (Convolutional Neural Networks, Transformer Encoders and Attentional BLSTM models) developed by our team and initially trained for the sentiment analysis problem in Twitter [9–11].

7 Experimental Results

We applied the cross-domain polarity models presented in Sect. 5 to the proposed task which consists of determining the polarity (positive, negative, or neutral) of the users' opinions about the learning platform. The opinions, as explained in Sect. 3, are processed as the whole observation from one user or, if the observation is composed of more than one sentence, each single sentence. The results obtained with each type of polarity models for observations and sentences in the two considered languages are shown in Tables 5, 6, 7 and 8.

First of all, it is important to highlight the good results obtained by all the systems for the Positive class (F_1^{pos} row in all tables), which is also the highest frequency class, as observed

Table 5 Experiments at observation level on the Spanish samples

Spanish observations						
	CNN	TE	BLSTM	Mean.C.	Google	Microsoft
<i>Acc</i>	76.68	77.94	78.15	78.57	77.73	70.59
<i>MF</i> ₁	53.06	61.72	58.55	59.47	55.64	51.98
<i>F</i> ₁ ^{pos}	87.52	88.86	88.24	88.32	87.72	83.21
<i>F</i> ₁ ^{neg}	66.33	66.67	68.82	68.35	65.25	60.87
<i>F</i> ₁ ^{neu}	5.33	29.63	18.60	21.74	13.95	11.86

Table 6 Experiments at sentence level on the Spanish samples

Spanish sentences						
	CNN	TE	BLSTM	Mean.C.	Google	Microsoft
<i>Acc</i>	78.19	76.49	80.07	76.15	76.49	68.14
<i>MF</i> ₁	55.38	58.73	61.62	55.69	51.71	49.35
<i>F</i> ₁ ^{pos}	87.30	87.41	87.65	87.75	87.20	82.38
<i>F</i> ₁ ^{neg}	70.83	67.92	74.13	66.94	60.71	55.74
<i>F</i> ₁ ^{neu}	8.00	20.87	23.08	12.39	7.23	9.93

in Table 2. However, the Neutral class has a totally different behavior. All systems obtained much worse results for this class, with differences of almost 50 points for the same system between F_1^{pos} and F_1^{neu} . It should be noted that the *Acc* confidence intervals (excluded from the tables) are wide, ranging from $\pm 3.23\%$ to $\pm 9.38\%$, mainly due to the small size of the corpora. Therefore, there is a considerable amount of uncertainty on the results and the following conclusions should be taken with caution.

Regarding the experiments using the Spanish observations, BLSTM, TE and Meaning-Cloud systems achieved the best results (see Table 5), being the TE system the best one in terms of *MF*₁ and the MeaningCloud system the best one considering only *Acc*. In the case of the Spanish sentence level experiments (Table 6), our BSLTM system achieved the best performance. As in general in all the experimentation, the Positive class (both in observations and sentences experiments) obtained the highest *F*₁ results, and the Neutral class the one with the lowest. At sentence level, our three cross-domain models trained with tweets showed a better behavior than those from general-domain commercial systems. We hypothesize that this is due to the greater similarity of sentences and tweets compared to the whole observation, with larger length than one tweet. That is, the test samples at sentence level are more similar to those used to learn our models and therefore these models perform better.

Regarding the experiments using the English samples, the best accuracy and *MF*₁ values were obtained by Google, Microsoft and our BLSTM model. At observation level, Google system obtained the best results, whereas at sentence level the most competitive system was BLSTM. The worst results were obtained by the CNN and the MeaningCloud system.

Microsoft model was the best detecting Negative samples. The TE system was the worst system detecting the negative samples at observation level, while the CNN model was the worst for detecting this class at the sentence level. As in the case of the Spanish samples, in the English samples the Neutral class was the class with the worst results, being the performance

English observations						
	CNN	TE	BLSTM	Mean.C.	Google	Microsoft
<i>Acc</i>	57.01	57.94	68.22	60.75	66.36	71.03
<i>MF₁</i>	49.01	47.55	56.02	50.87	58.54	56.13
<i>F₁^{pos}</i>	74.78	78.33	86.61	76.42	77.69	82.54
<i>F₁^{neg}</i>	60.00	43.48	52.17	57.14	67.93	69.84
<i>F₁^{neu}</i>	12.24	20.83	29.27	19.05	30.00	16.00

Table 8 Experiments at sentence level on the English samples

English sentences						
	CNN	TE	BLSTM	Mean.C.	Google	Microsoft
<i>Acc</i>	57.61	61.41	70.65	63.04	69.57	73.91
<i>MF₁</i>	50.16	51.44	58.14	49.60	55.55	54.35
<i>F₁^{pos}</i>	80.23	81.97	89.12	79.19	80.42	82.13
<i>F₁^{neg}</i>	53.57	54.70	63.87	57.38	72.59	70.92
<i>F₁^{neu}</i>	16.67	17.65	21.43	12.25	13.64	10.00

of all the systems for this class much lower than for the other classes. The Neutral samples are more complex in structure due to the neutralization of positive and negative elements in the same unit or the absence of polarity. If we compare the Neutral results for English and Spanish it is possible to see that the commercial system, with the exception of MeaningCloud, achieved results that are slightly higher for the English opinions, suggesting that these systems have been better adjusted for processing English documents.

It is important to highlight the good behavior obtained by our cross-domain models in comparison with the general-domain commercial systems that have been used in the experimentation. They are quite competitive despite having been only trained with tweets. Different from user opinions expressed in VLE, where a formal communication is carried out addressing a set of topics related to the course that they have taken, the tweets are informal and they express opinions of many different topics in a way influenced by the behavior of the Twitter social network (slang, user mentions, hashtags, lexical errors, elongations, etc.). This seems to suggest that there are related properties among the opinions expressed on VLE and those expressed on Twitter.

As stated in [38]: “However, sentiment accuracies even for binary positive/negative classification for single sentences has not exceeded 80% for several years. For more difficult multiclass case including a neutral class, accuracy is often below 60% for short messages on Twitter (Wang et al. 2012)”. The accuracies of our models are in the state of the art for sentiment analysis in Twitter, and the evaluation metrics are also very similar when the models are applied to our UX dataset, which is a multiclass problem with 3 classes (Positive, Negative, Neutral).

8 Conclusions and Future Work

In this paper, we have presented a sentiment analysis task to opinions written in natural language extracted from questionnaires of postgraduate biomedical and MOOCs online learning students. As stated in the introduction, the application of sentiment analysis tools on UX comments will provide a better and accurate understanding on human needs in the interaction with VLEs. Three Learning Management Systems have been evaluated, both in Spanish and English, applying cross-domain polarity models trained with a corpus of a different domain (tweets for each language) and general models for the language. The obtained results are very promising and they give an insight to keep going the research of applying sentiment analysis tools on User eXperience evaluation.

The ultimate goal is to develop further tools of automatic feed-back of user perception for designing virtual learning environments valued by users for its usability, quality and pleasure of use. For this, as a future work, we will address automatic aspect detection (*pleasure of use, pleasure of learning, learning platforms, video, slides, usability, etc.*) and its polarity will be analyzed in order to capture relevant aspects that influence on the students and, possibly, were not considered during the questionnaire, or to analyze which aspects of a course tend to be more negative or positive for the students. Finally, we are now continuing working with questionnaires on MOOCs to collect larger amounts of data in order to train models for the task of sentiment analysis at global and aspect level on VLEs. A transfer learning approach from models trained with data of other domains could also be applied in order to have more robust models for the task.

Acknowledgements Special thanks to the following biomedical organizations: Fundación IVI and Medigene Press S.L.; both have provided data from their Master and Postgraduate Courses through the academic stay research of Rosario Sanchis-Font, during 2017 and 2018. Many thanks to Carlos Turró-Ribalta and Ignacio Despujol-Zabala for supporting this research with data from UPV MOOCs.

References

1. Ba J, Kiros JR, Hinton GE (2016) Layer normalization. [arxiv:1607.06450](https://arxiv.org/abs/1607.06450)
2. Bahdanau D, Cho K, Bengio Y (2015) Neural machine translation by jointly learning to align and translate. In: 3rd international conference on learning representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, conference track proceedings. [arxiv:1409.0473](https://arxiv.org/abs/1409.0473)
3. Baziotis C, Pelekis N, Doukeridis C (2017) Datastories at SemEval-2017 task 4: deep LSTM with attention for message-level and topic-based sentiment analysis. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017). Association for Computational Linguistics, Vancouver, Canada, pp 747–754
4. Cliche M (2017) BB_twttr at SemEval-2017 task 4: Twitter sentiment analysis with CNNs and LSTMs. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017). Association for Computational Linguistics, Vancouver, Canada, pp 573–580. <https://doi.org/10.18653/v1/S17-2094>
5. Cohen J (1960) A coefficient of agreement for nominal scales. *Educ Psychol Meas* 20(1):37
6. Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 conference of the North American chapter of the Association for Computational Linguistics: Human Language Technologies, volume 1 (long and short papers). Association for Computational Linguistics, Minneapolis, Minnesota, pp 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
7. Diaz-Galiano MC, et al (2019) Overview of TASS 2019: one more further for the global Spanish sentiment analysis corpus. In: Proceedings of the Iberian languages evaluation forum (IberLEF 2019), CEUR-WS, Bilbao, Spain, CEUR workshop proceedings, pp 550–560
8. Godin F, Vandersmissen B, De Neve W, Van de Walle R (2015) Multimedia lab @ ACL WNUT NER shared task: named entity recognition for Twitter microposts using distributed word representations. In:

- Proceedings of the workshop on noisy user-generated text. Association for Computational Linguistics, Beijing, China, pp 146–153. <https://doi.org/10.18653/v1/W15-4322>. <https://www.aclweb.org/anthology/W15-4322>
9. González J, Pla F, Hurtado L (2018) Elirf-upv en TASS 2018: Análisis de sentimientos en twitter basado en aprendizaje profundo (elirf-upv at TASS 2018: sentiment analysis in Twitter based on deep learning). In: Proceedings of TASS 2018: workshop on semantic analysis at SEPLN, TASS@SEPLN 2018, co-located with 34nd SEPLN conference (SEPLN 2018), Sevilla, Spain, September 18th, 2018, pp 37–44. http://ceur-ws.org/Vol-2172/p2_elirf_tass2018.pdf
 10. González J, Hurtado L, Pla F (2019) Elirf-upv at TASS 2019: transformer encoders for Twitter sentiment analysis in Spanish. In: Proceedings of the Iberian languages evaluation forum co-located with 35th conference of the Spanish Society for Natural Language Processing, IberLEF@SEPLN 2019, Bilbao, Spain, September 24th, 2019, pp 571–578. http://ceur-ws.org/Vol-2421/TASS_paper_2.pdf
 11. González JÁ, Pla F, Hurtado LF (2017) ELiRF-UPV at SemEval-2017 task 4: sentiment analysis using deep learning. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017). Association for Computational Linguistics, Vancouver, Canada, pp 723–727. <https://doi.org/10.18653/v1/S17-2121>. <https://www.aclweb.org/anthology/S17-2121>
 12. González JÁ, Hurtado LF, Pla F (2019) ELiRF-UPV at TASS 2019: transformer encoders for Twitter sentiment analysis in Spanish. In: Proceedings of the Iberian languages evaluation forum (IberLEF 2019), CEUR-WS, Bilbao, Spain, CEUR workshop proceedings
 13. GoogleCloud (2019) Cloud natural language API. <https://cloud.google.com/natural-language/>. Accessed 27 Dec 2019
 14. Hassenzahl M, Tractinsky N (2006) User experience—a research agenda. *Behav Inf Technol* 25(2):91–97. <https://doi.org/10.1080/01449290500330331>
 15. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
 16. Hurtado Oliver LF, Pla F, González Barba J (2017) ELiRF-UPV at TASS 2017: sentiment analysis in Twitter based on deep learning. In: TASS 2017: workshop on semantic analysis at SEPLN, pp 29–34
 17. IBM (2019) Natural language understanding. <https://www.ibm.com/watson/services/natural-language-understanding/>. Accessed 27 Dec 2019
 18. ISO 9241-210:2019 (2019) Ergonomics of human–system interaction—part 210: human-centred design for interactive systems. International Standardization Organization (ISO). <https://www.iso.org/standard/77520.html>. Accessed 27 Dec 2019
 19. Kim Y (2014) Convolutional neural networks for sentence classification. In: Proceedings of the 2014 conference on empirical methods in natural language processing, EMNLP 2014, October 25–29, 2014, Doha, Qatar, a meeting of SIGDAT, a special interest group of the ACL, pp 1746–1751. <http://aclweb.org/anthology/D/D14/D14-1181.pdf>
 20. Krippendorff K (2004) Reliability in content analysis. *Hum Commun Res* 30(3):411–433
 21. Kujala S, Roto V, Väänänen-Vainio-Mattila K, Karapanos E, Sinnelä A (2011) UX curve: a method for evaluating long-term user experience. *Interact Comput* 23(5):473–483
 22. Liu B (2012) Sentiment analysis and opinion mining. A comprehensive introduction and survey. Morgan & Claypool Publishers, San Rafael
 23. Liu B, Hu M, Cheng J (2005) Opinion observer: analyzing and comparing opinions on the web. In: Proceedings of the 14th international conference on world wide web. ACM, New York, NY, USA, WWW '05, pp 342–351. <https://doi.org/10.1145/1060745.1060797>
 24. Luque FM (2019) Atalaya at TASS 2019: data augmentation and robust embeddings for sentiment analysis. In: Proceedings of the Iberian languages evaluation forum (IberLEF 2019), CEUR-WS, Bilbao, Spain, CEUR workshop proceedings
 25. Manning CD, Surdeanu M, Bauer J, Finkel J, Bethard SJ, McClosky D (2014) The Stanford CoreNLP natural language processing toolkit. In: Association for computational linguistics (ACL) system demonstrations, pp 55–60. <http://www.aclweb.org/anthology/P/P14/P14-5010>
 26. Martínez-Cámara E, Díaz-Galiano M, García-Cumbreras M, García-Vega M, Villena-Román J (2017) Overview of TASS 2017. In: Proceedings of TASS 2017: workshop on semantic analysis at SEPLN (TASS 2017), CEUR-WS, Murcia, Spain, CEUR workshop proceedings, vol 1896
 27. MeaningCloud (2019) Demo de Análítica de Textos. <https://www.meaningcloud.com/es/demos/demo-analitica-textos>. Accessed 27 Dec 2019
 28. MeaningCloud (2019) MeaningCloud: Servicios web de análítica y minería de textos. <https://www.meaningcloud.com/>. Accessed 27 Dec 2019
 29. MicrosoftAzure (2019) Text analytics API. <https://azure.microsoft.com/es-es/services/cognitive-services/text-analytics/>. Accessed 27 Dec 2019

30. Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 conference on empirical methods in natural language processing, vol 10. Association for Computational Linguistics, pp 79–86
31. Pla F, Hurtado LF (2018) Spanish sentiment analysis in Twitter at the TASS workshop. *Lang Resour Eval* 52(2):645–672. <https://doi.org/10.1007/s10579-017-9394-7>
32. Rauschenberger M, Schrepp M, Cota MP, Olschner S, Thomaschewski J (2013) Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (UEQ). Example: Spanish language version. *Int J Interact Multimed Artif Intell* 2(1):39–45. <https://doi.org/10.9781/ijimai.2013.215>
33. Rosenthal S, Farra N, Nakov P (2017) SemEval-2017 task 4: sentiment analysis in Twitter. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017). Association for Computational Linguistics, Vancouver, Canada, pp 502–518. <https://doi.org/10.18653/v1/S17-2088>. <https://www.aclweb.org/anthology/S17-2088>
34. Sadr H, Pedram MM, Teshnehlab M (2019) A robust sentiment analysis method based on sequential combination of convolutional and recursive neural networks. *Neural Process Lett* 50:2745–2761. <https://doi.org/10.1007/s11063-019-10049-1>
35. Sanchis-Font R, Castro-Bleda M, González J (2019) Applying sentiment analysis with cross-domain models to evaluate user experience in virtual learning environments. In: Rojas I, Joya G, Catala A (eds) *Advances in computational intelligence. IWANN (2019). Lecture notes in computer science*, vol 11506. Springer, Cham, pp 609–620
36. Schuster M, Paliwal K (1997) Bidirectional recurrent neural networks. *Trans Signal Process* 45(11):2673–2681. <https://doi.org/10.1109/78.650093>
37. Scott WA (1955) Reliability of content analysis: the case of nominal scale coding. *Public Opin Q* 19(3):321–325. <https://doi.org/10.1086/266577>
38. Socher R, Perelygin A, Wu J, Chuang J, Manning CD, Ng A, Potts C (2013) Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the 2013 conference on empirical methods in natural language processing. Association for Computational Linguistics, Seattle, Washington, USA, pp 1631–1642. <https://www.aclweb.org/anthology/D13-1170>
39. Turney PD (2002) Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: ACL, pp 417–424. <http://www.aclweb.org/anthology/P02-1053.pdf>
40. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. In: Proceedings of the 31st international conference on neural information processing systems, NIPS'17. Curran Associates Inc., USA, pp 6000–6010. <http://dl.acm.org/citation.cfm?id=3295222.3295349>
41. Wilson T, Hoffmann P, Somasundaran S, Kessler J, Wiebe J, Choi Y, Cardie C, Riloff E, Patwardhan S (2005) OpinionFinder: a system for subjectivity analysis. In: Proceedings of HLT/EMNLP on interactive demonstrations. Association for Computational Linguistics, pp 34–35
42. Zaharias P, Mehlenbacher B (2012) Editorial: exploring user experience (UX) in virtual learning environments. *Int J Hum Comput Stud* 70(7):475–477. <https://doi.org/10.1016/j.ijhcs.2012.05.001>
43. Zhang L, Wang S, Liu B (2018) Deep learning for sentiment analysis: a survey. *Wiley Interdiscip Rev Data Min Knowl Discov* 8(4):e1253

2.5. Publication 5 | E-LEARNING UNIVERSITY EVALUATION THROUGH SENTIMENT ANALYSIS CENTERED ON USER EXPERIENCE DIMENSIONS

Sanchis-Font, R., Castro-Bleda, M., Jorda-Albiñana, M., Lopez-Cuerva, L. (0). DYNA, DYNA-ACELERADO (0). [5 pp]. (2022)

DOI: <https://dx.doi.org/10.6036/10603>

FONT, R. S., BLEDA, M. J. C., ALBIÑANA, M. B. J., & CUERVA, L. L. (2023). EVALUACIÓN DEL APRENDIZAJE EN LÍNEA UNIVERSITARIO MEDIANTE ANÁLISIS DE SENTIMIENTO CENTRADO EN DIMENSIONES DE EXPERIENCIA DE USUARIO. DYNA, 98(2), 147-153.

Publication 5 proposes the basis for a method to automatically evaluate the UX of university online learning platforms by analysing student sentiment, focusing on specific aspects of their e-learning experience. Out of 2,035 online university students surveyed, the opinions of 476 users in Spanish were collected. These comments were processed with the commercial natural language processing tool MeaningCloud to analyse sentiment (positive, negative or neutral) about aspects of their online experience.

The results of this article present a new model that, on the one hand, ontologically classifies categories and aspects of online education with sentiment analysis techniques, and, on the other hand, the model groups these categories according to UX criteria presenting its own classification to facilitate the evaluation of online learning experiences in an accurate and automatic way and adjusted to specific characteristics.

E-LEARNING UNIVERSITY EVALUATION THROUGH SENTIMENT ANALYSIS CENTERED ON USER EXPERIENCE DIMENSIONS

Rosario Sanchis-Font¹, María José Castro-Bleda², Begoña Jordá-Albiñana³ y Luis López-Cuerva⁴

¹Universitat Politècnica de València. Escuela de Doctorado. Investigadora predoctoral. Camino de Vera, s/n. València 46022.

²Universitat Politècnica de València. Instituto Universitario Valenciano de Investigación en Inteligencia Artificial. Catedrática de Universidad, PDI. Camino de Vera, s/n. València 46022.

³Universitat Politècnica de València. Centro de Investigación en Tecnologías Gráficas. Dpto. de Ingeniería Gráfica. Catedrática de Universidad, PDI. Camino de Vera, s/n. València 46022.

⁴Universidad Miguel Hernández de Elche. Departamento Ingeniería de Comunicaciones. Investigador. Avda. Universidad, s/n. Edificio Innova, módulo 3, 03202 Elche (Alicante).

ABSTRACT:

The COVID-19 crisis increased the number of users of university online teaching, enhancing the importance of this learning format. Additionally, ISO 9241-210:2019 standard sets the international standards for the design of products, services and interaction systems from usability, accessibility, and user experience (User eXperience - UX) perspective. Then, in order to design interfaces and learning experiences that include motivations, feelings and needs of end users, it is necessary to previously evaluate the UX of these environments, with less general and/or laborious methods than those that currently exist.

Therefore, this work aims to establish the basis of a method that allows automatically to evaluate the UX of online teaching platforms by analyzing the users' sentiment about specific aspects of their virtual learning experience.

RESUMEN:

La crisis COVID-19 incrementó el número de usuarios de enseñanza online universitaria potenciando la importancia de este formato de aprendizaje. Adicionalmente, la normativa ISO 9241-210:2019 marca los estándares internacionales para el diseño de productos, servicios y sistemas de interacción desde la usabilidad, accesibilidad y la experiencia de usuario (User eXperience - UX). Así, con el fin de diseñar interfaces y experiencias de aprendizaje que incluyan motivaciones, sentimientos y necesidades de los usuarios finales se precisa evaluar previamente la UX de estos entornos, con métodos menos generalistas y/o laboriosos de los que actualmente existen.

Este trabajo pretende establecer las bases de un método que permita evaluar automáticamente la UX de plataformas de enseñanza en línea mediante el análisis del

<p><i>To do this, 2,035 users were surveyed about their online learning experience with a questionnaire and an open text field to give their opinion. The population surveyed were online postgraduate students of the Universitat de València and the Universidad Rey Juan Carlos, and university students of massive open online courses of the Universitat Politècnica de València. The opinions collected in Spanish from 476 students were processed with the commercial sentiment analysis and natural language processing tool MeaningCloud, to analyze the sentiment (positive, negative, or neutral) about aspects of their experience.</i></p> <p><i>The results present a new model that, on the one hand, ontologically classifies categories and aspects of online education with sentiment analysis techniques, and on the other hand, the model groups these categories according to UX criteria, presenting its own classification to facilitate the evaluation of online learning experiences in a concrete and automatic way.</i></p> <p><i>Keywords: user experience, UX, e-learning, virtual learning, sentiment analysis, data mining, MeaningCloud, natural language processing, university online learning, user centered design, UCD, NLP, VLE</i></p>	<p>sentimiento de los usuarios sobre aspectos concretos de su experiencia de aprendizaje virtual. Para ello, se encuestaron a 2.035 usuarios sobre su experiencia de aprendizaje en línea con un cuestionario campo abierto de texto para opinar. La población encuestada eran estudiantes online de posgrados de la Universitat de València y de la Universidad Rey Juan Carlos, y estudiantes universitarios de cursos masivos abiertos en línea de la Universitat Politècnica de València. Las opiniones recopiladas en español de 476 estudiantes se procesaron con la herramienta comercial de procesamiento del lenguaje natural MeaningCloud, para analizar el sentimiento (positivo, negativo o neutro) sobre aspectos de su experiencia.</p> <p>Los resultados presentan un nuevo modelo que, por un lado, clasifican ontológicamente categorías y aspectos de la educación en línea con técnicas de análisis de sentimiento, y por otra parte, el modelo agrupa estas categorías según criterios UX presentando una clasificación propia para facilitar la evaluación de experiencias de aprendizaje online de manera precisa y automática.</p> <p>Palabras clave: experiencia de usuario, aprendizaje virtual, análisis de sentimiento, minería de datos, enseñanza en línea, MeaningCloud, procesamiento lenguaje natural, aprendizaje en línea universitario, diseño centrado en el usuario, UX, DCU, PLN, VLE</p>
--	---

1.- INTRODUCTION

Online learning in higher education is a teaching model increasingly in demand due to the flexibility and personalization of the study system. Distance learning through virtual environments (e-learning) allows learning at the pace, time and place of each student, Berge. [1]. The COVID-19 world crisis triggered the number of new users in university institutions that began to use the online or hybrid teaching format (face-to-face and online simultaneously) with the consequent massive use of ICT tools (Information and Communication Technologies) applied to education, Ntshwarang. [2].

In recent years, human-computer interaction (Human Computer-Interaction -HCI-) in the field of higher education is the subject of research in order to ensure the success of the training of university students and the quality of online learning. In this sense, some factors evaluated in the quality of interactive learning are the faculty, the e-learning system and the educational content according to Alebeisat [3]. The specifications for the quality of HCI systems are detailed in the ISO 9241-210:2019 standard. [4]with a framework on interaction that considers usability, accessibility and user experience based on user-centered design. ISO describes user experience-User eXperience (UX)-as all the emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and achievements that occur before, during and after the use of interactive systems. For Zaharias and Mehlenbacher [5] UX is a multidimensional concept focused on human needs and aspects of beauty, fun, pleasure and personal growth that are experienced in human interaction with the computer. Other authors, such as Rauschenberger et al. [6] define UX as a concept that encompasses both pragmatic (clarity, efficiency and dependence) and hedonic (stimulation and novelty) qualities. Along these lines, Hassenzahl [7] highlights the importance of the pragmatic aspect of interactive products and the hedonic aspect for the design of experiences, understanding the UX concept as a consequence of the user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (complexity, purpose, usability, functionality, etc.), and the context.

Moreover, in 2015, the UN adopted the 2030 Agenda with 17 Sustainable Development Goals (SDGs) to improve the lives of all. Goal 4 aims to ensure inclusive, equitable and quality education and promote lifelong and lifewide learning opportunities for all, including university education. UNESCO's regular report on SDG 4 [8] highlights the need for inclusive and quality ICT tools for learning.

Students access the digital medium with different abilities, age, culture, location, context, browsing behavior, study, socialization and interaction within the same virtual learning environment, Sanchis-Font et al. et al [9]. Therefore, learning environments designed with interfaces that include the motivations, feelings and needs of all end users are required, with inclusive experiences that facilitate a successful learning process through university online training platforms. A deep understanding of the UX of these platforms will enhance the design of environments that meet the functional, aesthetic and emotional characteristics and needs required by users. In order to understand the user's opinion from this perspective, Spallazo [10] brings together the 129 UX evaluation methods for interactive systems, the most prominent of which by several authors, Diaz-Oreiro [11] AtrakkDiff, UEQ or meCUE. Although these UX evaluation tools are very widespread, their application to the field of online education is very limited, as they are laborious to process and not very specific.

On the other hand, Natural Language Processing (NLP) technologies allow to analyze the user's opinion from text and sentiment analysis. Thus, Clarizia [12] proposes to use data mining to analyze the sentiment present through text comments among students in order to allow the teacher to better match the mood of online learners. The authors aim to outline a model that successfully responds to the analysis task, rather than the results of the experience, by obtaining an overall analysis of user sentiment. Depending on the level at which one wishes to treat the text, one can extract the polarity of the whole document, Moraes [13], the polarity of each sentence in Sanchis-Font et al. [14] or the polarity of each aspect appearing in the text. In other research Sanchis-Font et al. [15] apply NLP models with data mining to evaluate such systems automatically in order to know the polarity of each sentence expressed by users of university e-learning platforms, improving automation and efficiency in the UX evaluation process. The results of the research were very encouraging, but only allowed to know the assessment of the experience in a generic way as positive, negative or neutral and still did not show concrete aspects of a virtual learning experience such as the concepts "video", "slides", "teacher", etc.

Other preliminary studies on the application of sentiment analysis techniques to analyze the perceptions of distance learners have also been published, Magayon [16] and Mac Kim [17]. Both works focus on general aspects of the experience without addressing specific characteristics of the online learning process from a user experience (UX) perspective. So, the problem of detecting concrete areas of improvement of the e-learning user experience in an automatic way remains unresolved.

Therefore, this work seeks more efficient and accurate automatic methods and tools for UX evaluation in e-learning contexts with respect to certain aspects associated with the experience.

Currently, there are several commercial NLP tools on the market that analyze sentiment on texts automatically. Thus, MeaningCloud, Google Cloud Natural Language or Microsoft Azure Text Analytics stand out for their widespread use. In this area, we found studies of sentiment analysis of user comments on their experience with commercially available techniques, Zulkifli et al. [18]. These authors analyze the polarity of English customer reviews from Amazon, Yelp and IMDb on social networks using three tools: Python NLTK Text Classification, Myopia and MeaningCloud. The study showed that MeaningCloud is the technique with the highest accuracy at 82.1%. MeaningCloud is a commercial text analytics NLP API and tool, released in 2015 as an evolution of an earlier product called Textalytics. Both tools have been validated by several authors in the study of their performance, Joshi [19], Singh [20], or Zulkifli [18].

In order to evaluate the UX in e-learning, authors such as Zaharias [5], Ovesleová [21] or Mtebe [22] have established a series of UX categories or criteria applied to the field of online education that clarify the guidelines to ensure a user-centered design of

interactive learning platforms and a positive and quality experience, but without developing automated methods. Therefore, taking as a reference the UX characteristics relevant to the e-learning domain defined by these authors, the present study aims to classify the opinions in favor, against or neutral in an automated way focusing on specific characteristics and dimensions of the user experience of online learning in university environments. For this purpose, this research presents the creation and application of an ontology model of UX in e-learning that allows using NPL tools to extract the themes of user comments in Spanish to detect the UX aspects of interactive learning systems associated with three dimensions and their aspects (VLE - Virtual Learning Environment-, Social Connections -Student, Communication and Teacher- and Learning Resources and Tools -Exercise, Assessment, Image, Material and Sound-) and analyzes the polarity of user sentiment for these dimensions being positive, negative or neutral.

In short, the present work offers, for the first time, the basis of an automatic artificial intelligence model that would allow the evaluation of user experience from sentiment analysis limited to university online learning environments and their related aspects, with the ultimate goal of improving the design of the online university student experience.

2.- MATERIALS AND METHODS

2.1.- CREATION OF THE CORPUS BASED ON E-LEARNING USER OPINIONS

For the creation of the corpus, we used the data previously obtained in the research by Sanchis-Font [15] on the experience of using e-learning platforms for graduate and undergraduate courses. The data came from the validated UEQ and UEQ-S user experience questionnaire with an open field for learner observations that was provided to 2,035 users during 2016 and 2018, which were biomedical postgraduate students coming from the Universitat de València (UV) and the Universidad Rey Juan Carlos (URJC), and university students of massive open online courses (Massive Open Online Course -MOOC-) of the Universitat Politècnica de València (UPV). The online training was carried out on three different platforms depending on the university providing the course. Thus, the questionnaires were adapted and integrated into the ad-hoc system called "Conecto" (in Spanish and English) for the UV master's degree courses, the open source Moodle platform (in Spanish) for the URJC postgraduate courses, and the edX platform for the UPV MOOCs. [23].

Of the 2,035 respondents, only 583 users made comments, of which 476 commented in Spanish and 107 in English. For the present study, only the data collected from the opinions of users who responded in Spanish to a free text field with their comments

(476 users) were used. Comments without information or sentences that were not meaningful (e.g., "nothing") were manually cleaned. In this way we obtained 410 useful comments from different users. These opinions formed a corpus consisting of 599 utterances (including phrases and sentences -with verb-). Thus, machine learning models of sentiment analysis would allow us to classify the aspects found in these sentences based on their Positive (P), Neutral (NEU) or Negative (N) polarity. Three examples from the corpus and how they were desired to be classified are shown in Table A of the appended supplementary material.

2.2.- CREATION OF AN ONTOLOGY OF E-LEARNING ASPECTS

The ontology was created based on the corpus found in the user comments (599 statements). The corpus consists of a closed set of texts or data intended for scientific research and the ontology catalogs the variables required for some set of computation and establishes the relationships between them. The authors created the ontology in order to be able to categorize the aspects present in the text and focused on nine categories or aspects related to the main peculiarities of a virtual learning system. Each sentence was classified into one aspect or category. For this purpose, entries or entities in the text were detected. An entry or entity is a word that can appear in the text denoting a category. For example, possible entries for the VLE aspect or category would be course, environment, experience, master, etc.

To select the different ontology entries, at a first stage the inclusion or exclusion criteria were based on the expert knowledge of the task, i.e. the authors' knowledge of which concepts may be the most important for online learning, and the entities that the authors know from their experience in the field were selected, e.g. entities such as "tutor" or "student" were chosen.

In this way, a proprietary ontology of e-learning aspects was created to evaluate the user experience based on the following categories or aspects: VLE, teacher, learner, sound, image, material, exercise, evaluation and communication. For each category, a list of entities or entries related to the e-learning aspect to be analyzed was detailed. The complete description of the ontology itself is detailed in table B of the supplementary material.

Next, an analysis of word frequency and its importance in the text was performed using the default ontology of MeaningCloud and Google Cloud Natural Language. Thus, in a first phase, the frequencies of occurrence of the various words present in the corpus were analyzed together with the metrics returned by the tools, which indicated how

important the words are in the utterances. The non-formal validation of the ontology was given by the high concordance between the a priori knowledge together with the most repeated words that the tools considered most important. There were 248 occurrences of aspects related to the categories of the e-learning ontology itself, as shown in Fig. 2.

2.3.- APPLICATION AND EVALUATION OF PNL TOOL FOR THE ANALYSIS OF SENTIMENT AT ASPECTUAL LEVEL IN E-LEARNING

To analyze the user experience aspects of online learning platforms in our data corpus, consisting of 599 utterances, we applied the commercial NLP MeaningCloud tool, specifically its free text sentiment analysis application programming interface (API) [24]. This is one of the tools on the market for text analysis with data mining that allows the own creation of ontologies through a web interface, the authors being able to select the key aspects associated with the UX evaluation and enter the ontologies used. The software used for sentiment analysis at the aspectual level in e-learning was MeaningCloud, Sentiment Analysis API version 2.1. [25]. The tool returns a file with the information that has to be processed, analyzed and converted to graphs. The processing, analysis and creation of graphs was done with proprietary software that the authors programmed in Python.

The tool uses a polarity scale consisting of six elements, but we have compressed them into the three basic polarities, using proprietary tools developed for the task:

- Negative: it joins the negative polarity N and very negative N+.
- Positive: it links the positive polarity P and very positive polarity P+.
- Neutral: joins polarity NEU and no NONE.

2.4.- CLASSIFICATION OF CATEGORIES OF SENTIMENT ANALYSIS IN UX DIMENSIONS IN E-LEARNING

Subsequently, to automatically evaluate e-learning systems from the UX perspective, we propose the classification and relationship of the categories present in our ontology grouped into three UX dimensions present in the UX e-learning research of Zaharias [5]Ovesleová [21] and Mtebe [22]. This classification of categories is shown in Table I. The three authors largely agree on the importance of three areas: the online teaching platform, social relations, and the availability of resources and didactic tools for the quality of online learning. For this reason, we propose a grouping of the categories of our ontology into three UX dimensions in e-learning, which we will call "**UXEL Dimensions**":

- **VLE (Virtual Learning Environment)** refers to the practical and functional properties of the virtual learning platform and includes only the VLE category.

- **Social Connections** groups together the whole set of tools that enable interpersonal communication and collaboration in e-learning systems. This dimension includes the categories Student, Communication and Teacher.
- **Learning resources and tools** encompasses the whole set of materials and resources used in the learning process, from exams or exercises to videos and their graphic and sound quality. This dimension includes the categories Exercise, Evaluation, Image, Material and Sound.

Own e-learning ontological categories	Categories Zaharias-Pappas Model	Requirements Ovesleová	Mtebe Criteria	UXEL dimension
VLE	Pragmatics	Capturing attention	-	VLE
Professor	Autonomy and relationships	Guided teaching	Collaborative learning	Social connections
Student	Autonomy and relationships	Guided teaching	Collaborative learning	Social connections
Sound	Real learning	Stimulus presentation	Didactic Materials Accessibility	Learning resources and tools
Image	Real learning	Stimulus presentation	Didactic Materials Accessibility	Learning resources and tools
Material	Real learning	Stimulus presentation	Didactic Materials	Learning resources and tools
Exercise	Real learning	Performance achievement. Improve knowledge retention and transfer	Didactic Materials Feedback and evaluation	Learning resources and tools
Evaluation	Pragmatics	Evaluation performance	Feedback and evaluation	Learning resources and tools
Communication	Autonomy and relationships	Guided teaching	Collaborative learning	Social connections

Table I: Relationship between e-learning UX categories and ontological categories and dimensions in online learning platforms ("UXEL Dimensions").

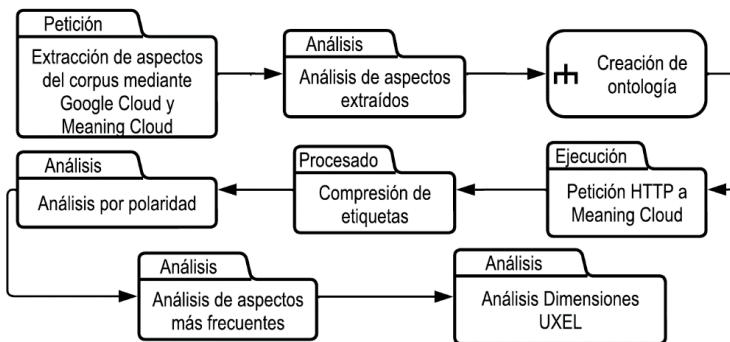


Diagram of the experimentation process at the programmatic level.

3.- RESULTS AND DISCUSSION

3.1.- OWN ONTOLOGY OF E-LEARNING CATEGORIES AND ASPECTS

The ontology defined in this work was based on nine categories that encompass the main peculiarities of an e-learning system. Part of the work has focused on reorganizing the aspects that make up our dataset of the three UX dimensions of online learning (VLE, Learning Resources and Tools, and Social Connections) in order to analyze their polarity using the MeaningCloud tool and evaluate the UX of the analyzed university courses.

As for the distribution by categories, Fig. 2 shows that the most frequent category is VLE, with 115 appearances. The second most commented category is "Image", which has appeared a total of 47 times and the rest of the categories have 25 or less appearances.

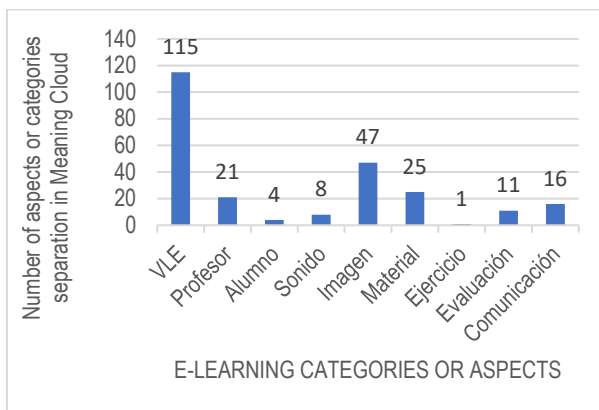


Figure 2. Histogram of frequency of appearance by category in MeaningCloud.

In Fig. 2 we observe how 248 aspects have been classified by MeaningCloud, representing VLE 46.37%, image 18.96%, material 10.09%, teacher 8.47%, communication 6.45%, evaluation 4.43%, sound 3.22%, learner 1.61% and exercise 0.40%.

This means that, among the 599 statements collected in this research belonging to a total of 476 students, MeaningCloud detected 248 aspects or categories of the self-ontology associated with the UXEL dimensions. The most named category was the virtual learning platform (VLE) which appeared 115 times. The second most mentioned UXEL dimension was Learning Resources and Tools with 92 different entries (Exercise,

Assessment, Image, Material and Sound). Aspects related to interpersonal communication and collaboration offered by the university learning experience, grouped in the Social Connections category (Teacher, Student, Communication) were named 41 times. These data are shown in Fig. 4.

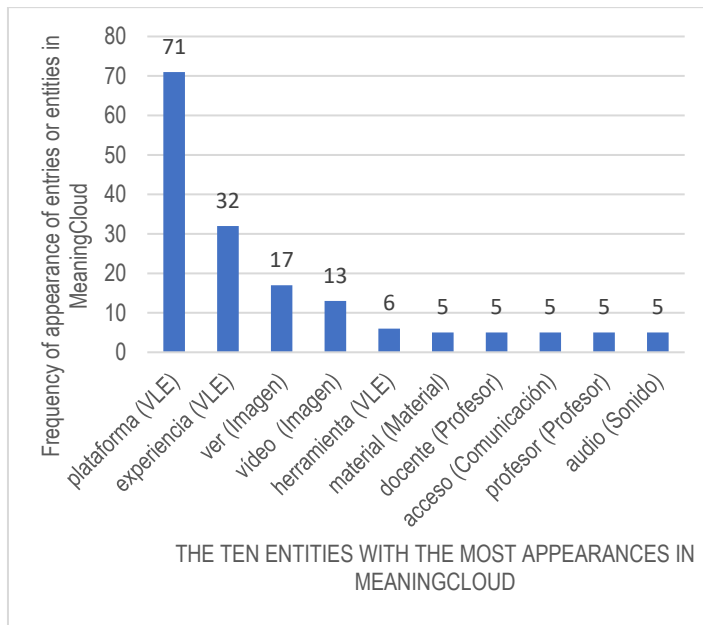


Figure 3. The ten entities or entries with the highest frequency of occurrence detected in MeaningCloud. The data show the number of times each entity has been detected by MeaningCloud and the category to which they belong (in parentheses).

Fig. 3 shows how "platform" is the most detected aspect followed by "experience". It is important to note that the data used come from various virtual learning platforms (Moodle, Ad-hoc -Conecto and edX). However, the results of frequency of most detected entries show the object of greatest interest for students, with the entities "platform", "experience" and "viewing" and "video" being the most mentioned in the data set coming from the three university e-learning platforms.

3.2.- UXEL DIMENSIONS

The total aspect-level polarity for the 248 e-learning aspects rated by MeaningCloud was 113 aspects rated as positive (45.56%), 99 neutral (39.91%) and 36 negative (14.5%), highlighting an overall positive e-learning experience for the surveyed students.

The results of sentiment analysis by UXEL categories and dimensions are shown in Fig. 4 to 7. Fig. 4 shows how MeaningCloud distributes the number of aspects and ranks them in the three dimensions. "Social connections" presents a lower number of aspects, while VLE is the most mentioned dimension.

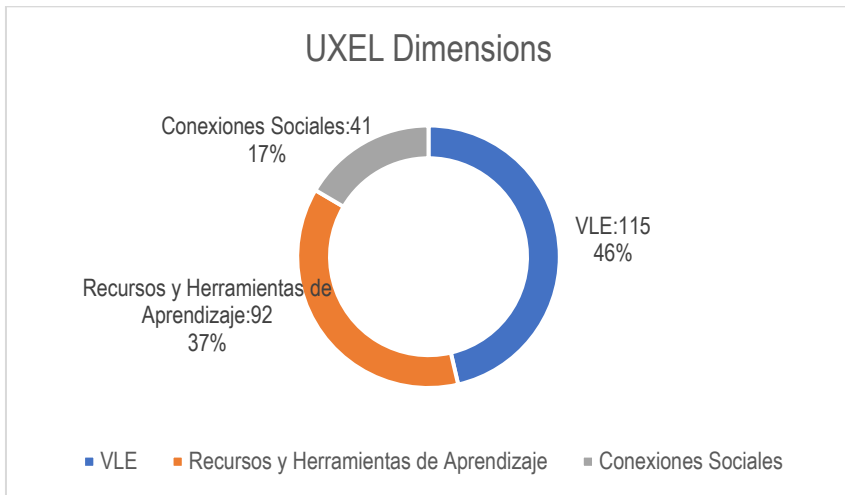


Figure 4: Relative distribution of UXEL dimensions, with the number of entries and their percentage of mention.

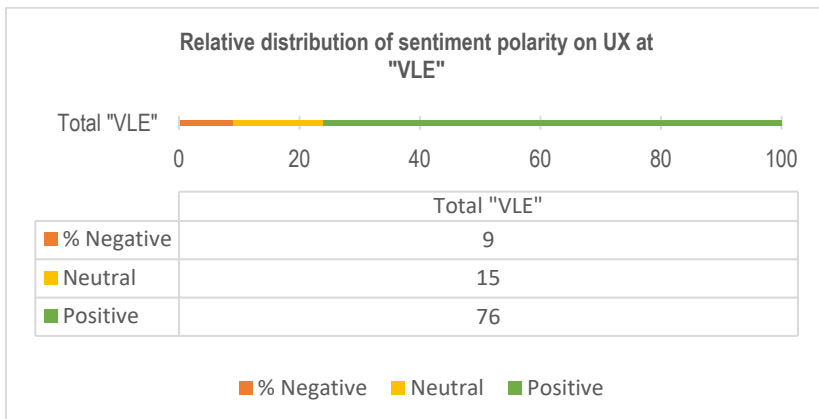


Figure 5: Relative distribution of sentiment of the UX category and dimension "VLE" in percentage values.

Fig. 5 shows the extent to which the comments on the VLE category and the UXEL dimension are positive, neutral or negative. As this figure shows, the virtual learning platform (VLE) of these university courses is the highest rated dimension with 76% of positive opinions. This reinforces the positive polarity results of 45.56% obtained in the overall experience.

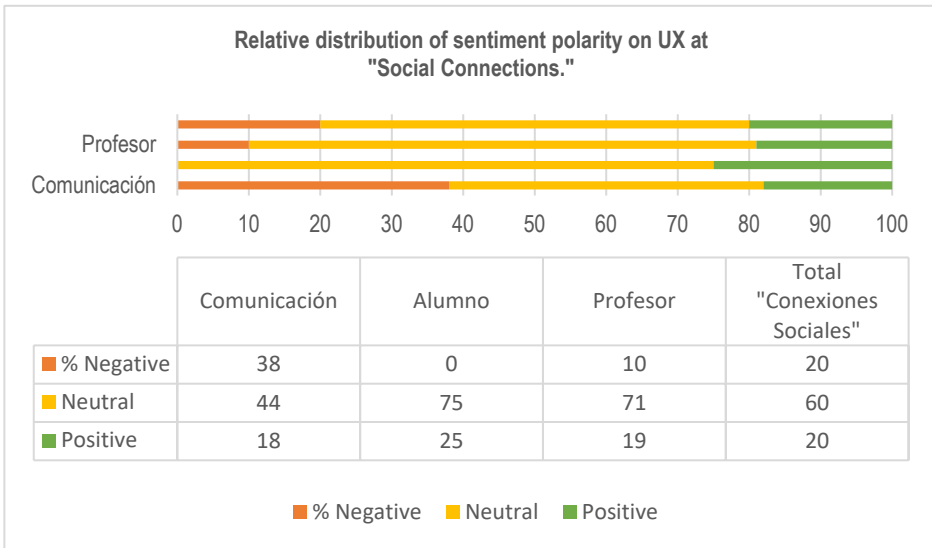


Figure 6: Relative distribution of the categories of the "Social connections" dimension in percentage values.

Fig. 6 shows how the most abundant rating of both the complete dimension and its different categories is neutral. Furthermore, it can be seen that the "Student" category does not have any negative rating and the "Communication" category receives the highest percentage of criticisms, despite this there is a balance between the proportion of positive and negative aspects of the dimension as a whole.

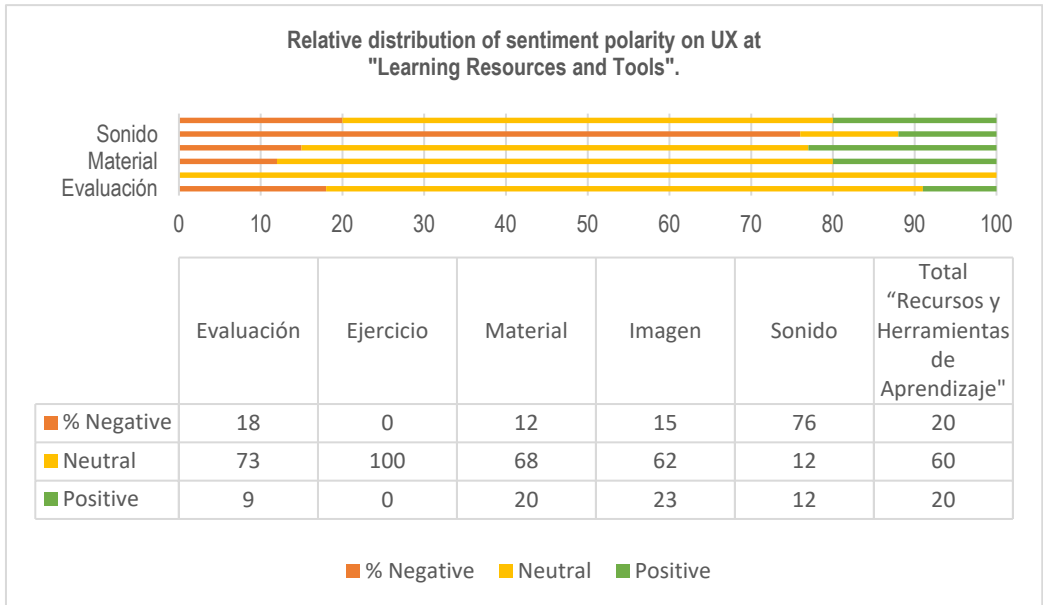


Figure 7: Relative distribution of the categories of the "Resources and Learning Tools" dimension in percentage values.

Analyzing the "Material and learning tools" dimension, we obtain a similar percentage of aspects with positive and negative evaluations. The "Exercise" category stands out, receiving only neutral opinions, and "Sound", which receives 76% of negative reviews, being the category with the worst rating of the three dimensions analyzed.

In summary, the main novelty of this approach consists in the UX evaluation through a proprietary ontology that allows to know the e-learning user experience dimensions automatically with NPL. The model reveals the aspects of improvement or success that directly impact the online experience of university students through three specific dimensions (VLE, Social Connections, and Learning Resources and Tools). In the case of these 476 university students, the results show that they perceived a positive experience when using the virtual environment, but on the other hand they expressed a neutral opinion with respect to the rest of the UXEL dimensions concerning communication tools and interpersonal collaboration or the set of materials and resources used.

4.- CONCLUSIONS

This research has made it possible to automatically evaluate the user experience in university online teaching by means of a proprietary e-learning ontology and the use of

free-to-use commercial tools for sentiment analysis. For the first time, the polarity of student opinions (positive, neutral, negative) has been classified by key categories of digital learning (VLE, Teacher, Student, Sound, Image, Material, Exercise, Assessment, and Communication) and grouped the comments into three UX e-learning or UXEL dimensions (VLE, Social Connections, and Learning Resources and Tools).

This new model provides concrete information on the characteristics of the virtual learning experience that can be improved for this type of university users and which dimensions are valued, as is the case of the dimension "virtual learning environment" which was rated positively by the majority of university students.

The application of sentiment analysis tools focused on aspects of online student opinions from the UX perspective will provide accurate and automatic information allowing to focus on the needs of human-computer interaction to design university e-learning experiences that meet ISO quality standards and ensure UNESCO's SDG 4 with inclusive, equitable and quality education for all.

In future work we intend to improve the model and compare the results between different artificial intelligence tools, include more samples that will allow us to extend the ontology with new aspects of UX e-learning and develop a more accurate, robust and efficient automatic system that investigates user sentiment in order to obtain information that facilitates the design of inclusive and user-centered virtual e-learning environments.

REFERENCES

- [1] BERGE, Zane L. Active, interactive, and reflective elearning. *The Quarterly Review of Distance Education*. 2002. Vol. 3, no. 2, p. 181–190.
- [2] NTSHWARANG, Poloko N., MALINGA, Tumani and LOSIKE-SEDIMO, Nonofu. eLearning Tools at the University of Botswana: Relevance and Use Under COVID-19 Crisis. *Higher Education for the Future*. 2021. Vol. 8, no. 1. <https://doi.org/10.1177/2347631120986281>
- [3] ALEBEISAT, Farhan, ALTARAWNEH, Haroon, ALHALHOULI, Zaid T., QATAWNEH, Ali and ALMAHASNE, Mohmmad. The Impact of Human and Computer Interaction on eLearning Quality. *International Journal of Interactive Mobile Technologies*. 2022. Vol. 16, no. 4, p. 58–69. <https://doi.org/10.3991/ijim.v16i04.27309>
- [4] ISO - ISO 9241-210:2019 - Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. Online. [Accessed 19 April 2022]. <https://www.iso.org/standard/77520.html>

- [5] ZAHARIAS, Panagiotis and MEHLENBACHER, Brad. *Exploring User Experience (UX) in virtual learning environments*. Online. 1 July 2012. Academic Press. ISBN 1071-5819. <https://doi.org/10.1016/j.ijhcs.2012.05.001>
- [6] RAUSCHENBERGER, Maria, SCHREPP, Martin, PEREZ-COTA, Manuel, OLSCHNER, Siegfried and THOMASCHEWSKI, Jorg. Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ). Example: Spanish Language Version. *International Journal of Interactive Multimedia and Artificial Intelligence*. Online. 2013. Vol. 2, no. 1, p. 39. <http://www.ijimai.org/journal/node/418>. DOI: 10.9781/ijimai.2013.215
- [7] HASSENZAHL, Marc and TRACTINSKY, Noam. User experience - a research agenda. *Behaviour & Information Technology*. Online. March 2006. Vol. 25, no. 2, p. 91–97. <https://doi.org/10.1080/01449290500330331>
- [8] EQUIPO DEL INFORME DE SEGUIMIENTO DE LA EDUCACIÓN EN EL MUNDO. *Informe de seguimiento de la educación en el mundo 2021/2: los actores no estatales en la educación: ¿quién elige? ¿quién pierde?* Paris, 2022. [Accessed 20 September 2022]. Available from: https://unesdoc.unesco.org/ark:/48223/pf0000380076_spa
- [9] SANCHIS-FONT, Rosario, JORDÁ-ALBIÑANA, Begoña, GONZÁLEZ-DEL-RÍO, Jimena and AMPUERO-CANELLAS, Olga. IMPROVING THE VIRTUAL LEARNING EXPERIENCE: USER-CENTERED DESIGN IN E-LEARNING. March 2017. p. 9902–9907. Available from: <http://library.iated.org/view/SANCHISFONT2017IMP>. DOI: 10.21125/inted.2017.2356.
- [10] SPALLAZZO, Davide, SCIANNAMÈ, Martina, AJOVALASIT, Marco, CECONELLO, Mauro, VITALI, Ilaria and ARQUILLA, Venanzio. UX Evaluation Methods Mapping. 9 April 2021. [Accessed 25 May 2022]. Available from: https://figshare.com/articles/dataset/UX_Evaluation_Methods_Mapping/14350553/1
- [11] DIÁZ-OREIRO, Ignacio, LÓPEZ, Gustavo, QUESADA, Luis and GUERRERO, Luis A. UX Evaluation with Standardized Questionnaires in Ubiquitous Computing and Ambient Intelligence: A Systematic Literature Review. *Advances in Human-Computer Interaction*. 2021. Vol. 2021. <https://doi.org/10.1155/2021/5518722>
- [12] CLARIZIA, Fabio, COLACE, Francesco, DE SANTO, Massimo, LOMBARDI, Marco, PASCALE, Francesco and PIETROSANTO, Antonio. E-Learning and Sentiment Analysis: A Case Study. 2018. <https://dl.acm.org/doi/10.1145/3178158.3178181>
- [13] MORAES, Rodrigo, João Francisco VALIATI y Wilson P. GAVIÃO NETO. Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications* [en línea]. 2013, 40(2), 621–633. ISSN 0957-4174. <https://doi.org/10.1016/j.eswa.2012.07.059>
- [14] SANCHIS-FONT, Rosario, CASTRO-BLEDA, Maria Jose and ANGEL GONZÁLEZ, José-'. Applying Sentiment Analysis with Cross-Domain Models to Evaluate User eXperience in Virtual Learning Environments. 2052. https://doi.org/10.1007/978-3-030-20521-8_50

- [15] SANCHIS-FONT, Rosario, CASTRO-BLEDA, Maria Jose, JOSÉ-ÁNGEL GONZÁLEZ, FERRAN PLA, and HURTADO, Lluís-F. Cross-Domain Polarity Models to Evaluate User eXperience in E-learning. *Neural Processing Letters*. 2021. Vol. 53, p. 3199–3215. <https://doi.org/10.1007/s11063-020-10260-5>
- [16] MAGAYON, MA. VICTORIA CASTILLO, SACCUAN, Reynaldo and CARBONELL, Aileen. Expectation vs. Reality: A Sentiment Analysis of Students' Experience on Distance Learning. *International Journal of Learning and Teaching*. 2021. Vol. 13, no. 4. <https://doi.org/10.18844/ijlt.v13i4.5979>
- [17] MACKIM, Sunghwan and CALVO, Rafael A. Sentiment analysis in student experiences of learning. In : *Educational Data Mining 2010 - 3rd International Conference on Educational Data Mining*. 2010. <http://dx.doi.org/10.7238/rusc.v12i3.2515>
- [18] ZULKIFLI, Nor Saradatul Akmar and LEE, Allen Wei Kiat. Sentiment Analysis in Social Media Based on English Language Multilingual Processing Using Three Different Analysis Techniques. In : *Communications in Computer and Information Science*. 2019. https://doi.org/10.1007/978-981-15-0399-3_30
- [19] JOSHI, Omkar Sunil and SIMON, Garry. Sentiment Analysis Tool on Cloud: Software as a Service Model. *2018 International Conference On Advances in Communication and Computing Technology, ICACCT 2018*. Online. 9 November 2018. P. 459–462. <http://dx.doi.org/10.1109/icacct.2018.8529649>
- [20] SINGH, Prabhsimran, SAWHNEY, Ravinder Singh and KAHLON, Karanjeet Singh. Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. *ICT Express*. 1 September 2018. Vol. 4, no. 3, p. 124–129. <https://doi.org/10.1016/j.ict.2017.03.001>
- [21] OVESLEOVÁ, Hana. User-Interface Supporting Learners' Motivation and Emotion: A Case for Innovation in Learning Management Systems. In : Online. Springer, Cham, 2016. p. 67–75. https://doi.org/10.1007/978-3-319-40355-7_7
- [22] MTEBE, Joel S. and KISSAKA, Mussa M. Heuristics for evaluating usability of Learning Management Systems in Africa. *2015 IST-Africa Conference, IST-Africa 2015*. 11 August 2015. DOI <http://dx.doi.org/10.1109/istafrica.2015.7190521>
- [23] SANCHIS FONT, Rosario, CASTRO BLEDA, María José, TURRÓ RIBALTA, Carlos and DESPUJOL ZABALA, Ignacio. Integración del “User Experience Questionnaire Short” en MOOCs UPV*. *IN-RED 2018. IV Congreso Nacional de Innovación Educativa y Docencia en Red*. Online. 26 September 2018. P. 393–404. Available from: <https://riunet.upv.es/handle/10251/112803> DOI 10.4995/INRED2018.2018.8840.
- [24] Meaning Cloud Analítica de Texto. <https://www.meaningcloud.com/es/soluciones/analitica-de-texto>. Online. [Accessed 12 October 2022]. Available from: <https://www.meaningcloud.com/es/soluciones/analitica-de-texto>
- [25] MEANINGCLOUD. Sentiment Analysis API version 2.1. Online. [Accessed 12 September 2022]. Available from:

<https://learn.meaningcloud.com/developer/sentiment-analysis/2.1/doc/what-is-sentiment-analysis>

ACKNOWLEDGMENTS

To thank the institutions Fundación IVI and Medigene Press S.L. for the data obtained during the research stay of Rosario Sanchis-Font with the aim of investigating the user experience of the online postgraduate university training provided by both organizations. Also, we would like to thank the UPVX team of the Information and Communication Systems Area of the Universitat Politècnica de València for their collaboration in the implementation of the questionnaire and obtaining data on the user experience of the MOOCs UPV courses.

SUPPLEMENTARY MATERIAL

Phrase	Aspect or Category	Polarity
The experience is going well.	VLE	Positive
It would seem better to me that the exams do not have a time limit to be taken, since being an online course one should be able to take the classes according to one's availability.	EVALUATION COURSE MATERIAL	Negative Positive Negative
I really don't have much practice with using the platform .	VLE	Neutral

Table A. Examples of the classification corpus for the comments of the surveyed e-learning users (VLE, Virtual Learning Environment).

CATEGORÍAO ASPECTO	VLE	PROFESOR	ALUMNO	SONIDO	IMAGEN	MATERIAL	EJERCICIO	EVALUACIÓN	COMUNICACIÓN
ENTRADA O ENTIDAD	aplicación curso entorno experiencia herramienta interfaz máster navegador plataforma posgrado	docente docencia enseñanza instructor maestro ponente presentador profesor tutor	alumno estudiante integrante participante	audio audiovisual entender escuchar oír sonido volumen	apuntador audiovisual cámara dispositivo_ñeñalador enfocar enfoque filmación filmado HD imagen láser material_audiovisual mostrar pantalla pantallazo powerpoint ppt proyectar puntero resolución reproducir señalador señalar slide (ENG) ver vídeo	apuntes archivo clase conferencia contenido charla disertación explicación exposición diapositiva lección material módulo presentación recursos tema transparencia	actividad cuestionario ejercicio encuesta entregable entrega preguntas preguntar respuesta tarea	calificación certificado corrección corregir evaluación examen pruebas revisión test nota	acceso aviso anuncio chat comunicación correo correo-electrónico email foro interacción mail mensaje notificación post

Table B. Aspects detected with domain knowledge and ontologies in the classification corpus for the comments of the surveyed e-learning users (VLE, Virtual Learning Environment).

3. DISCUSSION

The contributions made in this thesis consist of five milestones marked by the chapters of this research work. These findings delimit the basis of a method that allows the automatic evaluation of the UX of university online learning platforms through the analysis of student sentiment, focusing on specific aspects of their virtual learning experience.

In a first stage, it has been presented an overview of different methodologies and features to design a user experience-oriented e-learning that integrates aspects of user-centred design.

In a second phase, the work provides the adapted integration of the User Experience Questionnaire Short (UEQ-S) (Schrepp et al., 2017) on the evaluation questionnaire for MOOCs users of the Universitat Politècnica de València (UPV). Thus, the integration of the UEQ-S into the general student evaluation questionnaire has provided new data to investigate the interaction with the platform in order to enhance and improve the future of the learning experience in UPV MOOCs. This application was pioneer in integrating and adapting the short version of the UEQ questionnaire to Spanish (Rauschenberger, 2013) and integrating it into MOOCs environments. Currently, it continues to be a pioneering work, as none of the other integrations that have been made of the UEQ-S to date have been applied in e-learning environments, despite they have been applied in educational environments (Astuti et al. 2021) to assess other commercial applications such as Google Maps (<https://www.ueq-online.org/>).

In a third phase, and after surveying the population, data were obtained on the experience of using e-learning platforms for postgraduate and university courses. The data came from the validated UEQ and UEQ-S user experience questionnaire with an open field for student observations provided to 2,035 users from 2016 to 2018, who were biomedical postgraduate students from the Universitat de València (UV) and the Universidad Rey Juan Carlos (URJC), and university students of Massive Open Online Courses (MOOC) from the Universitat Politècnica de València (UPV). The online training evaluated was carried out on three different learning management systems (LMS) depending on the university providing the course. Thus, the questionnaires

were adapted and integrated into the ad-hoc system called "Conecto" (in Spanish and English) for the UV master's degrees, into the open-source Moodle platform (in Spanish) for the URJC postgraduate courses, and into the edX platform for the UPV MOOCs. The great diversity of the population evaluated a priori, either by the content of the studies (biomedical postgraduate courses and different disciplines of the UPV MOOCs - See UPV News-), the type of platforms (Moodle, ad-hoc LMS and edX) and the different type of universities that deliver the courses it guarantees a greater heterogeneity of responses received.

In publication 3 we investigate with data from a preliminary population of a smaller size, only taking students from the biomedical postgraduate programmes of the UV and the URJC. For this, we tested the application of machine learning tools for sentiment analysis to compare the results and their performance according to models, tools and language analysed (English and Spanish).

In this stage, we analysed the sentiment of free text comments from 133 users of postgraduate courses and MOOCs (37 in English and 96 in Spanish) in order to obtain their polarity (positive -P-, negative -N- or neutral -NEU-). We created a corpus by manually labelling user opinions with their sentiment polarity (positive, negative or neutral), for the two languages (English and Spanish), and for the three different semantic levels of decreasing complexity: observation, sentence and meaningful unit.

We used Convolutional Neural Networks (CNN) models trained with a corpus of different domain (Twitter posts for each language) and MeaningCloud models. These models were widely used and recognised by the scientific community, either because they were proposed in national and international NLP competitions such as CNN (Rosenthal, 2017; Martínez-Cámara 2017), or because they originated from R&D in data mining such as MeaningCloud (<https://www.meaningcloud.com>).

The results were very promising as the performance accuracy of the sentiment polarity analysis models for both models were optimal and well above the expected value by chance. The CNN and MeaningCloud models offered an accuracy above 50%

compared to an accuracy of 45% for CNN previously demonstrated by other authors (Ouyang, 2015). The cross-domain models showed better results for this sample of users, and specifically for English-language opinions. Therefore, they would initially be the most recommended for this use.

The CNN models showed in this first stage a better behaviour in English than the MeaningCloud models. But this result could be due to the fact that there were many more training samples for English than for Spanish, so the trained model was able to generalise better. For Spanish, on the other hand, the accuracy of the models was better with MeaningCloud. In this case, the CNN models had been trained with fewer samples in Spanish than in English and the generalisation was worse than MeaningCloud.

These results opened the way for us to continue advancing in the application of sentiment analysis models for the UX evaluation of different university online courses and platforms by including more samples and more NLP models.

In the quest for automatic evaluation of e-learning user experience features our next challenge was to include a larger number of sentiment analysis techniques and a more robust corpus. This is detailed in publication four. In that next study we evaluated the opinions of 583 users (107 English speakers and 476 Spanish speakers) and their positive, neutral or negative polarity by applying six different models (3 deep neural network models and 3 sentiment analysis commercial tools). The results improved their performance. All models had an accuracy around or above 70%, reaching up to 80% for Spanish opinions in some models. However, while still optimal, the models performed less well in detecting the sentiment polarity of English comments (<57%).

However, these accuracies correspond to those expected from existing state-of-the-art sentiment analysis models trained with Twitter for binary positive/negative combinations (Wang, 2012). Thus, the evaluation metrics of these models were very similar to when we applied them with our UX dataset, which also features three classes of sentiment polarity (positive, negative and neutral).

With these results and this sample, we were able to move forward with the objective of analysing the feeling of characteristic aspects in e-learning in university environments using artificial intelligence techniques.

Therefore, the last milestone of the research, which is described in publication five, consisted of defining e-learning aspects related to user experience dimensions for evaluating university students' perceptions using sentiment analysis techniques.

Therefore, it is presented the basis for a method to automatically evaluate the UX of university online learning platforms by analysing students' feelings, focusing on specific aspects of their e-learning experience. Key categories of digital learning (VLE, Teacher, Student, Sound, Image, Material, Exercise, Evaluation and Communication) are extracted from the corpus of student comments and these comments are grouped into three UX e-learning or UXEL dimensions (VLE, Social Connections, and Learning Resources and Tools). From 2,035 online university students surveyed, the opinions of 476 users in Spanish were collected. These comments were processed with the commercial natural language processing tool MeaningCloud to analyse the sentiment (positive, negative or neutral) about aspects of their online experience.

The results present a procedure that, on the one hand, ontologically classifies categories and aspects of online education with sentiment analysis techniques, and, on the other hand, the model groups these categories according to UX criteria. This procedure represents a scientific breakthrough by presenting a proprietary classification to facilitate the evaluation of online learning experiences in an accurate and automatic way and adjusted to specific characteristics of university students.

The results presented in this thesis offer an original evaluation model that will provide the keys for the design of attractive, quality, inclusive and user-centred e-learning university platforms, mainly for the Spanish-speaking community.

4. CONCLUSIONS / CONCLUSIONES

4.1. General conclusions

In this chapter we recapitulate the main conclusions reached with this doctoral thesis.

- 1) There are specific characteristics in user experience on university e-learning environments, which are expressed by free text and obtained without a closed questionnaire, that allow it to be automatically analyzed to find out the polarity of a student sentiment (positive, negative and neutral).

- 2) There is an ontology detailing the concrete characteristics by which it is possible to classify and measure the polarity of students' opinions (positive, neutral, negative) through artificial intelligence methods. These features are key categories of university on-line learning: Virtual Learning Environment or VLE, Teacher, Student, Sound, Image, Material, Exercise, Assessment and Communication).

- 3) Based on these user experience characteristics, students' opinions can be measured with sentiment analysis techniques in three UX e-learning or UXEL dimensions: VLE, Social Connections, and Learning Resources and Tools.

- 4) There are artificial intelligence machine learning tools and models that allow to analyze with an accuracy between 57% and 80% the polarity of the sentiment of the users of university e-learning environments. This polarity is assessed on the basis of the opinions expressed by the students through free text, in both English and Spanish. These models and tools are: MeaningCloud, Google Cloud, Microsoft Text Analytics, Convolutional Neural Networks, Transformer Encoders and Attentional BLSTM.

- 5) The application of sentiment analysis tools focused on aspects of online student opinions from the UX perspective provides accurate and automatic information to focus on the needs of human-computer interaction to design university e-learning experiences that meet ISO quality standards and that guarantee UNESCO SDG 4 with an inclusive, equitable and quality education for all.

4.2. Achievement of Goals

The main objective of this doctoral thesis has been to identify and lay the foundations of the most relevant specific characteristics in the user experience of university e-learning environments that allow the automatic analysis of student sentiment.

This general objective has been carried out progressively throughout the five indexed publications presented in chapter 2:

- **Publication 1** (“IMPROVING THE VIRTUAL LEARNING EXPERIENCE: USER CENTERED DESIGN IN E-LEARNING”) presents a series of methods and characteristics existing in the scientific literature to measure and design an interactive learning experience centered on the student that allows generating good, attractive, engaging, fun and emotional virtual teaching environments.
- **Publication 2** (“INTEGRACIÓN DEL “USER EXPERIENCE QUESTIONNAIRE SHORT” EN MOOCS UPV”) integrates one of the measurement methods detailed in publication 1 in its abbreviated version (UEQ-S) to apply it on students of UPV MOOCs courses. The questionnaire will collect user data that will provide information about the interactive online learning experience and its features.
- In **publications 3 and 4**, (“APPLYING SENTIMENT ANALYSIS WITH CROSS-DOMAIN MODELS TO EVALUATE USER EXPERIENCE IN VIRTUAL LEARNING ENVIRONMENTS” and “CROSS-DOMAIN POLARITY

MODELS TO EVALUATE USER EXPERIENCE IN E-LEARNING”) the UX evaluation is applied to different groups of students from several universities and e-learning platforms. Opinions from more than 500 users are analyzed with six natural language processing techniques for sentiment analysis.

- And, finally, in **publication 5** (“E-LEARNING UNIVERSITY EVALUATION THROUGH SENTIMENT ANALYSIS CENTERED ON USER EXPERIENCE DIMENSIONS”) the user data from publication 4 is taken to process them with more appropriate automatic sentiment analysis techniques. It also identifies certain aspects and categories of the e-learning experience in university environments and measures the polarity of student opinion on these UX characteristics as positive, negative or neutral.

On the other hand, the specific objectives have been reached as follows:

- *Objective 1.2.2.1. Select the characteristics to study in our university population related to user experience in e-learning environments.*
This objective has been achieved in publications 1, 2, 3, 4 and 5.
- *Objective 1.2.2.2. Survey the university population that uses e-learning platforms in reference to the selected characteristics.*
This objective has been achieved in publications 2, 3 and 4.
- *Objective 1.2.2.3. Analyze the sentiment of the experience of university users with different machine learning models and tools in order to choose the most appropriate models and which better fit the reality that we want to investigate.*
This objective has been achieved in publications 3, 4 and 5.
- *Objective 1.2.2.4. Categorize the most relevant characteristics obtained with the most appropriate models of*

automatic learning of natural language processing in reference to the user experience in university e-learning environments.

This goal has been achieved in publication 5.

4.3. Contributions made

This work has generated four major contributions to the scientific community:

- 1) An **adaptation of the validated questionnaire UEQ-S** integrated and adapted to three e-learning platforms for specific postgraduate courses at the Universitat de València and at the Universidad Rey Juan Carlos; and for MOOCs at the Universitat Politècnica de València.
- 2) An **innovative application of sentiment analysis** for the evaluation of the user experience of university online students. Specifically, the novelty of the outcomes focuses on the use of deep neural network models (Convolutional Neural Networks, Transformer Encoder and Attentional Bidirectional Long Short Term Memory) and three commercial sentiment analysis systems (MeaningCloud, Google Cloud and Microsoft Text Analytics) to a new task: analyze the learners comments and classify them according to their polarity in positive, negative or neutral.
- 3) A **proprietary ontology of aspects** for the virtual learning experience associated with UX dimensions. This ontology used with sentiment analysis tools, allows classifying the polarity of student opinions (positive, neutral, negative) by key categories of e-learning (VLE, Teacher, Student, Sound, Image, Material, Exercise, Evaluation and Communication) and group the comments in three dimensions UX e-learning or UXEL (VLE, Social Connections, and Learning Resources and Tools).

4) **Five indexed publications** with scientific impact which are collected in this doctoral thesis. These publications have contributed in knowledge to the scientific community for the advancement of the user experience domain, virtual learning, and natural language processing with sentiment analysis techniques. Proof of this impact is the scientific work that cites us in about fourteen academic publications. However, the following nine **publications citing our scientific contributions** from 2019 to February of 2023 stand out for their scope and impact:

- *Multimodal Orthodontic Corpus Construction Based on Semantic Tag Classification Method.*
Springer
August 2022. Neural Processing Letters 54(11)
DOI: 10.1007/s11063-021-10558-y

- *Microblogging: an online resource to support education and training processes.*
July 2022. Campus Virtuales, 11(2)
DOI: 10.54988/cv.2022.2.1013

- *Shared Dictionary Learning Via Coupled Adaptations for Cross-Domain Classification.*
Springer
July 2022. Neural Processing Letters
DOI: 10.1007/s11063-022-10967-7

- *English Education Tutoring Teaching System Based on MOOC.*
Hindawi
May 2022. Computational Intelligence and Neuroscience 2022(9):1-8
DOI: 10.1155/2022/1563352

- *Towards Understanding of User Perceptions for Smart Border Control Technologies using a Fine-Tuned Transformer Approach.*
March 2022. Proceedings of the Northern Lights Deep Learning Workshop 3
DOI: 10.7557/18.6292

- *Estudio de la experiencia de usuario en los sistemas de gestión del aprendizaje.*
January 2022. IE Revista de Investigación Educativa de la REDIECH 12:e1358
DOI: 10.33010/ie_rie_rediech.v12i0.1358

- *Systematic literature review of sentiment analysis in the Spanish language*
February 2021. Data Technologies and Applications ahead-of-print(ahead-of-print)
DOI: 10.1108/DTA-09-2020-0200

- *Examining User Experience of eLearning Systems using EKhool Learners*
October 2020. Journal of Networking and Communication Systems
DOI: 10.46253/jnacs.v3i4.a4

- *A deep-belief network approach for course scheduling.*
September 2020. Journal of Applied Research on Industrial Engineering, 7
DOI: <https://doi.org/10.22105/jarie.2020.243184>.1185

- *Scientific production in the study of user experience in education: case Web of Science and Scopus*
December 2019 TransInformação 32 (Especial number)
DOI: 10.1590/2318-0889202032e190003

4.4. Further lines of research

We will continue developing the automatic process of evaluation and analysis of the user experience in university e-learning environments. With this approach, our next goal is to obtain more precise information from students and to facilitate the design of inclusive and user-centered online learning environments. Thus, **two future lines of research** are derived from this thesis based on the following objectives:

- 1) To enhance the **sentiment analysis model** with more aspects of the UX experience and associated UXEL dimensions. To include more samples (user opinions) that will allow the ontology to be extended with new aspects of UX e-learning.
- 2) To determine the existence of sociodemographic, context and time **variables** that could affect the user experience on different virtual learning university environments (postgraduate studies and MOOCs).

4.5. Conclusiones en español

a) *Conclusiones generales*

En este capítulo recopilamos las principales conclusiones alcanzadas con la presente tesis doctoral.

- 1) Existen características concretas en la experiencia de usuario de entornos e-learning universitarios que, expresadas mediante texto libre y sin hacer uso de un cuestionario cerrado, permiten ser analizadas automáticamente para conocer la polaridad del sentimiento de los estudiantes (positivo, negativo y neutro).
- 2) Existe una ontología que detalla las características concretas con las que es posible clasificar y medir con métodos de inteligencia artificial la polaridad de las opiniones de los estudiantes (positiva, neutra, negativa). Estas características son categorías clave de la enseñanza universitaria online: Virtual Learning Environment o VLE, Profesor, Alumno, Sonido, Imagen, Material, Ejercicio, Evaluación y Comunicación.
- 3) En base a estas características de la experiencia de usuario se puede medir la opinión de los estudiantes con técnicas de análisis de sentimiento en tres dimensiones UX e-learning o UXEL: VLE, Conexiones Sociales, y Recursos y Herramientas de Aprendizaje.
- 4) Existen modelos y herramientas de aprendizaje automático de inteligencia artificial que permiten analizar con una precisión de entre 57% y 80% la polaridad del sentimiento de los usuarios de entornos e-learning universitarios. Esta

polaridad se evalúa a partir de las opiniones realizadas por los estudiantes mediante texto libre, tanto en idiomas inglés como español. Estos modelos y herramientas son: MeaningCloud, Google Cloud, Microsoft Text Analytics, Convolutional Neural Networks, Transformer Encoders and Attentional BLSTM.

- 5) La aplicación de herramientas de análisis de sentimiento centrada en aspectos de las opiniones de estudiantes en línea desde la perspectiva UX proporciona información precisa y automática permitiendo centrarse en las necesidades de la interacción persona-ordenador. Esta información permite diseñar experiencias e-learning universitarias que cumplan con los estándares de calidad ISO y que garanticen el ODS 4 de la UNESCO con una educación inclusiva, equitativa y de calidad para todos.

b) Cumplimiento de los objetivos

El **principal objetivo** de esta tesis doctoral ha sido el de *identificar y asentar las bases de las características más relevantes en la experiencia de usuario de entornos e-learning universitarios que permitan analizar automáticamente el sentimiento de los estudiantes.*

Este objetivo general se ha llevado a cabo de modo progresivo a lo largo de las cinco publicaciones indexadas presentadas en el capítulo 2:

- La **publicación 1** plantea una serie de métodos y características existentes en la literatura científica para la medir y diseñar una experiencia interactiva de aprendizaje centrada en el estudiante que permita generar entornos de enseñanza virtual buenos, atractivos, vinculantes, divertidos y emocionales.

- En la **publicación 2** se integra uno de los métodos de medición detallados en la publicación 1 en su versión abreviada (UEQ-S) para aplicarlo en los estudiantes de los cursos MOOCs UPV. El cuestionario recogerá datos de los usuarios que facilitará información sobre la experiencia interactiva de aprendizaje online y sus características.
- En las **publicaciones 3 y 4** se aplica la evaluación UX a diferentes grupos de estudiantes de diferentes universidades y plataformas e-learning. Las opiniones de más de 500 usuarios son analizadas con seis técnicas del procesamiento del lenguaje natural para el análisis del sentimiento.
- Y, finalmente, en la **publicación 5** se toman los datos de los usuarios de la publicación 4 para procesarlos con técnicas automáticas de análisis de sentimiento más adecuadas. Así mismo, se identifican determinados aspectos y categorías de la experiencia de aprendizaje e-learning en entornos universitarios y se mide la polaridad de la opinión de los estudiantes sobre estas características UX como positiva, negativa o neutra.

Por otra parte, los **objetivos específicos** se han cumplido de la siguiente manera:

- *Objetivo 1.2.2.1. Seleccionar las características a estudiar en nuestra población universitaria relacionadas con la experiencia del usuario en entornos e-learning.*
Este objetivo se ha logrado en las publicaciones 1, 2, 3, 4 y 5.
- *Objetivo 1.2.2.2. Encuestar a la población universitaria usuaria de las plataformas e-learning en referencia a las características seleccionadas.*
Este objetivo se ha logrado en las publicaciones 2, 3 y 4.
- *Objetivo 1.2.2.3. Analizar el sentimiento de la experiencia de los usuarios universitarios con diferentes modelos de aprendizaje automático para elegir los modelos más adecuados y que mejor se ajustan a la realidad que queremos investigar.*

Este objetivo se ha alcanzado en las publicaciones 3, 4 y 5.

- *Objetivo 1.2.2.4. Categorizar las características más relevantes obtenidas con los modelos más adecuados de aprendizaje automático del procesamiento del lenguaje natural en referencia a la experiencia del usuario universitario en entornos e-learning.*

Este objetivo se ha conseguido en la publicación 5.

c) Aportaciones realizadas

El presente trabajo ha generado a la comunidad científica cuatro grandes aportaciones:

- 1) Una **adaptación del cuestionario validado UEQ-S** integrado y adaptado a tres plataformas e-learning para determinados cursos de posgrado de la Universitat de València y de la Universidad Rey Juan Carlos; y para los MOOCs de la Universitat Politècnica de València.
- 2) Una **innovadora aplicación de métodos de análisis de sentimiento** para la evaluación de la experiencia de usuario de estudiantes on-line universitarios. En concreto, la novedad se centra en el uso de los modelos de redes neurales profundas (Convolutional Neural Networks, Transformer Encoder y Attentional Bidirectional Long Short Term Memory) y tres sistemas comerciales de análisis de sentimiento (MeaningCloud, Google Cloud y Microsoft Text Analytics) para una nueva tarea: analizar los comentarios de los estudiantes y clasificarlos según su polaridad en positivo, negativo o neutro.
- 3) Una **ontología propia** de aspectos de la experiencia de aprendizaje virtual asociadas a dimensiones UX. Esta ontología utilizada con herramientas de análisis de sentimiento permite clasificar la polaridad de las opiniones de los estudiantes (positiva, neutra, negativa) por categorías claves de la enseñanza digital (VLE, Profesor, Alumno, Sonido, Imagen, Material, Ejercicio, Evaluación y Comunicación) y agrupar los comentarios en tres

dimensiones UX e-learning o UXEL (VLE, Conexiones Sociales, y Recursos y Herramientas de Aprendizaje).

4) **Cinco publicaciones indexadas y de impacto** científico, que se recogen en esta tesis doctoral. Estas publicaciones han aportado conocimiento a la comunidad científica para el progreso del dominio de la experiencia usuario, del aprendizaje virtual y de procesamiento del lenguaje natural con técnicas de análisis del sentimiento. Muestra de este impacto son los trabajos científicos que nos citan en unas catorce publicaciones académicas. Pero destacamos por su alcance e impacto las siguientes **nueve publicaciones que citan** nuestras contribuciones científicas desde 2019 hasta febrero de 2023:

- *Multimodal Orthodontic Corpus Construction Based on Semantic Tag Classification Method.*
Springer
August 2022. Neural Processing Letters 54(11)
DOI: 10.1007/s11063-021-10558-y
- *Microblogging: an online resource to support education and training processes.*
July 2022. Campus Virtuales, 11(2)
DOI: 10.54988/cv.2022.2.1013
- *Shared Dictionary Learning Via Coupled Adaptations for Cross-Domain Classification.*
Springer
July 2022. Neural Processing Letters
DOI: 10.1007/s11063-022-10967-7
- *English Education Tutoring Teaching System Based on MOOC.*
Hindawi
May 2022. Computational Intelligence and Neuroscience 2022(9):1-8
DOI: 10.1155/2022/1563352

- *Towards Understanding of User Perceptions for Smart Border Control Technologies using a Fine-Tuned Transformer Approach.*
March 2022. Proceedings of the Northern Lights Deep Learning Workshop 3
DOI: 10.7557/18.6292

- *Estudio de la experiencia de usuario en los sistemas de gestión del aprendizaje.*
January 2022. IE Revista de Investigación Educativa de la REDIECH 12:e1358
DOI: 10.33010/ie_rie_rediech.v12i0.1358

- *Systematic literature review of sentiment analysis in the Spanish language*
February 2021. Data Technologies and Applications ahead-of-print(ahead-of-print)
DOI: 10.1108/DTA-09-2020-0200

- *Examining User Experience of eLearning Systems using EKhool Learners*
October 2020. Journal of Networking and Communication Systems
DOI: 10.46253/jnacs.v3i4.a4

- *A deep-belief network approach for course scheduling.*
September 2020. Journal of Applied Research on Industrial Engineering, 7
DOI: <https://doi.org/10.22105/jarie.2020.243184.1185>

- *Scientific production in the study of user experience in education: case Web of Science and Scopus*
December 2019 TransInformação 32 (Especial number)
DOI: 10.1590/2318-0889202032e190003

d) Líneas de investigación futuras

Continuaremos desarrollando el proceso automático de evaluación y análisis de experiencia de usuario en entornos e-learning universitarios. Con ello, nuestro siguiente reto es obtener información de los estudiantes y facilitar el diseño de entornos virtuales de enseñanza en línea inclusivos y centrados en el usuario. Así, de esta tesis se derivan **dos líneas de investigación** futuras basándose en los siguientes objetivos:

- 1) Mejorar el **modelo de análisis de sentimiento** con más aspectos de la experiencia UX y dimensiones UXEL asociadas. Incluir más muestras (opiniones de los usuarios) que permitirán ampliar la ontología con nuevos aspectos de la UX e-learning.
- 2) Determinar la **existencia de variables** sociodemográficas, de contexto y tiempo que pudieran afectar a la experiencia de usuario en diversos entornos de universitarios de aprendizaje virtual (estudios de posgrado y MOOCs).

ABBREVIATIONS AND ACRONYMS

API: Application Programming Interface.

CNN: Convolutional Neural Networks

COVID-19: Coronavirus disease 2019 caused by SARS-CoV-2 virus.

edX: American massive open online course (MOOC) provider created by Harvard and MIT. It hosts online university-level courses in a wide range of disciplines to a worldwide student body, including some courses at no charge.

GUX: Global User eXperience

HCI: Human-Computer Interaction

ISO: International Organization for Standardization

ITC: Information and Communication Technologies

JCR: Journal Citations Report

LMS: Learning Management System

LNCS: Lecture Notes in Computer Science

LSTM: Long Short Term Memories

N: Negative polarity

NEU: Neutral polarity

NLP: Natural Language Processing

MOOC: Massive Online Open Course

ONU: Organización de las Naciones Unidas. (See UN).

ODS: Objetivo de Desarrollo Sostenible (See SDG)

P: Positive polarity

SemEval: International Workshop on Semantic Evaluation.

SEPNL: Sociedad Española para el Procesamiento del Lenguaje Natural

SDG: Sustainable Development Goals for 2030 by UN

SJR: SCImago Journal Rank

TASS: Taller de Análisis Semántico en la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)

UCD: User-Centered Design

UEQ: User Experience Questionnaire

UEQ-S: User Experience Questionnaire Short

UN: United Nations

UNESCO: United Nations Educational, Scientific and Cultural Organization

UOC: Universitat Oberta de Catalunya

UPV: Universitat Politècnica de València

URJC: Universidad Rey Juan Carlos

UV: Universitat de València

UX: User eXperience

UXEL: User eXperience en E-Learning.

VLE: Virtual Learning Environment.

REFERENCES (Introduction and Discussion)

A

Adelson, B. (2010). The emotional design of everyday things: The 2006 Benjamin Franklin Medal in computer and cognitive science presented to Donald A. Norman. *Journal of the Franklin Institute*, 347(4), 681-687.

Alebeisat, F., Altarawneh, H., Alhalhouli, Z. T., Qatawneh, A., & Almahasne, M. (2022). The Impact of Human and Computer Interaction on eLearning Quality. *International Journal of Interactive Mobile Technologies*, 16(4).

Ali, S., Uppal, M. A., & Gulliver, S. R. (2018). A conceptual framework highlighting e-learning implementation barriers. *Information Technology & People*, 31(1), 156-180.

Al-Fraihat, D., Joy, M., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in human behavior*, 102, 67-86.

Astuti, I. A., Farida, L. D., & Hidayat, T. (2021, April). Measuring the UX of Mobile Application Attendance Lectures Feature Using Short-User Experience Questions (UEQ-S). In *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)* (pp. 286-291). IEEE.

Al-Fraihat, D., Joy, M., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in human behavior*, 102, 67-86.

B

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.

Bates, A. T. (2005). Technology, e-learning and distance education. Routledge.

Baziotis, C., Pelekis, N., & Doukeridis, C. (2017, August). Dastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis. In Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017) (pp. 747-754).

Brown, T., & Katz, B. (2011). Change by design. Journal of product innovation management, 28(3), 381-383.

C

Clarizia, F., Colace, F., De Santo, M., Lombardi, M., Pascale, F., & Pietrosanto, A. (2018, January). E-learning and sentiment analysis: a case study. In Proceedings of the 6th international conference on information and education technology (pp. 111-118).

Cliche M (2017) BB_twtr at SemEval-2017 task 4: Twitter sentiment analysis with CNNs and LSTMs. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017). Association for Computational Linguistics, Vancouver, Canada, pp 573–580. <https://doi.org/10.18653/v1/S17-2094>. <https://www.aclweb.org/anthology/S17-2094>

Courage, C., & Baxter, K. (2005). *Understanding your users: A practical guide to user requirements methods, tools, and techniques*. Gulf Professional Publishing.

D

Damodaran, L. (1996). User involvement in the systems design process-a practical guide for users. *Behaviour & information technology*, 15(6), 363-377..

de Lera, E., Almirall, M., Fernàndez, C., & Gisbert, M. (2012). Designing innovative open spaces for learning: the case of the Open University of Catalunya. *Procedia-Social and Behavioral Sciences*, 55, 1000-1008.

Devlin J, Chang MW, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional Transformers for language understanding. In: *Proceedings of the 2019 conference of the North American chapter of the Association for Computational Linguistics: Human Language Technologies*, volume 1 (long and short papers). Association for Computational Linguistics, Minneapolis, Minnesota, pp 4171–4186. <https://doi.org/10.18653/v1/N19-1423>. <https://www.aclweb.org/anthology/N19-1423>

Diaz-Galiano MC, et al (2019) Overview of TASS 2019: one more further for the global Spanish sentiment analysis corpus. In: *Proceedings of the Iberian languages evaluation forum (IberLEF 2019)*, CEUR-WS, Bilbao, Spain, CEUR workshop proceedings, pp 550–560

Díaz-Oreiro, I., López, G., Quesada, L., & Guerrero, L. A. (2021). UX evaluation with standardized questionnaires in ubiquitous

computing and ambient intelligence: a systematic literature review. *Advances in Human-Computer Interaction*, 2021, 1-22.

E

EQUIPO DEL INFORME DE SEGUIMIENTO DE LA EDUCACIÓN EN EL MUNDO. Informe de seguimiento de la educación en el mundo 2021/2: los actores no estatales en la educación: ¿quién elige? ¿quién pierde? Paris, 2022. [Accessed 20 September 2022]. Available from: https://unesdoc.unesco.org/ark:/48223/pf0000380076_spa)

F

Fensel, D. (2001). *Ontologies* (pp. 11-18). Springer Berlin Heidelberg.

G

Gena, C. (2006, June). A user-centered approach for adaptive systems evaluation. In *5th Workshop on User-Centred Design and Evaluation of Adaptive Systems* (pp. 20-23).

González J, Pla F, Hurtado L (2018) Elirf-upv en TASS 2018: Análisis de sentimientos en twitter basado en aprendizaje profundo (elirf-upv at TASS 2018: sentiment analysis in Twitter based on deep learning). In: *Proceedings of TASS 2018: workshop on semantic analysis at SEPLN, TASS@SEPLN 2018, collocated with 34nd SEPLN conference (SEPLN 2018), Sevilla, Spain, September 18th, 2018*, pp 37–44. http://ceur-ws.org/Vol-2172/p2_elirf_tass2018.pdf

González-Barba, J. Á., Arias-Moncho, J., Hurtado Oliver, L. F., & Pla Santamaría, F. (2020, September). *Elirf-upv at tass 2020*:

Twilbert for sentiment analysis and emotion detection in spanish tweets. In Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2020) (pp. 179-186). CEUR. https://www.ceur-ws.org/Vol-2664/tass_paper2.pdf

Gonzalez, J. A., Hurtado, L. F., & Pla, F. (2021). TWilBert: Pre-trained deep bidirectional transformers for Spanish Twitter. *Neurocomputing*, 426, 58-69.

Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., & Cajander, Å. (2003). Key principles for user-centred systems design. *Behaviour and Information Technology*, 22(6), 397-409.

H

Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality. In *Mensch & computer* (Vol. 57, pp. 187-196). Heifelberg: Springer.

Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

I

ISO - ISO 9241-210:2019 - Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems. Online. [Accessed 19 April 2022]. <https://www.iso.org/standard/77520.html>

J

Joshi, O. S., & Simon, G. (2018, February). Sentiment analysis tool on cloud: software as a service model. In 2018 International Conference On Advances in Communication and Computing Technology (ICACCT) (pp. 459-462). IEEE.

K

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2022). Natural language processing: State of the art, current trends and challenges. *Multimedia tools and applications*, 1-32.

Kim Y (2014) Convolutional neural networks for sentence classification. In: *Proceedings of the 2014 conference on empirical methods in natural language processing, EMNLP 2014, October 25–29, 2014, Doha, Qatar, a meeting of SIGDAT, a special interest group of the ACL*, pp 1746–1751. <http://aclweb.org/anthology/D/D14/D14-1181.pdf>

Kujala S, Roto V, Väänänen-Vainio-Mattila K, Karapanos E, Sinnelä A (2011) UX curve: a method for evaluating long-term user experience. *Interact Comput* 23(5):473–483

L

Law, E. L. C., Van Schaik, P., & Roto, V. (2014). Attitudes towards user experience (UX) measurement. *International Journal of Human-Computer Studies*, 72(6), 526-541.

Liu B (2012) *Sentiment analysis and opinion mining. A comprehensive introduction and survey*. Morgan & Claypool Publishers, San Rafael

Liu, B., Hu, M., & Cheng, J. (2005, May). Opinion observer: analyzing and comparing opinions on the web. In Proceedings of the 14th international conference on World Wide Web (pp. 342-351).

M

Mac Kim, S., & Calvo, R. A. (2010, June). Sentiment Analysis in Student Experiences of Learning. In EDM (pp. 111-120).

Magayon, M. A. V. C., Saccuan, R., & Carbonell, A. (2021). Expectation vs. Reality: A Sentiment Analysis of Students' Experience on Distance Learning. *International Journal of Learning and Teaching*, 13(4), 260-275.

Martinez-Camara, E., Diaz-Galiano, M. C., Garcia-Cumbreras, M. A., Garcia-Vega, M., & Villena-Román, J. (2017). Overview of TASS 2017. *Proceedings of TASS*, 13-21.

Mejova, Y. (2009). Sentiment analysis: An overview. University of Iowa, Computer Science Department.

Moraes, R., Valiati, J. F., & Neto, W. P. G. (2013). Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, 40(2), 621-633.

Mtebe, J. S., & Kissaka, M. M. (2015, May). Heuristics for evaluating usability of learning management systems in Africa. In 2015 IST-Africa Conference (pp. 1-13). IEEE.

Muriel, G. D., & Enric, M. (2007). User centered design in e-learning environments: from usability to learner experience. In Proceedings of the EDEN 2007 annual conference, Naples, Italy, June (pp. 13-16).

N

Ntshwarang, P. N., Malinga, T., & Losike-Sedimo, N. (2021). eLearning tools at the University of Botswana: Relevance and use under COVID-19 crisis. *Higher Education for the Future*, 8(1), 142-154.

Norman D. A. & Draper S. W. (Eds.) (1986) *User centered system design: New perspectives on human-computer interaction*. Hillsdale, NJ, Lawrence Erlbaum Associates.

Noticia UPV: Dos MOOC de la UPV, entre los 100 con más inscripciones de todos los tiempos a nivel mundial | Universitat Politècnica de València. Available at: <http://www.upv.es/noticias-upv/noticia-12013-ranking-class-es.html> (Accessed: February 19, 2023).

O

Ovesleová, H. (2016). User-Interface Supporting Learners' Motivation and Emotion: A Case for Innovation in Learning Management Systems. In *Design, User Experience, and Usability: Novel User Experiences: 5th International Conference, DUXU 2016, Held as Part of HCI International 2016, Toronto, Canada, July 17–22, 2016, Proceedings, Part II 5* (pp. 67-75). Springer International Publishing.

Ouyang, X., Zhou, P., Li, C. H., & Liu, L. (2015, October). Sentiment analysis using convolutional neural network. In 2015

IEEE international conference on computer and information technology; ubiquitous computing and communications; dependable, autonomic and secure computing; pervasive intelligence and computing (pp. 2359-2364). IEEE.

P

Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 conference on empirical methods in natural language processing, vol 10. Association for Computational Linguistics, pp 79–86

R

Rauschenberger, M., Schrepp, M., Pérez Cota, M., Olschner, S., & Thomaschewski, J. (2013). Efficient measurement of the user experience of interactive products. How to use the user experience questionnaire (UEQ). Example: Spanish language version.

Rosenthal S, Farra N, Nakov P (2017) SemEval-2017 task 4: sentiment analysis in Twitter. In: Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017). Association for Computational Linguistics, Vancouver, Canada, pp 502–518. <https://doi.org/10.18653/v1/S17-2088>. <https://www.aclweb.org/anthology/S17-2088>

S

Sadr H, Pedram MM, Teshnehlab M (2019) A robust sentiment analysis method based on sequential combination of convolutional and recursive neural networks. *Neural Process Lett* 50:2745–2761.

Schaik, P. V., & Ling, J. (2012). An experimental analysis of experiential and cognitive variables in web navigation. *Human-Computer Interaction*, 27(3), 199-234.

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4 (6), 103-108.

Singh, P., Sawhney, R. S., & Kahlon, K. S. (2018). Sentiment analysis of demonetization of 500 & 1000 rupee banknotes by Indian government. *ICT Express*, 4(3), 124-129.

Socher R, Perelygin A, Wu J, Chuang J, Manning CD, Ng A, Potts C (2013) Recursive deep models for semantic compositionality over a sentiment treebank. In: Proceedings of the 2013 conference on empirical methods in natural language processing. Association for Computational Linguistics, Seattle, Washington, USA, pp 1631–1642. <https://www.aclweb.org/anthology/D13-1170>

Squires, D. (1999). Usability and educational software design: special issue of interacting with computers. *Interacting with computers*, 11(5), 463-466.

T

Turney PD (2002) Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In: ACL, pp 417–424. <http://www.aclweb.org/anthology/P02-1053.pdf>

U

UPV News: See “Noticia UPV”.

V

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. In: Proceedings of the 31st international conference on neural information processing systems, NIPS'17. Curran Associates Inc., USA, pp 6000–6010. <http://dl.acm.org/citation.cfm?id=3295222.3295349>

W

Wang, J., Yu, L. C., Lai, K. R., & Zhang, X. (2016, August). Dimensional sentiment analysis using a regional CNN-LSTM model. In Proceedings of the 54th annual meeting of the association for computational linguistics (volume 2: Short papers) (pp. 225-230).

Wilson T, Hoffmann P, Somasundaran S, Kessler J, Wiebe J, Choi Y, Cardie C, Riloff E, Patwardhan S (2005) OpinionFinder: a system for subjectivity analysis. In: Proceedings of HLT/EMNLP on interactive demonstrations. Association for Computational Linguistics, pp 34–35

Z

Zaharias, P., & Mehlenbacher, B. (2012). Exploring User Experience (UX) in virtual learning environments. *International Journal of Human-Computer Studies*, 70(7), 475-477.

Zaharias, P., & Poylymenakou, A. (2009). Developing a usability evaluation method for e-learning applications: Beyond functional usability. *Intl. Journal of Human-Computer Interaction*, 25(1), 75-98.

Zhang L, Wang S, Liu B (2018) Deep learning for sentiment analysis: a survey. *Wiley Interdiscip Rev Data Min Knowl Discov* 8(4):e1253

