

Article

A Review on MAS-Based Sentiment and Stress Analysis User-Guiding and Risk-Prevention Systems in Social Network Analysis

Guillem Aguado * , Vicente Julián , Ana García-Fornes  and Agustín Espinosa 

Valencian Research Institute for Artificial Intelligence (VRAIn), Universitat Politècnica de València, Camino de Vera s/n, 46022 Valencia, Spain; vinglada@dsic.upv.es (V.J.); agarcia@dsic.upv.es (A.G.-F.); aespinos@upvnet.upv.es (A.E.)

* Correspondence: guiagsar@dsic.upv.es

Received: 21 July 2020; Accepted: 22 September 2020; Published: 26 September 2020



Abstract: In the current world we live immersed in online applications, being one of the most present of them Social Network Sites (SNSs), and different issues arise from this interaction. Therefore, there is a need for research that addresses the potential issues born from the increasing user interaction when navigating. For this reason, in this survey we explore works in the line of prevention of risks that can arise from social interaction in online environments, focusing on works using Multi-Agent System (MAS) technologies. For being able to assess what techniques are available for prevention, works in the detection of sentiment polarity and stress levels of users in SNSs will be reviewed. We review with special attention works using MAS technologies for user recommendation and guiding. Through the analysis of previous approaches on detection of the user state and risk prevention in SNSs we elaborate potential future lines of work that might lead to future applications where users can navigate and interact between each other in a more safe way.

Keywords: Multi-Agent System; social networks; sentiment analysis; stress analysis

1. Introduction

Since people are constantly navigating through online applications and Social Network Sites (SNSs) are between the most common of them, several efforts from the research community have been made in trying to prevent risks and aiming to lead to more satisfactory and safe user experience. The question of whether users are at risk or not in SNSs caused by the social interaction has been reviewed in Reference [1], where risks and negative outcomes that arise from this interaction between users in SNSs have been explored. In References [2,3] risk factors are reviewed. Content risks refer to the risks of receiving inappropriate content, which can be of a variety of types (e.g., racism, violence). Contact risks arise from interacting with strangers and refer to the risks that users face from this interaction (e.g., cyber-harassment, privacy issues). Commercial risks are in line with aggressive marketing techniques, that involve spam or getting asked for personal information to be used for commercial affairs for the interest of a third person or organization different from the user that is asked for this kind of data. Moreover, it has been reported in the literature that teenagers face several risks when interacting in a SNS and that said teenagers have characteristics that make them more vulnerable to those risks [4]. Additionally, regret and negative consequences can arise from the fact of publishing a post in a SNS [5].

The decision-making process drives the interaction of users navigating in online social environments. It determines, for example, the information that users share, the users that are contacted, and the kind of interaction that users do with each other. Decision making has been reported to be

affected by the emotional state of the person making the decision. In Reference [6], authors show that incidental moods, discrete emotions, integral affect, and regret have an effect on decision making. Additionally, stress has been associated with a specific emotional state (high arousal and negative valence) and has been used in Reference [7] to construct an adaptation (TensiStrength) of the sentiment strength detection software called SentiStrength [8], for which reason it might be a potential candidate (in addition to sentiment polarities) to offer useful information in a system that attempts to prevent risks in online social environments through user data analysis.

In Reference [9] works on sentiment analysis using text, audio, visual and physiological signals are reviewed, including multi-modal sentiment analysis, which combines different data sources. Applications of sentiment analysis are also commented and future lines of work are highlighted. To the best of our knowledge, there is a lack of a review in user risk prevention when navigating online social environments, therefore in our proposal, works that address this line of work are reviewed and works that use Multi-Agent Systems (MAS) based user recommendation highlighted. This leads to also review works not only in sentiment analysis but in the lines of automatic user state detection, concretely the cases of sentiment analysis and stress analysis, so the state-of-art techniques used for detection can be later linked to prevention approaches in Sections 5 and 6. In those sections, conclusions about how current technologies help prevent issues in online social platforms are drawn and potential new lines of work elaborated. Therefore, the aim of this work is to review the current state-of-art works in risk prevention and recommendation in online social environments using MAS-based approaches, and also review the literature in sentiment and stress detection, for being able to link them and to emphasize potential new lines of work, unexplored at the moment. Consequently, the present survey might help future researchers and developers to build online social platforms that could guide the users and prevent them from suffering negative consequences of interacting, thus leading to a more satisfactory and safe social experience.

The rest of the paper is structured as follows—Section 2 presents the main topics of this literature review. Section 3 reviews a series of recent works in the lines of sentiment and stress analysis. Section 4 reviews works in risk prevention in SNSs, and usage of MAS technologies for user guidance and recommendation. Section 5 makes an overview of the cases of study of automatic user state detection and gives insight on how could risk prevention and user guidance be addressed by detecting the user state in SNSs. Finally in Section 6 future potential lines of work are extracted and developed and conclusions are extracted.

2. Problem Statement

The present work aims to review the current literature in three different topics and to link them together, for extracting potential future lines of work in risk prevention in online social environments. The three topics are user state detection, risk prevention, and recommendation.

- User state detection: refers to the automatic detection of an aspect of the user state by the system. Has many variations, but since in this work we aim to address risk prevention we focus on the detection of the sentiment and stress levels of users, which are sentiment analysis and stress analysis. Those two techniques address the problem of detecting automatically the sentiment and stress level of users by employing different techniques (e.g., machine learning, natural language processing (NLP)) on different data sources (e.g., text, audio, images), either using one data modality or multiple.
- Risk prevention: addresses the prevention of risks that users on a system can suffer. In our case, we focus on the prevention of risks that users can suffer while navigating online social environments, such as SNSs. It can be performed by employing user state detection and applying feedback to users when necessary or giving recommendations to them, which is the focus of the present survey, but it can also be addressed by performing analysis of relations between users and warning users about dangerous people, as an example.

- Recommendation: encompasses the techniques used by a system to give recommendations about different matters to users (e.g., what to buy, when to invest, whom to trust). User state detection can be used by the system to perform recommendations, and recommendations can in turn be used to prevent risks that users could be exposed to in a system.

In the following two sections, and to be able to later draw conclusions on how the existing state-of-art technologies can help risk prevention and user guidance in SNSs, and elaborate potential new lines of work, we will review works related to the user guiding process. Firstly, we start reviewing works on user automatic sentiment polarity and stress level detection. We also review works in Case-Based Reasoning (CBR) based sentiment analysis. Later we review works in risk prevention, and MAS-based recommendation and guiding systems.

3. Detection Approaches Review

In the following subsections, approaches to the detection of the state of the user by analyzing different sources of data and fusion techniques will be reviewed. Performance of the reviewed approaches and techniques used are summarized in Tables 1 and 2, respectively. Datasets used, their characteristics and the partitions used for training and testing are summarized in Tables 3 and 4 (these two later tables will be referred to as the ‘dataset tables’ in the following text of this paper). From the dataset tables, it can be seen that annotated customer reviews about products in e-commerce websites such as Amazon are very useful to construct datasets for sentiment analysis models. Annotated reviews extracted from various sites not related to e-commerce can also be useful for this purpose, and they can be found from very varied sources (e.g., the Internet Movie Database, epinions.com, CitySearch.com). Moreover, in the dataset tables it is shown that when building a dataset, researchers can make use of data streaming services from SNSs, such as Twitter, or download videos and images from online platforms that allow users to share images and videos (e.g., Youtube, Flickr). Then for labeling the data researchers can make use of labeling services such as the one offered by Amazon (Amazon Mechanical Turk or AMT). If there is no need for very large datasets, controlled laboratory experiments with a set of people to generate a dataset are always an option, or reusing existing datasets, as is shown as well in the dataset tables. Additionally, in the dataset tables can be seen that specific requirements such as people reacting to a specific set of emotions, or collecting data under certain conditions of physical and cognitive stress might require to conduct a laboratory experiment, since those might be hard to find and reuse. Stress-related data can be collected online like in the case of sentiment analysis datasets, using, for example, SNSs data.

Firstly, we review sentiment analysis on text data, which is one of the most predominant and well-established lines of work in recent years regarding data analysis for detecting the state of the user. Following we review visual sentiment analysis, a more recent line of work on sentiment analysis that gives a new approach to the problem of detecting sentiment polarity of the users. In the same way, we later review works on another approach to the sentiment analysis task, which is sentiment analysis on audio data. We follow with a review of works on multi-modal fusion sentiment analysis, which refers to works that address sentiment analysis with a combination of techniques, using analysis of text and audio data and other sources of data. Also in these works, different levels of fusion are investigated. Next, we continue reviewing sentiment analysis works that use CBR technologies for their approaches. Finally, we finish with a review of works on stress analysis and works that use keystroke dynamics data to detect sentiment and stress. For covering the works most relevant to the aim of this review, we define an inclusion/exclusion criteria as follows:

- Is it relevant to the works being reviewed in a given section? (e.g., a work included in the sentiment analysis on audio data subsection has to focus on sentiment analysis techniques applied to audio).
- The work uses a technique or techniques to address the problem which are different than the ones used in other works reviewed previously in the same section? (e.g., dictionary-based methods are different from machine learning models for sentiment classification in sentiment analysis).

There can be works using the same technique as another reviewed work if the problem addressed by the paper is different (e.g., one work addresses emotion detection using big data, which is different from other work performing emotion detection on stored data or on single users).

- Does the work provide experiments and data that gives insight on the usefulness of the used technique to address the problem at its focus? (e.g., accuracy, precision, and recall of the proposed technique or method on the addressed problem, or data on the significance of an effect, for example data on the significance of the effect of emotion on keystroke latency).

For searching for works, different databases were used, which were Google Scholar and Web Of Science. Intensive search was performed on both of them for finding relevant works for the aim of this literature review.

Table 1. Performance of the detection approaches.

Reference	Metrics	Values
[10]	Accuracy, precision, recall and F1	0.6778–0.7961, 0.628–0.8905, 0.5931–0.9231 and 0.5549–0.8591
[11]	Precision and recall	0.56–0.79 and 0.67–0.80
[12]	Precision, recall and F1	Single-domain approach: 0.622–0.792, 0.414–0.661 and 0.497–0.702
		Cross-domain approach: 0.565–0.678, 0.273–0.435 and 0.36–0.518
[13]	Accuracy, precision, recall and F1	Multi-aspect sentence labeling: 0.477–0.83, 0.126–0.969, 0.179–1 and 0.148–0.887
	L1 error, ρ aspect, ρ preview and MAP @10	Multi-aspect rating prediction with indirect supervision: 0.238–0.645, –0.149–0.715, 0.454–0.846 and 0.129–0.429
	L1 error	Supervised multi-aspect rating prediction: 0.554–1.071
[14]	Precision and recall	Explicit feature extraction task: 93–95% and 73–80%
		Finding word semantic orientation: 72–88% and 55–83%
		Extracting opinion phrases: 79% and 76%
		Extracting opinion phrase polarity: 86% and 89%
[15]	Precision and recall	0.75–0.955 and 0.2–0.286
[16]	Precision, recall and F1	0.761–0.918, 0.37–0.82, and 0.498–0.837
[17]	Accuracy	Using visual features: 0.49–0.83
		Using visual and text features: 0.48–0.88
[18]	Precision, recall, accuracy and F1	0.691–0.797, 0.729–0.905, 0.667–0.783 and 0.722–0.846
[19]	Accuracy	0.68–0.82
[20]	Accuracy	0.452–0.851
[21]	Unweighted Average Recall (UAR)	0.579–0.627
[22]	Accuracy (over automatic speech recognition output and human transcripts)	Using text and acoustic features: 0.825 and 0.81
		Using only text: 0.75 and 0.844
[23]	Mean linear error and correlation between estimates and the reference for valence, activation, and dominance	Acoustic emotion estimation: 0.13, 0.16, 0.14 and 0.53, 0.82, 0.78
		Visual emotion estimation (eyes region): 0.18, 0.19, 0.13 and 0.57, 0.58, 0.57
		Visual emotion estimation (lips region): 0.18, 0.19, 0.14 and 0.58, 0.62, 0.53
		Decision-level fusion acoustic and visual emotion estimation: 0.14, 0.12, 0.09 and 0.7, 0.84, 0.8

Table 1. *Cont.*

Reference	Metrics	Values
[24]	Accuracy for arousal and valence	Unimodal (Electroencephalogram): 0.65–0.7563 and 0.73–0.74
		Unimodal (Peripheral physiological signals): 0.6689–0.6905 and 0.6689–0.6905
		Feature-level fusion: 0.7844–0.8563 and 0.8279–0.8398
		Decision-level fusion: 0.66–0.73 and 0.5–0.64
[25]	Root Mean Squared Error (RMSE), Correlation Coefficient (COR) and Sign Agreement Metric (SAGR) for valence and arousal	Single-cue prediction: 0.17–0.22, 0.444–0.712, 0.648–0.841 and 0.24–0.29, 0.411–0.586, 0.681–0.764
		Support vector regression: 0.21–0.25, 0.146–0.551, 0.538–0.740 and 0.26–0.27, 0.388–0.419, 0.667–0.716
		Feature-level fusion: 0.19–0.21, 0.583–0.681, 0.733–0.856 and 0.24–0.28, 0.461–0.589, 0.685–0.763
		Model-level fusion: 0.16–0.19, 0.653–0.782, 0.830–0.892 and 0.22–0.26, 0.479–0.639, 0.637–0.8
[26]	Accuracy and speed up factor (big data tools)	0.831 and 75.55
	Precision and recall (sentiment analysis)	0.751 and 0.758
[27]	Reduction of product attribute redundancy	0.41
[28]	Accuracy	0.62–0.7258
[29]	Accuracy, precision, recall and F1	0.85–0.91 0.85–0.899, 0.847–0.92 and 0.85–0.91
[7]	Exact matches with strength label, 1 level away of strength matches (+ 1 of the label), correlation and MAD	Stress strength detection: 0.31–0.579, 0.791–0.939, 0.329–0.505 and 0.502–0.893
		Relaxation strength detection: 0.482–0.717, 0.916–0.963, 0.332–0.466 and 0.338–0.515
[30]	Accuracy	Detecting cognitive stress: 0.521–0.615 (raw data), 0.641–0.75 (normalized data)
		Detecting physical stress: 0.541–0.625 (raw data), 0.542–0.625 (normalized data)
	False Acceptance Rates (FARs) and False Rejection Rates (FRRs)	Detecting cognitive stress: 0.375–0.75 and 0.208–0.479 (raw data), 0.187–0.333 and 0.229–0.479 (normalized data)
		Detecting physical stress: 0.437–0.771 and 0.125–0.375 (raw data), 0.25–0.437 and 0.333–0.5 (normalized data)

Table 2. Techniques of detection approaches.

Reference	Technique
[10]	Sentence-level sentiment analysis based on lexicon and syntactic structures. Document-level sentiment analysis using a weighted sum of the polarities of sentences within the document.
[11]	Part of Speech (POS) tagging. Frequent feature generation using association rule mining. Feature pruning as pruning meaningless and redundant features. Opinion words extraction, extracting words that are adjacent to frequent features and are adjectives that modify the feature. Infrequent feature identification using opinion words to find the nearest noun/noun phrase of the opinion word.
[12]	Conditional Random Fields (CRFs) for extracting aspects. Features: token string text, POS tag of the token text, label of existence of short dependency path (between the token and opinion expressions), word distance label (token appear in the closest word distance to an opinion expression or not), opinion sentence label (does the token appear in a sentence with opinion expressions?).
[13]	Multi-aspect sentence labeling: Latent Dirichlet Allocation (LDA), Multi-grain LDA (MG-LDA), Segmented Topic Model (STM) and Local LDA models weakly supervised using seed words, supervised Support Vector Machine (SVM) and a majority baseline that assigns the most common aspect label. For multi-aspect rating prediction with indirect supervision: LDA, MG-LDA, STM, and local LDA weakly supervised with seed words are used to label sentences with aspects, and a Support Vector Regression (SVR) model is trained with the combined vectors of each entity and their overall ratings. Supervised multi-aspect rating prediction: Perceptron Ranking (PRank) and linear SVR are used, trained with and without features derived from LDA, MG-LDA, STM, and local LDA, that do not make use of seed words, and trained with unigram baseline features.
[14]	Feature extraction: the algorithm extracts frequent noun phrases from parsed reviews. It also examines opinion phrases associated with explicit features in order to extract implicit properties. Finding opinion phrases: if there is an explicit feature in a sentence, the algorithm applies extraction rules to find opinion phrases. Each head word, together with its modifiers, is returned as a potential opinion phrase. Opinion phrases polarity detection: relaxation labeling.
[15]	POS tagging and shallow parser. Syntactic parsing and sentiment lexicon used for sentiment analysis and relating sentiment expressions to subjects.
[16]	CRFs. Joint extraction of opinions and object features.
[17]	Linear SVMs and Logistic Regression (LR) for learning Ajective Noun pairs (ANPs) detectors for visual sentiment analysis. Visual Sentiment Ontology (VSO) based on ANPs. SentiBank, visual concept detection library that can detect 1200 ANPs in images using ANP detectors, based on VSO.
[18]	Progressive Convolutional Neural Network (PCNN) for image sentiment classification trained using weakly labeled data. Data from the output of the model is used to fine-tune it. Previously trained CNNs are fine-tuned with a small set of manually labeled images for addressing domain transfer.
[19]	Automatic Speech Recognition (ASR) for obtaining transcriptions of videos. POS tagging for extracting text-based sentiment features. A Maximum Entropy (ME) model with feature tuning for sentiment classification.
[20]	Algorithm for emotion classification using region switching between Vowel-like Regions (VLR) and non-VLRs from audio data.
[21]	Sparse autoencoder-based feature transfer learning method, using a single-layer autoencoder to find a common structure in small target data and then applying such structure to reconstruct source data.

Table 2. Cont.

Reference	Technique
[22]	ASR engine: The AT&T Watson speech recognizer was used to convert spoken review summaries to text. Linear interpolation of the three Katz's backoff language models. In the experiments, Ada-boost with acoustic features combined with a text-based prediction feature was used and compared to LR, SVM, and C4.5 decision tree.
[23]	Acoustic emotion estimation: SVR. Visual emotion estimation: SVR. Decision-level fusion: weighted linear combination of the acoustic and visual estimations for a given sentence using different weights for the estimation of valence, activation and dominance.
[24]	Hidden Markov Models (HMMs) using multi-modal feature sets. Electroencephalogram (EEG) from the central nervous system and three kinds of Peripheral physiological signals (PERI) from the peripheral nervous system (Respiration or RSP, Electromyogram or EMG, and Skin Temperature or TMP) are used. Fusion at feature level is performed, and at decision level employing six different strategies, classification is performed using feature fusion, decision fusion, and non-fusion models.
[25]	SVR and bidirectional Long Short-Term Memory Neural Networks (BLSTM-NN) both used for single-cue prediction of valence and arousal. BLSTM-NNs are used for feature-level and model-level fusion, as well as for output-associative fusion of different cues (facial Expressions, shoulder cues, and audio cues). Model-level fusion performs fusion of the output of BLSTM-NNs predicting valence or arousal, using different cues (one cue in each BLSTM-NN), and uses it as the input for another BLSTM-NN. Output-associative fusion fuses the output of BLSTM-NNs predicting valence and BLSTM-NNs predicting arousal using different cues (again one cue in each BLSTM-NN).
[26]	Audio-visual big data emotion recognition system, using Multi-directional Regression (MDR) features for speech and Weber Local Descriptor (WLD) features for face images. SVM classifiers are used for each modality and decision-level fusion using Bayesian sum rule.
[27]	Fuzzy SVMs for sentiment analysis on customer reviews of products. Case-based Reasoning (CBR) for generating ordinary and extraordinary use cases, from the sentiment labeled product attributes obtained from the SVMs.
[28]	CBR to compare text documents and different sentiment lexicons for sentiment classification associated to the cases.
[29]	Domain ontology and POS tagging for sentiment analysis, using CBR for reusing past cases of sentiment detection on text.
[7]	Algorithmic approach using a lexicon of stress and relaxation terms to detect stress and relaxation magnitude on text. The value predicted on a sentence is based on the score of the highest stress or relaxation term found within that sentence. Sentiment on a text with more than one sentence is computed as the highest value from any constituent sentence. Corrections such as negation of stress terms or spelling correction are applied.
[30]	Decision Tree (DT), SVM, k-Nearest Neighbor (kNN), AdaBoost, using DecisionStump as a base classifier, and ANN are used. DT was used to select features as input for the other methods.

Table 3. Datasets used and partitions for training and testing of the detection approaches.

Reference	Dataset or Datasets	Partitions
[10]	Euthanasia dataset: 851 Chinese articles on “euthanasia”, manually labeled into 502 positive and 349 negative articles. AmazonCN dataset: 458,522 reviews from six categories (book, music, movie, electrical appliance, digital product, and camera), labeled according to Amazon user’s five-star rating into 310,390 positive and 29,540 negative reviews.	Euthanasia dataset: Standard 10-fold cross-validation was performed. AmazonCN dataset: Up to 200 positive and 200 negative randomly selected reviews of each product category as the training dataset, and up to 500 positive and 500 negative randomly selected reviews of each product category as the test dataset.
[11]	Customer reviews from Amazon.com and C net.com about five products (2 digital cameras, 1 DVD player, 1 mp3 player, and 1 cellular phone). 100 reviews for each product. A person extracted features manually for evaluation, resulting in 79, 96, 67, 57, and 49 manual features for Digital camera1, Digital camera2, Cellular phone, Mp3 player, and DVD player, respectively.	The data was used in the proposed system to perform feature extraction and compare it to the manually extracted features.
[12]	Four datasets annotated with individual opinion target instances on a sentence level. Movies: reviews for 20 movies from the Internet Movie Database (1829 documents containing 24,555 sentences). Annotated with opinion target—opinion expression pairs. Web-services: reviews for two web-services collected from epinions.com (234 documents containing 6091 sentences). Cars: Reviews of cars (336 documents containing 10,969 sentences). Cameras: blog postings regarding digital cameras (234 documents containing 6091 sentences). In order for datasets movies, web-services, cars and cameras: sentences with targets: 21.4%, 22.4%, 51.1% and 54.0%; sentences with opinions: 21.4%, 22.4%, 53.5% and 56.1%.	As development data for the CRF model, 29 documents from the movies dataset, 23 documents from the web-services dataset, and 15 documents from the cars and cameras datasets were used. 10-fold cross-validation in single-domain (single dataset) experiments. In the cross-domain experiments, the system is trained on the complete set of data from one or various datasets and tested on all the data of a dataset not used in training.

Table 3. Cont.

Reference	Dataset or Datasets	Partitions
[13]	<p>OpenTable: 73,495 reviews (29,596 after excluding excessively long and short reviews) and their associated overall, food, service, and ambiance aspect ratings for all restaurants in the New York/Tri-State area appearing on OpenTable.com. Not labeled;</p> <p>CitySearch: 652 restaurant reviews from CitySearch.com. Each sentence manually labeled with one of six aspects: food, service, ambiance, price, anecdotes, or miscellaneous;</p> <p>TripAdvisor: 66,512 hotel reviews. Each review labeled with overall rating and ratings for 7 aspects: value, room, location, cleanliness, check-in/front desk, service, and business services.</p>	<p>Multi-aspect sentence labeling: For evaluation, 1490 singly-labeled sentences from the annotated portion of the CitySearch corpus were used. Inference is performed on all 652 documents of CitySearch; multi-aspect rating prediction with indirect supervision: OpenTable and TripAdvisor datasets sentences are labeled with aspects using weakly supervised topic models. All reviews for each entity (hotel or restaurant) are combined into a single review, aspect ratings are obtained by averaging the overall/aspect ratings for each combined review. 5-fold cross-validation is performed; supervised multi-aspect rating prediction: 5-fold cross-validation on subsets of the OpenTable and TripAdvisor data.</p>
[14]	<p>Two sets of 1307 reviews downloaded from tripadvisor.com for Hotels and amazon.com for Scanners.</p> <p>Two annotators labeled a set of 450 feature extractions from the algorithm as correct or incorrect. The annotators extracted explicit features from 800 review sentences (400 for each domain);</p> <p>Word semantic dataset: 13,841 sentences and 538 previously extracted features;</p> <p>Opinion phrase dataset: 550 sentences containing previously extracted features. The sentences were annotated with opinion phrases corresponding to the known features and with opinion polarity.</p>	<p>Explicit feature extraction: the algorithm was evaluated on the two sets from TripAdvisor and amazon.</p> <p>Finding word semantic orientation: the algorithm was evaluated on the Word semantic dataset. Extracting pinion phrases and opinion phrase polarity detection: the algorithm was evaluated on the Opinion phrase dataset.</p>
[15]	<p>Benchmark Corpus: 175 samples of subject terms within context text. Contains 118 favorable sentiment samples and 58 unfavorable samples;</p> <p>Open Test Corpus: 2000 samples related to camera reviews. Half the samples are labeled favorable or unfavorable and the other half neutral; 6415 web pages with 16,862 subject references, 1618 news articles with 5600 subject references, 1198 pharmaceutical web pages with 3804 subject references.</p>	<p>The system was directly used with data from the datasets.</p>

Table 3. Cont.

Reference	Dataset or Datasets	Partitions
[16]	<p>Movies dataset: 500 reviews about 5 movies. Contains 2207 sentences;</p> <p>Products dataset: 601 reviews about 4 products. Contains 2533 sentences.</p> <p>Both datasets are labeled manually by humans. Labels for object features, positive opinions, negative opinions, and the object feature-opinion pairs for all sentences are given.</p>	<p>Each dataset is split into 5 parts, and four are used for training while one for testing.</p>
[17]	<p>Flickr dataset: 150,034 images and videos with 3,138,795 tags; YouTube dataset: 166,342 images and videos with 3,079,526 tags; Amazon Mechanical Turk (AMT) experiment: randomly sampled images of 200 Adjective Noun Pair (ANP) concepts from the Flickr images, manually labeled by AMT crowdsourcer; Twitter Images dataset: Tweets containing images crawled using popular hashtags. Three labeling runs using AMT, namely image-based, text-based, and joint text-image based are performed. The dataset includes 470 positive tweets and 133 negative tweets over 21 hashtags; ArtPhotos dataset: ArtPhotos retrieved from DeviantArt.com. Contains 807 images from 8 emotion categories.</p>	<p>Training dataset with Flickr ANP labeled images: 80% of pseudo positive images of each ANP and twice as many negative images. Test datasets (full and reduced test sets): both use 20% of pseudo positive samples of a given ANP as positive test samples. The full test set includes 20% pseudo positive samples from each of the other ANPs (except those with the same adjective or noun) as negative samples. The reduced test set contains twice as many negative samples for each ANP as the positive samples. 5 versions of the reduced test set are created varying the negative samples.</p>
[18]	<p>Half million Flickr images weakly labeled with one ANP; Image tweets dataset: Tweets that contain images. The total is 1269 images. AMT is used to generate sentiment labels. Three sub-datasets are created: 581 positive and 301 negative images where 5 labelers agree, 689 positive and 427 negative images where at least 4 labelers agree, and 769 positive and 500 negative images where at least 3 labelers agree.</p>	<p>Randomly chosen 90% of the images from Flickr are the training dataset. The remaining 10% images are used as the testing dataset in the experiments with CNN and PCNN without domain transfer. 5-fold cross-validation is performed with Twitter images, using the training images to fine-tune a pre-trained model on Flickr images and the testing images to validate this model.</p>
[19]	<p>Amazon product reviews dataset: contains review comments about a large range of products including books, movies, electronic goods, apparel, and so forth; Pros and Cons and Comparative Sentence Set databases containing a list of positive and negative sentiment words/phrases; selected 28 youtube videos rated manually (16 negative and 12 positive sentiment) containing expressive speakers sharing their opinion on a wide variety of topics including movies, products, and social issues.</p>	<p>From the combination of the Amazon product reviews, Pros and Cons and Comparative Sentence Set datasets, extracted 800,000 reviews for training, and 250,000 reviews for evaluation were used.</p>

Table 3. Cont.

Reference	Dataset or Datasets	Partitions
[20]	EMODB database: Ten professional speakers for ten German sentences, 535 speech files, seven emotions (anger, anxiety, boredom, disgust, happiness, neutral and sadness), recorded at 48 kHz; IEMOCAP database: Audio-visual data in English, only audio track considered for this work, five male speakers and five female speakers, six emotions of the IEMOCAP database are considered (anger, excited, frustration, happiness, neutral and sadness), recorded at 16 kHz; FAU AIBO database: spontaneous emotional speech, contains recordings of 51 German children (21 male and 30 female) at the age of 10–13 years interacting with a pet robot. Contains 9959 training chunks and 8257 testing chunks with length approximately 1.7 s. Chunks are categorized into five different emotions (anger, emphatic, neutral, positive, and rest).	Leave-one-speaker-out cross-validation protocol was used for EMODB, IEMOCAP, and FAU AIBO databases. Additionally, with FAU AIBO database, a predefined partition of one children's data is used for validation purposes, and the remaining children's data is used for training purposes.

Table 4. Datasets used and partitions for training and testing of the detection approaches (continuation).

Reference	Dataset or Datasets	Partitions
[21]	FAU AEC database: based on the FAU AIBO emotion corpus, which contains recordings of children interacting with a pet robot in German speech. In the training set there are 6601 instances of positive and 3358 negative valence, and in the test set 5792 positive 2465 negative valence; TUM Audio-Visual Interest Corpus (TUM AVIC), Berlin Emotional Speech Database (EMO-DB), eNTERFACE, Speech Under Simulated and Actual Stress (SUSAS), and the "Vera am Mittag" (VAM) database. The age, language, kind of speech, emotion type, number of positive and negative utterances and sampling rate are: children, German, variable, natural, 5823, 12393 and 16 kHz for FAU AIBO; adults, English, variable, natural, 553, 2449 and 44 kHz for TUM AVIC; adults, German, fixed, acted, 352, 142 and 16 kHz for EMO-DB; adults, English, fixed, induced, 855, 422 and 16 kHz for eNTERFACE; adults, English, fixed, natural, 1616, 1977 and 8 kHz for SUSAS; adults, German, variable, natural, 876, 71 and 16 kHz for VAM.	FAU AEC is chosen as target set and the rest are used as source sets.

Table 4. Cont.

Reference	Dataset or Datasets	Partitions
[22]	<p>Corpus of 3,268 textual review summaries produced by 384 annotators, resulting in 1055 rated as negative, 1600 as positive, and 613 as mixed;</p> <p>CitySearch dataset: 87000 reviews describing more than 6000 restaurant businesses from the citysearch.com website; AMT text dataset: short text reviews summarized by Amazon turkers; GoodRec dataset: a set of short restaurant and bar recommendations mined from the goodrec.com website; short reviews of restaurants dataset: 84 participants made short reviews of restaurants on phone, answering questions, rating them and making a short free review. Resulted in 52 positive and 32 negative reviews.</p>	<p>The text-based classification is done training with the complete set of textual review summaries. The speech recognition models were trained using the CitySearch, AMT, and GoodRec datasets. Sentiment analysis from acoustic features models were trained on the short reviews of restaurants dataset, performing 10-fold cross-validation.</p>
[23]	<p>VAM corpus: consists of audio-visual spontaneous speech. signals were sampled at 16 kHz and 16 bit resolution. Facial image sequences were taken at a rate of 25 fps. Labeled with emotion by human listeners. The signals were sampled at 16 kHz</p>	<p>The VAM corpus was used for all the experiments. 245 utterances of 20 speakers for acoustic emotion estimation were used, performing 10-fold cross-validation. For the visual emotion estimation, 1600 images were used, again performing 10-fold cross-validation. For audio-visual fusion emotion estimation, 234 sentences and 1600 images were used.</p>
[24]	<p>Database for Emotion Analysis using Physiological Signals (DEAP): physiological signals Electroencephalogram (EEG) and Peripheral physiological signals (PERI) are used. EEG was recorded from 32 active electrodes (32 channels). PERI (8 channels) from Peripheral Nervous System (PNS) include Galvanic Skin Response (GSR), Skin Temperature (TMP), Blood Volume Pulse (BVP), Respiration (RSP), Electromyogram (EMG) collected from zygomaticus major and trapezius muscles, and horizontal and vertical Electrooculograms (hEOG and vEOG). The signals were recorded while playing 41 different music clips, and self-report of valence and arousal was done by the participants. Ten participants did 400 self-reports on valence and 400 on arousal.</p>	<p>For the feature-level fusion method, the DEAP database was used in the training and the most significant feature sets were selected for testing. Nested five-fold cross-validation was used in the testing phase. For decision-level, features were extracted in the training as well by performing nested five-fold cross-validation. DEAP database was used for all the experiments.</p>

Table 4. Cont.

Reference	Dataset or Datasets	Partitions
[25]	Sensitive Artificial Listener Database: spontaneous audio-visual interaction between a human and an operator with different personalities (happy, gloomy, angry, and pragmatic). The Sampling rate for video is 25 fps, and for audio 16 kHz. A set of coders annotated the recordings in the continuous valence-arousal 2D space confined to $[-1,1]$, although not all the data in the database has been labeled.	For validation, a subset of the SAL-DB that consists of 134 audiovisual segments (a total of 30,042 video frames) obtained by automatic segmentation was used. In this work subject-dependent leave-one-out-validation evaluation was used for the experiments.
[26]	eNTERFACE database: 42 non-professional subjects, 81% were male and 19% female, reacting to 5 sentences for each emotion between anger, disgust, fear, happiness, sadness, and surprise. The average length of the samples was 3 seconds. Berlin's emotional speech database and Kanade-Cohn emotional face database were used as single modality databases. Subjects were well-trained for acting according to the emotions. The extra emotion category is found in the Berlin database. Additionally, a massive amount of continuous video (voice or speech and facial video) generated from a video camera or smart mobile-phone based cameras, while the person is using social network service or smart health monitoring services was compiled into five datasets of different sizes.	In the experiments without using big data tools, four-fold validation was performed on the eNTERFACE, Berlin, and Kanade-Cohn databases. In the experiment with big data tools, the five datasets of continuous video generated were used, and a block replication number of three and a block size of 64 MB were used as the default settings in Hadoop. The authors varied the settings in the experiments in order to examine how the performance varies with respect to various: cluster sizes, block sizes, and block replication numbers.
[27]	Kindle Fire HD 7 reviews. Unstructured review data collected from 2 October 2012 to 20 November 2013. User-provided ratings.	A ten-fold cross-validation method is adopted for fine-tuning the parameters of the fuzzy Support Vector Machine (SVM) models and for sentiment prediction, using the Kindle Fire reviews data.
[28]	6 Text user review datasets: IMDB dataset of film reviews (2000 reviews); hotel reviews (2874 reviews); Amazon apparel products reviews (2072 reviews); Amazon music products reviews (2034 reviews); Amazon book products reviews (566 reviews); Amazon electronic products reviews (5902 reviews). All of the datasets have an equal number of positive and negative labeled reviews.	6 distinct case bases are created by training on datasets of all but one of the domains, and then each case base is used to classify documents on the hold out domain, which is the domain not used for populating the case base of the Case-Based Reasoning (CBR) module.
[29]	1999 reviews about digital cameras labeled by users of Amazon with sentiment polarity. 1000 positive reviews and 999 negative reviews; 1991 reviews about DVD movies labeled by users of Amazon with sentiment polarity. 996 positive reviews and 995 negative reviews.	Leave-one-out cross-validation on the two datasets (cameras and movies) was performed.

Table 4. Cont.

Reference	Dataset or Datasets	Partitions
[7]	Development corpus: this corpus is a collection of 3000 stress-related tweets, manually classified by the author for stress and relaxation. These tweets were identified by monitoring a set of stress and relaxation keywords over a week; six corpora of English short text messages extracted from Twitter.com (tweets), and coded by humans with stress and relaxation strengths. They were extracted from Twitter monitoring certain keywords in a 1 month period in July 2015. The corpora are: Common short words (608 tweets); Emotion terms (619 tweets); Insults (180 tweets); Opinions (476 tweets); Stress terms (655 tweets); Transport (528 tweets).	For assigning term strengths, identify missing terms, and to refine the sentiment term scores the development corpus was used. The performance of the supervised version of TensiStrength was evaluated using 10-fold cross-validation 30 times on the data of the six English short text datasets, with the average scores across the 30 iterations recorded.
[30]	24 participants, with ages ranged from 18 to 56, being 14 female, 10 male and 22 right-handed. They were asked to type with a keyboard after cognitive and physical stress tasks. Sessions spread over at least 3 days, ranging from 3 to 22 per participant, with a median of 9 days. The data collected was information about event (key up or down), time stamp (10 ms resolution), and key code. After each task, participants self-reported their stress level.	Baseline condition, control condition, and 2 experimental conditions were used. Baseline: 10 samples under no stress. Control: two samples under no stress. Experimental: completed either a cognitively or physically challenging task prior to providing a typing sample. The performance of each machine learning model was evaluated with three-fold cross-validation.

3.1. Sentiment Analysis on Text

Sentiment analysis can be applied to different kinds of media. In this section, we will review state-of-art works in the line of applying sentiment analysis to texts and using distinct techniques. Sentiment analysis on texts has been assessed with four well-differentiated techniques in the literature, which are document-level sentiment analysis, sentence-level sentiment analysis, aspect-based sentiment analysis, and comparative sentiment analysis [31]. Starting from sentiment detected in an entire document, to sentiment in a sentence and finally in an aspect, which is a sequence of words representing an aspect in the text (e.g., the government, love, group of people). In other words, depending on the level of fine-grained analysis that we want to perform we will be choosing one or another. Comparative sentiment analysis does not apply to this concept, it is an exception to the techniques of sentiment analysis on text mentioned in this paragraph, where we use comparative words as the elements of the model with a sentiment polarity associated. The model is trained so we can learn which ones are the preferred entities using comparative sentences [31].

Document-level sentiment analysis has two important issues. Having to give an aggregate value of sentiment for an entire document, and that it may, and potentially will contain a varied set of polarities in different sections of it. For these reasons, researchers developed more fine-grained levels of sentiment analysis. Sentence-level is easier but still could potentially find more than one polarity in the same sentence, which may lead to a conflict when trying to generate an aggregated sentiment for the sentence. Thus, aspect-level sentiment analysis was created, which focuses on concrete aspects or entities and tries to give as output a sentiment polarity associated with them.

In Reference [10] sentence-level sentiment analysis based on a sentiment lexicon and sentence syntactic structures is performed on Chinese texts, and further used to calculate the aggregated sentiment of the document where the sentences are used to compute a weighted sum of the polarity of each sentence in the document, considering the importance of each sentence. This work, even when addresses sentence-level and document-level sentiment analysis using sentences as atomic units for semantic analysis, is unable to perform a more fine-grained analysis, which is achieved with aspect-based sentiment analysis.

For the case of aspect-based sentiment analysis, there are two main problems to address, which are aspect detection and sentiment classification. Aspect detection is the detection of aspects from the training data for the model and sentiment classification is the actual sentiment analysis on the aspects, to give a sentiment label to them. There are several approaches for solving both of those problems, and also hybrid approaches that try to assess them at the same time [32]. For the case of aspect detection, frequency-based methods use the terms found in the training corpus with higher frequency as aspects for the aspect set of the sentiment analysis model [11]. Generative models are used for detecting aspects as well, in Reference [12] Conditional Random Fields (CRFs) with a varied set of features are used. The frequency-based method generates aspects using the most frequent terms, while generative models use string tokens and their associated features for this matter. The frequency-based method presented in Reference [11] has slightly less precision but significantly more recall than the one achieved in Reference [12]. Sentiment classification is addressed with dictionary-based methods in Reference [11], where a dictionary is generated by propagating the sentiment of a set of seed words through WordNet synonym/antonym graph, counting only adjectives as sentiment words. Then the dictionary is used for calculating sentence-level sentiment analysis using majority voting of the adjectives detected in a sentence and found in the dictionary, with some refinements such as flipping the polarity of negated adjectives or multiple polarities finding when the number of positive and negative words found in a sentence is the same. Supervised and unsupervised machine learning methods are also used for sentiment classification. In Reference [13] a Support Vector Regression (SVR) model is used to find the sentiment score of aspects as a real number in the interval zero to five. In Reference [14] each aspect is used to find phrases that could contain sentiment, and then a label is assigned to them using a method from computer vision called relaxation labeling. Dictionary-based approaches use an existing set of labeled terms, while supervised and unsupervised machine learning options use different machine learning techniques to assign labels to new terms. The reported precision was higher for machine learning approaches in this case than the dictionary-based method, but the recall was similar, although it reached higher values in some cases for machine learning approaches. Hybrid approaches try to detect aspects and assign sentiment polarities to them at the same time. In Reference [15] a syntax-based method was used to extract other aspects from words associated with a sentiment, exploiting semantic relations. Generative models are also used, in Reference [16] CRF are used to relate sentiments to aspects, extracting information from the relations between words. Both of the hybrid methods reviewed shown a high precision compared to other reviewed methods, but only the CRF approach presented in Reference [16] reached high recall.

Applying the criteria for inclusion/exclusion of works relevant to this review, the work in Reference [10] was selected for applying sentiment analysis to at sentence-level and document-level, while others were selected for illustrating aspect-level sentiment analysis, using different techniques, and thus giving a global insight on the state-of-art in these lines of work. In Reference [12] a frequency-based method is presented for aspect detection, while in Reference [12] CRFs are used. Those two techniques are different as one uses frequent terms, while the other extracts features using a model and features related to text tokens. For the case of sentiment classification, different kinds of techniques are presented. In Reference [11] dictionary-based methods are used, while in Reference [13] supervised machine learning models are presented, and in Reference [14] an unsupervised method is used. Finally, two different works using hybrid techniques for aspect detection and sentiment

classification are reviewed. A syntax-based method is presented in Reference [15], while CRFs are used in Reference [16].

3.2. Visual Sentiment Analysis

In general, there are three main state-of-art approaches to visual sentiment analysis [33], which are mid-level sentiment ontology, deep sentiment prediction, and multi-modal sentiment prediction.

Mid-level sentiment ontology provides a set of mid-level features for visual sentiment analysis. Borth and Ji [17] proposed a mid-level feature named adjective-noun pairs (ANPs), which are related to sentiments and a visual sentiment ontology (VSO). These ANPs with sentiment are extracted from a set of annotated images, then they are used to train VSO detectors, and a set of detectors are used to construct SentiBank, which is finally used for sentiment prediction on images. Deep sentiment prediction uses deep learning models to predict sentiment on images. Yu et al. [18] proposed a Progressive Convolutional Neural Network (CNN) that used a selection of images from the training set, according to the output of the trained model on them to fine-tune the training, and also address domain transfer by using a set of manually labeled Twitter images to fine-tune a previously trained CNN model. Finally, multi-modal sentiment prediction refers to creating classifiers that use different kinds of data to predict a sentiment, such as text and images. A multi-modal correlation model based on Markov Random Fields (MRFs) was proposed in Reference [34]. Multi-modal features are extracted, with ANPs as image features, words as text features, and symbols as emoticons features. A MRF model is created and decomposed in 4 subgraphs (three single graphs and one correlation graph). The single graph denotes the contribution of each modality to the sentiment, and the correlation graph denotes the correlation contribution. According to the correlation of each modality, the final model is graphed and then learned. Following the inclusion/exclusion criteria, in this section the three different approaches used in the literature for visual sentiment analysis are presented in three different works. The three different strategies or techniques have clear differences, either use mid-level features that help predict an emotion from images, such as ANP that are text concepts related to the image (adjective-noun pairs), deep learning models, or a multi-modal model that analyzes both text and images to predict emotion. Deep models reported to have reached a better accuracy overall but slightly less maximum accuracy than mid-level sentiment ontology prediction in the reviewed works.

3.3. Sentiment Analysis on Audio Data

Sentiment analysis has been successfully performed using only audio data. In Reference [19] an Automatic Speech Recognition (ASR) system is used to convert youtube videos to transcribed text, and then a Part Of Speech (POS) tagger based feature extraction technique to identify useful sentiment features in the text, to later be classified into positive or negative sentiment polarity by a Maximum Entropy (ME) based sentiment classification model; a method is proposed in Reference [20] to perform speech emotion recognition. The proposed method performs emotion recognition by employing Vowel-Like Regions (VLRs) and Non-Vowel-Like Regions (non-VLRs), and by choosing the features of either VLRs or non-VLRs for each emotion; a sparse autoencoder based feature transfer learning method is proposed in Reference [21] that uses a single-layer autoencoder to find a common structure in small target data and then applies it to reconstruct source data for performing knowledge transfer from source data into target task. The authors used the reconstructed data to build a speech emotion recognition engine for a real-life task and performed experiments with six publicly available corpora. The experiments showed that the proposed algorithm enhances emotion classification accuracy of the speech emotion recognition engine significantly; sentiment analysis on short spoken reviews was performed in Reference [22], where authors collected manually a set of user reviews and extracted acoustic features from the spoken reviews using the openEAR/openSMILE toolkit. Various algorithms were used, which were logistic regression, AdaBoost, a C4.5 decision tree, and a Support Vector Machine (SVM) classifier with a radial basis function (RBF) kernel. The best performing algorithm was AdaBoost with speaker-dependent features after applying manual feature selection, reaching

an accuracy of 72.9%. The four works reviewed in this section give an overview of the different possibilities for emotion recognition on audio, having a classic ME model applied to POS features from transcripts, a work using different features and selecting concrete features for each emotion, a deep learning approach that leverages classic model detection, and a work using different algorithms. Therefore, we followed the inclusion/exclusion criteria in this section. The reported results shown a better accuracy reached by the work in Reference [20], using different features for each emotion, although the best overall results of accuracy are achieved by the work performed in Reference [22], where different algorithms are used.

3.4. Multi-Modal Fusion Sentiment Analysis

Regarding multi-modal fusion sentiment analysis, there are three main approaches in the literature, which are categorized by type of fusion (feature, decision, and hybrid) [9]. Feature level fusion fuses information creating features that have information of different data sources; Decision level fusion fuses different modalities in the semantic space; Hybrid level fusion combines both feature and decision level fusion.

A method for estimating spontaneously expressed emotions in audio-visual data was developed in Reference [23]. Support Vector Regression and decision level fusion achieved an average performance gain of 17.6% and 12.7% over the individual audio and visual emotion recognition methods respectively. In Reference [24], a feature-level fusion approach for recognizing emotion from video and physiological data was developed. The authors perform emotion recognition on the valence-arousal emotional space using Hidden Markov models (HMMs). The best recognition accuracies reported are of 85.63% for arousal and 83.98% for valence. A comparison with the proposed feature-level fusion approach against a decision-level fusion and non-fusion approaches was performed on the same DEAP database, showing that significant improvements in accuracy obtained by the feature-level fusion approach were observed. A hybrid output-associative fusion method for emotion prediction in the valence and arousal space was proposed in Reference [25]. The authors used facial expression, shoulder gesture, and audio cues, and used Bidirectional Long Short-Term Memory Neural Networks (BLSTM-NNs) and Support Vector Regression models for performing sentiment classification. The authors claim that their hybrid method outperforms results from predicting either valence or arousal alone (both for feature-level and model level fusion). A model that combines emotion aware big data and cloud technology with 5G is proposed in Reference [26]. The authors claim that the proposed approach achieves 83.10% emotion recognition accuracy. The different alternatives of multi-modal sentiment analysis, which are presented in Reference [9] are used in the works presented in this section, achieving different results. The decision-level fusion method presented in Reference [23] reported a higher overall correlation between the estimates and reference values of emotion than the hybrid output-associative fusion method presented in Reference [25], and the feature-level fusion method presented in Reference [24] not only achieved higher accuracy than the unimodal and decision-level fusion approaches compared in the same work, but also achieved higher accuracy than the big data approach presented in Reference [26]. These works were selected to give insight on the different alternatives of multi-modal sentiment analysis according to the inclusion/exclusion criteria, since each one uses one alternative, except the work in Reference [26] which follows a different aim, that is applying emotion recognition combined with big data technology.

3.5. CBR for Sentiment Analysis

The CBR approach has successfully been applied in the past to predict sentiment. In Reference [27] explicit customer needs are extracted from customer reviews of products, by means of performing sentiment analysis on them using fuzzy SVMs. Then, a CBR module constructs a case on its case base according to the ordinary use cases of products detected in the previous step. When extraordinary use cases are detected, the CBR module searches for the most similar ordinary case in the case base, and then elicits the extraordinary customer needs in the extraordinary use case based on the

ordinary case employing substitution, rule-based adaptation, and design engineer evaluation of the adapted extraordinary cases. In Reference [28] sentiment classification is addressed using a CBR-based approach. First, the case base is populated using labeled customer reviews and five different sentiment lexicons. If a document is correctly classified by at least one lexicon, a case is created containing document statistics and writing style of the review that generated it, and also the case solution which is the information about which sentiment lexicons generated a successful prediction on the review associated to the case. Prediction on new reviews is performed retrieving the k most similar cases (1, 2, or 3 in the reported experiments), and querying the lexicons of the retrieved solutions for sentiment information on the terms of the new review. Domain ontology with natural language processing techniques are combined in Reference [29] to perform sentiment analysis. Case-based reasoning is also used to learn from past sentiment polarizations. The authors claim that the accuracy obtained by the proposed model overcomes standard statistical approaches. Different ways of using a CBR-based approach for sentiment analysis are shown in the works presented in this section. While in Reference [27] CBR is used for generating extraordinary cases (cases of use of products not frequently found), based on ordinary or frequent cases, in Reference [28] the CBR approach is combined with sentiment lexicons to form cases of document statistics and writing style associated to lexicons that correctly classified the document that generated the case. Finally in Reference [29] CBR is used to learn from past sentiment polarizations while using domain ontology combined with NLP techniques for performing sentiment analysis. In the results can be observed that the later work outperforms the CBR approach for generating extraordinary use cases in precision and recall, and also outperforms the lexicon CBR approach in terms of accuracy. The inclusion/exclusion criteria was applied in this section selecting works that applied a CBR-based approach for sentiment analysis in different ways.

3.6. Stress Analysis and Keystroke Dynamics

Stress analysis has been addressed in the literature using different sources of information, like in the case of sentiment analysis. In Reference [7] text is used as the source of information for an algorithm that employs a lexicon of stress-related terms for calculating the stress score of a sentence, based on the score of the highest stress term found, with some rules that modify the base approach (e.g., spelling correction or negating stress words). Keystroke dynamics refers to the way an user types at a keyboard and includes different features such as timing features of key press and release and accuracy rate while typing. As an example of keystroke dynamics applied to stress analysis, in Reference [35] authors show through a Wilcoxon Signed Rank test that several keystroke dynamics features reject the null hypothesis that stress and non-stress data does not have a significant difference at least at 90 % confidence level. Another example is Reference [30], where the authors used keystroke dynamics and linguistic features for the analysis of free text, and demonstrated that such techniques can be effectively used to detect cognitive and physical stress from free-text data. The authors claim that the accuracy of detection of cognitive stress was consistent with those obtained using affective computing methods, and that the accuracy for the detection of physical stress, while not being as high as the ones obtained for cognitive stress, still encourages further research. There are works that perform sentiment analysis using a model trained with keystroke dynamics data. In Reference [36], the authors used IADS [37] sounds for inducing sentiments to a series of users and record their keystroke dynamics after hearing them. They showed that the effect of arousal on keystroke duration and keystroke latency was significant but the one on the accuracy rate of keyboard typing was not. The works performed in Reference [35] and Reference [36] demonstrate that there is an effect of stress and sentiment, respectively, on keystroke dynamics data, showing that models can be built to use this kind of data for performing sentiment and stress analysis. Keystroke dynamics data used in Reference [30] to detect stress, and text data used in Reference [7] for this same purpose can be compared in Table 1, showing that the text-based method outperforms the keystroke method when using stress strength detection while allowing the matches to be ± 1 stress level away of the label, while being outperformed when considering exact stress level matches.

Related to the inclusion/exclusion criteria, References [35] and [36] are selected for the review because they perform an analysis on the significance of the effect of stress and sentiment on keystroke data, respectively. For reviewing a method for detecting stress, References [7] and [35] were selected, since a method that addressed stress analysis on text data is presented in the former and one addressing stress analysis on keystroke data is shown in the later.

4. MAS-Based Prevention and Recommendation Systems Review

In the literature, we can find several works addressing user guiding or recommendation using the MAS technologies for it, and in this section, we review works in this line. The problem addressed, techniques, and contributions of each work are summarized in Table 5. Moreover, the automatic detection of sentiment polarities and stress levels by the system could be used to achieve a more satisfactory and safe user experience, by preventing potential risks that could arise from the interaction (e.g., triggering contact risks by publishing information that you do not really want to post because of cognitive distortions, attracting sexual predators). In this section previous works in the lines of applying user state detection to risk prevention will be reviewed.

For the works reviewed in this section, the last rule of the inclusion/exclusion criteria presented in Section 3 does not apply, since in this section works are presented which apply risk prevention, recommendation and user guiding approaches using MAS-based technologies, and thus examine different applications of the existing technologies for addressing these problems. Additionally, some works in this section cannot be directly compared (they address different topics such as user protection against cyber-bullying or online grooming and group recommendation, which are completely different problems). Therefore, only the techniques used are analyzed in detail, and not a performance comparison like the one performed with detection approaches in Section 3.

Table 5. Problem addressed, techniques and contributions of the prevention approaches.

Reference	Problem to Address	Technique	Contributions of the Proposal
[38]	Enhancing communication between users in Social Network Sites (SNSs).	Multi-agent System (MAS) and agents working as mediators.	Enhanced user engagement and collaboration in SNSs.
[39]	Collaborative filtering recommendation.	User ontology created by monitoring user behavior, and calculation of inter-ontology similarities. MAS that integrates the previous tasks, representing users as agents with an ontology representing their behavior.	Automatically creating a model of users by creating ontologies monitoring users, and computing similarities between users using such ontologies for recommendation.
[40]	Trust and reputation in MAS.	MAS implementing agents that perform certified recommendations. The certifications are achieved by using signed transactions or witnessed transactions by other agents as certificate.	Certified recommendations and possibility of the MAS to determine how much agents can be trusted as experts.
[41]	Business-to-customer (B2C) e-commerce activities.	XML-based MAS architecture with users personalized profiles. Such profiles are built and updated by weighting activities performed in B2C processes.	Implementation of business-to-customer e-commerce through using user profiles built with information of user actions in previous transactions.
[42]	Privacy-preserving recommendation systems.	Privacy-preserving protocol for information filtering processes that makes use of a MAS architecture and suitable filtering techniques (feature-based approaches and knowledge-based filtering).	The proposed approach provides information filtering while preserving privacy. An application of the proposal supporting users in planning entertainment-related activities is presented.

Table 5. Cont.

Reference	Problem to Address	Technique	Contributions of the Proposal
[43]	Content-based recommendation system, aiming to solve the new user and overspecialization problems.	MAS architecture as a recommendation system. Semantic enhancement of user preference through domain ontology and semantic association discovery in user profile database.	Addresses two existing problems in an existing technique, and experimental results show an improvement in positive feedback rate.
[44]	Group recommendation.	MAS approach based on negotiation techniques. A multilateral monotonic concession protocol is used to combine individual recommendations into a group recommendation.	Implementation of group recommendation using a MAS architecture and a multilateral protocol. Testing the proposed approach in the movies domain, users were found more evenly satisfied in the groups than with ranking aggregation.
[45]	Detecting the social emotion of a group of entities and influencing them.	Social-emotional model computed using an ANN, based on the pleasure, arousal, and dominance (PAD) three-dimensional emotional space. Application of the model in a group of Human-Immersed agents.	Artificial Neural Network (ANN) that computes the social emotion of a group of agents. Experiments show that using the proposed model to predict the emotion of a group of agents and computing the distance to a target emotion 'happiness' for selecting the action for the system to take achieves the distance to the target emotion to diminish after few iterations.
[46]	Cyber-bullying and online grooming prevention in SNSs through the use of different techniques.	Sentiment analysis on text by using different text mining modules, adult image detection using Skin Tone Pixels detection and message classification using Natural Language Processing (NLP) algorithms, through keyword search in the text.	Combination of different data analysis techniques including text and image analysis for prevention of user negative behaviors such as bullying and grooming.
[47]	Prevention of negative outcomes in SNSs, negative sentiment, and high stress levels through decision-level fusion of sentiment and stress analysis on text.	Sentiment, stress, and combined analysis of sentiment and stress using decision-level fusion on text using ANNs. MAS architecture with agents integrating different unimodal analyses, and an agent performing decision-level fusion and feedback generation to users in SNSs.	Combination of different data analysis techniques and a fusion technique with a MAS architecture for prevention of negative outcomes in SNSs. Experiments with data from Twitter that show significant differences between the analyzers predicting negative outcomes, and with a real-life SNS.
[48]	Prevention of negative outcomes in SNSs, negative sentiment, and high stress levels through decision-level fusion of sentiment and stress analysis on text and keystroke dynamics data.	Extension of a MAS architecture that employs ANNs for sentiment and stress analysis on text with new ANNs performing sentiment and stress analysis on keystroke dynamics data. Design of different decision-level fusion methods employing sentiment and stress analysis on text and keystroke dynamics data.	Addition of analyzers performing sentiment and stress analysis on keystroke dynamics data to a MAS with analysis on text data. Experiments performed with data from Twitter exploring different decision-level fusion methods, and proposal of a novel rule-based feedback generation agent in the MAS, in accordance with the results of the experiments.
[49]	Learning the sentiment associated with specific keywords from different data sources.	MAS architecture with agents implementing reinforcement learning algorithms, learning the sentiment associated with keywords, with each agent analyzing data from a different source.	Implements sentiment analysis on keywords by applying collective learning from different data sources and reinforcement learning algorithms in a MAS architecture.

Table 5. Cont.

Reference	Problem to Address	Technique	Contributions of the Proposal
[50]	Sentiment analysis on different SNSs using user opinion to construct a collective sentiment as the opinion of a product.	MAS architecture with agents implementing naïve Bayes classification for performing sentiment analysis on user opinions from different SNSs. A final sentiment is calculated using a common blackboard.	Collective sentiment or opinion about a product computed using sentiment analysis on different SNSs with a MAS architecture.
[51]	Design and implementation of an actor-based software library for building distributed data analysis applications.	Prototype library implemented using the ActoDeS software framework for the development of concurrent and distributed systems. The library implemented includes a MAS architecture and different implementations of five agent types, which are acquirer, preprocessor, engine, controller, and master.	Prototype of a library that provides a MAS architecture and agent implementations that wrap the different tasks of a data analysis application.
[52]	Framework for understanding and predicting the emergence of collective emotions.	Framework built as an agent-based model with agents modeled with individual emotion states and communication between agents.	Proposal of a framework that allows to understand and predict collective emotions, based on interactions between agents, who have individual emotional states.
[53]	Product opinion mining from SNSs data using big data analysis.	MAS architecture including a data extraction, analysis, management and manager agents. It is implemented using JADEX, an agent architecture for representing mental states in JADE agents. Agents make use of Hadoop MapReduce for data process and analysis, and HBase for data storage. Influence of the poster, knowledge about the topic, and sentiment analysis are computed on text messages in MapReduce.	Implementation of distributed data analysis using big data tools for opinion mining from SNSs.

In Reference [38] agents and a multi-agent system are suggested to work as communicator mediators between users of SNSs and social groups; in Reference [39] an ontology is constructed by monitoring user behavior, and later used in a task of collaborative filtering recommendation, by means of computing inter-ontology similarities; trust and reputation of agents in a MAS architecture are computed in Reference [40] on the basis of certified recommendations (e.g., based on signed or witnessed transactions), to make the system able to determine how much the agents can be trusted as experts; in Reference [41] an XML-based MAS architecture is proposed, which the authors called MAST. It supports business-to-customer (B2C) e-commerce activities, by means of user personalized profiles that are built and updated by weighting the activities performed in B2C processes. Mas architectures can be useful for several different purposes. The works already presented in this section showed that they can be successfully used for implementing interfaces between users and social content in SNSs, for guiding users; for monitoring the user behavior and later give advice; for addressing trust and reputation of agents in the system, so users can differentiate between which agents they should trust for a certain task; for addressing B2C e-commerce activities and other activities in which the user interacts with a business in a certain way, so these processes can be more guided. All of these applications of MAS technologies are useful for providing a more guided and satisfactory experience for users in a system in different ways. Additionally, since works addressing these different applications of MAS technologies for guiding users have been selected and reviewed, we followed our inclusion/exclusion criteria.

Moreover, MAS architectures have been used for creating recommendation systems. In Reference [42] an agent-based approach was developed for privacy-preserving recommendation

systems. The authors provide a privacy-preserving protocol for information filtering processes and make use of suitable filtering techniques, resulting in an approach that preserves privacy in information filtering architectures. An application of the proposed approach which supports users in planning entertainment-related activities is presented; a MAS architecture is proposed in Reference [43] as a content-based recommendation system, aiming to solve the new user and overspecialization problems existing in such systems. Semantic enhancement of the user preference through domain ontology and semantic association discovery in user profile database are performed, to address the new user problem. Experimental results suggested that there is an improvement in positive feedback rate. A multi-agent approach based on negotiation techniques for group recommendation is proposed in Reference [44]. Multilateral monotonic concession protocol is used in this approach to combine individual recommendations into a group recommendation. Applying the proposal to the movies domain, experimental results showed that applying this negotiation protocol, users were found more evenly satisfied in the groups than using traditional ranking aggregation approaches. To sum up, the MAS architecture proved to be useful for diverse aspects of recommendation systems, such as privacy-preserving recommendation, solving problems existing in traditional recommendation systems and to perform group recommendation. Since these topics of group recommendation using MAS technologies are different between them, and give insight of the usefulness of this technology for that purpose, References [42–44] were selected because every one addresses one of these topics, thus following our inclusion/exclusion criteria.

In Reference [45] a social-emotional model is computed using an Artificial Neural Network (ANN) for detecting the social emotion in a group of entities. The model is based on the pleasure, arousal, and dominance (PAD) three-dimensional emotional space. The authors show an example where they use the model to infer the emotion of a group of human-immersed agents when they hear a song, then predict the future emotion that the group would achieve after hearing new similar songs, and finally compute the distance of the predicted emotion to the target emotion 'happiness' for deciding which song to play, that would be the one that minimizes this distance. The reported distance of the emotion of agents detected against the target emotion on the experiments performed quickly diminishes after a few iterations of the system. Sentiment analysis was used in Reference [46] on the texts of users interacting inside a SNS, together with adult image detection and message classification to help the system ban users that incurred in either cyber-bullying or online grooming. Moreover, in Reference [47] a set of analyzers were built for performing sentiment analysis, stress analysis, and a decision-level fusion analysis, using sentiment and stress on the text messages of users interacting in a SNS. When the system detects negative sentiment or high levels of stress, a warning is sent to the user to prevent him from sending the message to the network as it is and avoid potential negative outcomes in the SNS. The authors performed experiments with data from Twitter to discover which analyzer of the proposed ones was able to detect a state of the user that propagated more to its replies in the SNS, finding significant differences between the analyzers. Experiments with a private SNS called Pesedia [54] were also performed to test the system in a real-life scenario. Additionally, in Reference [48] new agents were added to the system presented in Reference [47] to perform sentiment and stress analysis on keystroke dynamics data, proposing fusion analysis that employed analyzers working on text data and the ones working with keystroke data. Experiments were performed with data gathered from the private SNS Pesedia to find which of the proposed analyzers worked best at detecting states that propagated more in the SNS. Finally, a new version of the advisor agent was proposed, which generates feedback to users based on the input of the analyzers found best in the experiments. Thus, the MAS architecture is not only useful for guiding users or recommending them, but it also proved to be able to detect the state of users and use it for helping to prevent issues that could arise in a social environment, or make for a better social experience. This can be seen in Reference [45], where the goal of the system is to detect a group emotion and simulate agents interacting to achieve a better social experience; in Reference [46] where the system detects content in images and text to help itself detect dangerous users; and in References [47,48], where the system analyzes the sentiment and stress levels of users to

help prevent negative interactions in SNSs and risks. These four works were selected for assessing the usefulness of MAS technologies for detecting the state of the user to address the different topics mentioned. For the case of References [47,48], the former addresses sentiment and stress analysis on text data for guiding users, while the later uses both text and keystroke data combined for this purpose.

A MAS with agents that apply reinforcement learning algorithms to learn the sentiment pertaining to specific keywords was presented in Reference [49]. Agents collectively learn this sentiment since each agent processes its assigned subset of data. Experiments were conducted on abstracts from PubMed related to muscular atrophy, Alzheimer's disease, and diabetes, and results show that the system was able to learn the sentiment score related to specific keywords. A MAS architecture with a set of agents that work with opinion data from different SNSs is proposed in Reference [50]. The proposed system has an agent extracting opinion about a product from Twitter, another from Wikipedia, and another from Facebook. All agents compute sentiment using machine learning techniques and are able to communicate between them by using a blackboard. In this way, they can generate a more complete opinion about the product with sentiment computed on additional features by other agents. In Reference [51] ActoData (Actor Data Analysis), an actor-based software library for building distributed data analysis applications is presented, and a prototype is implemented. This library provides a multi-agent architecture and different implementations of five agent types, which are acquirer, preprocessor, engine, controller, and master. Each agent acts as a wrapper for components that perform the different tasks of a data analysis application. A framework built as an agent-based model is presented in Reference [52]. The framework is used to understand and predict the emergence of collective emotions based on interactions between agents, who have individual emotional states. For helping enterprises be aware of their customer's opinions about products or services, an agent-based social framework that extracts reviews from social media is presented in Reference [53]. Data analysis and storage are performed by using a framework based on Hadoop MapReduce and HBase, respectively, which allows the efficient manipulation of big amounts of data, which is the case for the task of opinion mining from SNS data. As can be seen, the MAS architecture can be useful to model the emotion of users or mine opinions in different ways that are related to computing this emotion or opinion in a distributed way. It can help compute collectively emotion using different sources of data, can help build distributed data analysis applications, predict the emergence of collective emotions based on interactions between agents with individual emotions, or create big data distributed applications to perform opinion mining with big volumes of data, such as SNS data. Therefore, the emotion of users can be computed in different ways and in a distributed architecture using MAS technologies and machine learning or other detection technologies, and then use this information to guide users, recommend, and prevent risks and potential issues in SNSs or other environments with emotional entities. For addressing the different topics presented related to using MAS technologies to detect emotion or mine opinion in a distributed way, a set of works was selected according to our inclusion/exclusion criteria: Reference [49] was selected for presenting a MAS with agents applying reinforcement learning algorithms to learn the sentiment of associated to keywords from different data; Reference [50] was selected for addressing distributed learning of a product opinion from different SNS data; Reference [51] presents an actor-based library for building distributed data analysis applications; Reference [52] addresses the problem of understanding and predicting the emergence of collective emotions in the basis of interactions between agents; Reference [53] addresses opinion mining from SNS data using big data techniques.

5. Discussion

As has been reviewed in the previous section, there is a substantial effort in the literature on detecting the sentiment polarity of people that create different kinds of content, using different sources of data. Between the data used to perform sentiment analysis, we can find extensive literature featuring approaches using text, audio, images, and in less quantity writing patterns. Some approaches only perform detection of sentiment on one data source, while others do it on multiple, which is

the case of multi-modal sentiment analysis. CBR technology has also been applied to performing sentiment analysis. Additionally, stress level detection has been performed on text data and writing patterns. Moreover, as has been commented in Section 1, the emotional state affects the decision making process, and there are risks that can arise from social interaction in online social environments such as SNSs. User state detection, as has been revised on this survey can be effectively used together with MAS technologies to guide and recommend users, and prevent potential risks or issues. Moreover, using different data modalities, which can be implemented in the MAS architecture, can improve the capacity of the system to detect a risk.

In the literature, there are works reviewing different previous approaches for sentiment analysis. On the one hand, there are surveys specialized in single data modality. Aspect-based sentiment analysis works on text data have been reviewed in Reference [32], while works on sentiment analysis applied to images have been reviewed in Reference [33]. On the other hand, in Reference [9] is a survey on works about sentiment analysis applied to different data modalities and also multi-modal sentiment analysis. Although there are works that review sentiment analysis using one or various data modalities, to the best of our knowledge, there is a lack of a review of works in user risk prevention when navigating online social environments, reviewing different strategies of sentiment and stress analysis, which is the case of the present survey. We also focus on works that use MAS-based techniques for recommending and guiding users, which is a technology that can be used together with SNSs as it fits the social network architecture by using agents to represent entities in the network, and guide or interact with users.

6. Conclusions and Future Lines of Work

Since we have discussed previous works in risk prevention, recommendation and user state detection, and highlighted the relations between them and potential uses for preventing users from suffering negative consequences from their interaction, and for helping improve future systems, following potential future works in three different lines will be discussed, which are:

- Using current technologies in user state detection for creating improved user guiding systems in online environments.
- Combining different technologies compatible with the architecture of SNSs with emotion detection techniques and testing their effectiveness in real-life scenarios as guiding or recommendation systems.
- Improving user state detection techniques.

Regarding future lines of work using current technologies in user state detection, it should be taken into account that automatic user state detection gives information to the system about a factor that directly influences decision making, and consequently the probability of incurring one of the risks that arise from the interaction between users. Therefore, it is desirable that a system that guides users navigating exploits the potential of the extensive amount of sentiment analysis techniques in the existing literature. Secondly, combining different sources of data has shown to help improve the system performance when detecting emotional states, and also using sentiment analysis together with stress analysis. In this way, future works could investigate new ways of combining different kinds of data in fusion approaches, and also use different analysis such as stress analysis together with sentiment analysis to try to improve the system performance. It could allow researchers to discover a correlation between different variables such as stress levels of users, sentiment or other factors, and at the same time it might improve the system performance to use detection of multiple aspects of the user state to perform a guiding or recommendation, therefore it is an interesting possibility to test.

Another potential branch of future lines of work where there is room for improvement is investigating the effect of a combination of different technologies compatible with the architecture of SNSs with emotion detection techniques, and the effectiveness of such systems in different real-life scenarios (to guide users by analyzing their data and giving them feedback). CBR techniques,

as has been reviewed, can be used to work together with sentiment analysis techniques for example exploiting different sentiment lexicons. Since CBR systems can be integrated easily into a SNS as a guiding module that helps users by monitoring their interactions and giving them feedback based on the different characteristics of the interaction and user states, they are potential candidates for designing user guiding systems that create a more satisfactory and safe user experience. Moreover, recommendation systems that use system data to give recommendations to users could potentially be improved, using, for example, persuasion techniques and sentiment or stress analysis. Finally, MAS-based approaches can fit in the SNS architecture, for example assigning agents to users and different system tasks to other agents. These systems can work together with user state detection techniques as user guiding and recommendation systems, by using the data collected by the automatic user state detection when users interact to give advice or recommendations. Therefore investigating the effect of combining different user state detection techniques with MAS-based approaches might improve the performance of the system as a guiding or recommendation system. It might be interesting to also test other technologies such as Peer-to-peer and Internet of things to work as guiding and recommendation systems, integrated into a SNS or other social environment, since they can allow users to share information in a distributed way, and the system could use this information to perform its guiding and recommendation functions.

Finally, related to the user state detection techniques, new fusion techniques, that use feature-level, decision-level or hybrid fusion could be tested to analyze a potential improvement in emotion detection, as has been shown in the literature that some approaches manage to beat the accuracy of non-fusion techniques with a fusion technique using the same data. It might be an interesting research to also apply fusion techniques to stress analysis and other aspects of the user state (e.g., fusion of text, keystroke dynamics data and images to determine the tiredness of users, or to perform fusion of image data and text data to determine the level of interaction with other users that a given user has in a SNS), and see if there are differences in accuracy achieved between unimodal and multi-modal techniques. To summarize, the recent literature contains several works in both automatic user state detection and user guiding and recommendation in online social environments, but there is still plenty of potential improvement in those lines of work, by improving the accuracy of user state detection models and testing the usefulness of different technologies and combinations in user guiding and recommendation systems. Moreover, improving the feedback given to users to create a better understanding of potential risks and a better response of the users for avoiding them could be implemented in guiding systems.

Author Contributions: Conceptualization, G.A., V.J., A.G.-F. and A.E.; Formal analysis, G.A., V.J., A.G.-F. and A.E.; Investigation, G.A., V.J., A.G.-F. and A.E.; Methodology, G.A., V.J., A.G.-F. and A.E.; Supervision, G.A., V.J., A.G.-F. and A.E.; Validation, G.A.; Visualization, G.A., V.J., A.G.-F. and A.E.; Writing—original draft, G.A.; Writing—review & editing, G.A., V.J., A.G.-F. and A.E. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the project TIN2017-89156-R of the Spanish government.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Vanderhoven, E.; Schellens, T.; Vanderlinde, R.; Valcke, M. Developing educational materials about risks on social network sites: A design based research approach. *Educ. Technol. Res. Dev.* **2016**, *64*, 459–480. [[CrossRef](#)]
2. De Moor, S.; Dock, M.; Gallez, S.; Lenaerts, S.; Scholler, C.; Vleugels, C. Teens and ICT: Risks and Opportunities. Belgium: TIRO. 2008. Available online: <http://www.belspo.be/belspo/fedra/proj.asp?l=en&COD=TA/00/08> (accessed on 25 April 2020).

3. Livingstone, S.; Haddon, L.; Görzig, A.; Ólafsson, K. Risks and Safety on the Internet: The Perspective of European Children: Full Findings and Policy Implications From the EU Kids Online Survey of 9–16 Year Olds and Their Parents in 25 Countries. EU Kids Online, Deliverable D4. EU Kids Online Network, London, UK. 2011. Available online: <http://eprints.lse.ac.uk/33731/> (accessed on 25 April 2020).
4. Vandenhoven, E.; Schellens, T.; Valacke, M. Educating teens about the risks on social network sites. *Media Educ. Res. J.* **2014**, *43*, 123–131. [[CrossRef](#)]
5. Christofides, E.; Muise, A.; Desmarais, S. Risky disclosures on Facebook: The effect of having a bad experience on online behavior. *J. Adolesc. Res.* **2012**, *27*, 714–731. [[CrossRef](#)]
6. George, J.M.; Dane, E. Affect, emotion, and decision making. *Organ. Behav. Hum. Decis. Process.* **2016**, *136*, 47–55. [[CrossRef](#)]
7. Thelwall, M. TensiStrength: Stress and relaxation magnitude detection for social media texts. *Inf. Process. Manag.* **2017**, *53*, 106–121. [[CrossRef](#)]
8. Thelwall, M.; Buckley, K.; Paltoglou, G.; Cai, D.; Kappas, A. Sentiment strength detection in short informal text. *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 2544–2558. [[CrossRef](#)]
9. Shoumy, N.J.; Ang, L.M.; Seng, K.P.; Rahaman, D.M.; Zia, T. Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals. *J. Netw. Comput. Appl.* **2020**, *149*, 102447. [[CrossRef](#)]
10. Zhang, C.; Zeng, D.; Li, J.; Wang, F.Y.; Zuo, W. Sentiment analysis of Chinese documents: From sentence to document level. *J. Am. Soc. Inf. Sci. Technol.* **2009**, *60*, 2474–2487. [[CrossRef](#)]
11. Hu, M.; Liu, B. Mining opinion features in customer reviews. In *AAAI*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 4, pp. 755–760.
12. Jakob, N.; Gurevych, I. Extracting opinion targets in a single-and cross-domain setting with conditional random fields. In Proceedings of the 2010 conference on empirical methods in natural language processing. Association for Computational Linguistics, Cambridge, MA, USA, 9–11 October 2010; pp. 1035–1045.
13. Lu, B.; Ott, M.; Cardie, C.; Tsou, B.K. Multi-aspect sentiment analysis with topic models. In Proceedings of the 2011 11th IEEE International Conference on Data Mining Workshops, Vancouver, BC, Canada, 11 December 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 81–88. [[CrossRef](#)]
14. Popescu, A.M.; Etzioni, O. Extracting product features and opinions from reviews. In *Natural Language Processing and Text Mining*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 9–28.
15. Nasukawa, T.; Yi, J. Sentiment analysis: Capturing favorability using natural language processing. In Proceedings of the 2nd international conference on Knowledge capture, Sanibel Island, FL, USA, 23–25 October 2003; ACM: New York, NY, USA, 2003; pp. 70–77. [[CrossRef](#)]
16. Li, F.; Han, C.; Huang, M.; Zhu, X.; Xia, Y.J.; Zhang, S.; Yu, H. Structure-aware review mining and summarization. In Proceedings of the 23rd International Conference on Computational Linguistics, Beijing, China, 23–27 August 2010; Association for Computational Linguistics: Stroudsburg, PA, USA, 2010; pp. 653–661.
17. Borth, D.; Ji, R.; Chen, T.; Breuel, T.; Chang, S.F. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In Proceedings of the 21st ACM International Conference on Multimedia, Barcelona, Spain, 21–25 October 2013; ACM: New York, NY, USA, 2013, pp. 223–232. [[CrossRef](#)]
18. You, Q.; Luo, J.; Jin, H.; Yang, J. Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks. In Proceedings of the AAAI'15: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, TX, USA, 25–30 January 2015; pp. 381–388.
19. Kaushik, L.; Sangwan, A.; Hansen, J.H. Sentiment extraction from natural audio streams. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 8485–8489.
20. Deb, S.; Dandapat, S. Emotion classification using segmentation of vowel-like and non-vowel-like regions. In *IEEE Transactions on Affective Computing*; IEEE: Piscataway, NJ, USA, 2019; Volume 10, pp. 360–373. [[CrossRef](#)]
21. Deng, J.; Zhang, Z.; Marchi, E.; Schuller, B. Sparse autoencoder-based feature transfer learning for speech emotion recognition. In Proceedings of the 2013 Humaine Association Conference on Affective Computing And Intelligent Interaction, Geneva, Switzerland, 2–5 September 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 511–516. [[CrossRef](#)]

22. Mairesse, F.; Polifroni, J.; Di Fabbrizio, G. Can prosody inform sentiment analysis? experiments on short spoken reviews. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 25–30 March 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 5093–5096.
23. Kanluan, I.; Grimm, M.; Kroschel, K. Audio-visual emotion recognition using an emotion space concept. In Proceedings of the 2008 16th European Signal Processing Conference, Lausanne, Switzerland, 25–29 August 2008; IEEE: Piscataway, NJ, USA, 2008; pp. 1–5.
24. Chen, J.; Hu, B.; Xu, L.; Moore, P.; Su, Y. Feature-level fusion of multimodal physiological signals for emotion recognition. In Proceedings of the 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, USA, 9–12 November 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 395–399.
25. Nicolaou, M.A.; Gunes, H.; Pantic, M. Continuous prediction of spontaneous affect from multiple cues and modalities in valence-arousal space. *IEEE Trans. Affect. Comput.* **2011**, *2*, 92–105. [[CrossRef](#)]
26. Hossain, M.S.; Muhammad, G.; Alhamid, M.F.; Song, B.; Al-Mutib, K. Audio-visual emotion recognition using big data towards 5G. *Mob. Netw. Appl.* **2016**, *21*, 753–763. [[CrossRef](#)]
27. Zhou, F.; Jianxin Jiao, R.; Linsey, J.S. Latent customer needs elicitation by use case analogical reasoning from sentiment analysis of online product reviews. *J. Mech. Des.* **2015**, *137*, 071401. [[CrossRef](#)]
28. Ohana, B.; Delany, S.J.; Tierney, B. A case-based approach to cross domain sentiment classification. In Proceedings of the International Conference on Case-Based Reasoning, Lyon, France, 3–6 September 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 284–296.
29. Ceci, F.; Goncalves, A.L.; Weber, R. A model for sentiment analysis based on ontology and cases. *IEEE Lat. Am. Trans.* **2016**, *14*, 4560–4566. [[CrossRef](#)]
30. Vizer, L.M.; Zhou, L.; Sears, A. Automated stress detection using keystroke and linguistic features: An exploratory study. *Int. J. Hum. Comput. Stud.* **2009**, *67*, 870–886. [[CrossRef](#)]
31. Feldman, R. Techniques and applications for sentiment analysis. *Commun. ACM* **2013**, *56*, 82–89. [[CrossRef](#)]
32. Schouten, K.; Frasincar, F. Survey on aspect-level sentiment analysis. *IEEE Trans. Knowl. Data Eng.* **2016**, *28*, 813–830. [[CrossRef](#)]
33. Ji, R.; Cao, D.; Zhou, Y.; Chen, F. Survey of visual sentiment prediction for social media analysis. *Front. Comput. Sci.* **2016**, *10*, 602–611. [[CrossRef](#)]
34. Li, L.; Cao, D.; Li, S.; Ji, R. Sentiment analysis of Chinese micro-blog based on multi-modal correlation model. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 4798–4802. [[CrossRef](#)]
35. Gunawardhane, S.D.; De Silva, P.M.; Kulathunga, D.S.; Arunatileka, S.M. Non invasive human stress detection using key stroke dynamics and pattern variations. In Proceedings of the 2013 International Conference on Advances in ICT for Emerging Regions (ICTer), Colombo, Sri Lanka, 11–15 December 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 240–247.
36. Lee, P.M.; Tsui, W.H.; Hsiao, T.C. The influence of emotion on keyboard typing: An experimental study using auditory stimuli. *PLoS ONE* **2015**, *10*, e0129056. [[CrossRef](#)]
37. Bradley, M.M.; Lang, P.J. *The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective Ratings of Sounds and Instruction Manual*; Tech. Rep. B-3; University of Florida: Gainesville, FL, USA, 2007.
38. Matsiola, M.; Dimoulas, C.; Kalliris, G.; Veglis, A.A. Augmenting user interaction experience through embedded multimodal media agents in social networks. In *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications*; IGI Global: Pennsylvania, PA, USA, 2018; pp. 1972–1993. [[CrossRef](#)]
39. Rosaci, D. CILIOS: Connectionist inductive learning and inter-ontology similarities for recommending information agents. *Inf. Syst.* **2007**, *32*, 793–825. [[CrossRef](#)]
40. Buccafurri, F.; Comi, A.; Lax, G.; Rosaci, D. Experimenting with certified reputation in a competitive multi-agent scenario. *IEEE Intell. Syst.* **2015**, *31*, 48–55. [[CrossRef](#)]
41. Rosaci, D.; Sarnè, G.M. Multi-agent technology and ontologies to support personalization in B2C E-Commerce. *Electron. Commer. Res. Appl.* **2014**, *13*, 13–23. [[CrossRef](#)]
42. Cissée, R.; Albayrak, S. An agent-based approach for privacy-preserving recommender systems. In Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems, Honolulu, HI, USA, 14–18 May 2007; Association for Computing Machinery: New York, NY, USA, 2007; pp. 1–8.

43. Singh, A.; Sharma, A. MAICBR: A multi-agent intelligent content-based recommendation system. In *Information and Communication Technology for Sustainable Development*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 399–411. [[CrossRef](#)]
44. Villavicencio, C.; Schiaffino, S.; Diaz-Pace, J.A.; Monteserin, A.; Demazeau, Y.; Adam, C. A MAS approach for group recommendation based on negotiation techniques. In *Proceedings of the International Conference on Practical Applications of Agents and Multi-Agent Systems*, Sevilla, Spain, 1–3 June 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 219–231. [[CrossRef](#)]
45. Rincon, J.; de la Prieta, F.; Zanardini, D.; Julian, V.; Carrascosa, C. Influencing over people with a social emotional model. *Neurocomputing* **2017**, *231*, 47–54. [[CrossRef](#)]
46. Upadhyay, A.; Chaudhari, A.; Ghale, S.; Pawar, S. Detection and prevention measures for cyberbullying and online grooming. In *Proceedings of the 2017 International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, India, 19–20 January 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–4.
47. Aguado, G.; Julian, V.; Garcia-Fornes, A.; Espinosa, A. A Multi-Agent System for guiding users in on-line social environments. *Eng. Appl. Artif. Intell.* **2020**, *94*, 103740. [[CrossRef](#)]
48. Aguado, G.; Julián, V.; García-Fornes, A.; Espinosa, A. Using Keystroke Dynamics in a Multi-Agent System for User Guiding in Online Social Networks. *Appl. Sci.* **2020**, *10*, 3754. [[CrossRef](#)]
49. Camara, M.; Bonham-Carter, O.; Jumadinova, J. A multi-agent system with reinforcement learning agents for biomedical text mining. In *Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics*, Atlanta, GA, USA, 9–12 September 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 634–643. [[CrossRef](#)]
50. Almashraee, M.; Diaz, D.M.; Unland, R. Sentiment Classification of on-line Products Based on Machine Learning Techniques and Multi-agent Systems Technologies. In *Proceedings of the Industrial Conference on Data Mining-Workshops*, Berlin, Germany, 13–20 July 2012; IBAI Publishing: Fockendorf, Germany, 2012; pp. 128–136.
51. Lombardo, G.; Fornacciari, P.; Mordonini, M.; Tomaiuolo, M.; Poggi, A. A multi-agent architecture for data analysis. *Future Internet* **2019**, *11*, 49. [[CrossRef](#)]
52. Schweitzer, F.; Garcia, D. An agent-based model of collective emotions in online communities. *Eur. Phys. J. B* **2010**, *77*, 533–545. [[CrossRef](#)]
53. El Fazziki, A.; Ennaji, F.Z.; Sadiq, A.; Benslimane, D.; Sadgal, M. A MULTI-AGENT BASED SOCIAL CRM FRAMEWORK FOR EXTRACTING AND ANALYSING OPINIONS. *J. Eng. Sci. Technol.* **2017**, *12*, 2154–2174.
54. Bordera, J. PESEDIA. Red Social Para Concienciar en Privacidad. Master's Thesis, Universitat Politècnica de València, Valencia, Spain, 2016.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).