Resumen

La manera en la cual se expresan los usuarios en redes sociales ha atraído la atención de varias disciplinas debido al gran potencial de investigación que esta representa. Las plataformas de redes sociales, como Twitter, ofrecen a los usuarios la posibilidad de expresarse de forma libre y espontanea haciendo uso de diferentes recursos lingüísticos como la ironía para lograr diferentes propósitos de comunicación. Las expresiones irónicas en dichas plataformas son generadas por usuarios que, la mayoría de las veces, tienen solamente una definición intuitiva de lo que es la ironía. Manejar ese tipo de contenido representa un gran reto para la lingüística computacional. La ironía está estrechamente vinculada con la expresión indirecta de sentimientos, emociones y evaluaciones (entendidas como la actitud o postura del autor hacia la entidad específica involucrada en la expresión irónica). Por consiguiente, el interés en detectar la presencia de ironía en textos de redes sociales ha aumentado significativamente en los últimos años, también por el impacto en áreas del procesamiento de lenguaje natural (PLN) relacionadas con el análisis de sentimientos, donde detectar la presencia de la ironía es importante para evitar malinterpretar expresiones irónicas como literales.

En esta tesis, introducimos el problema de detección de ironía en redes sociales desde una perspectiva de la lingüística computacional. Proponemos abordar dicha tarea enfocándonos, particularmente, en el rol de información relativa al afecto y las emociones para detectar la presencia de dicho recurso lingüístico. Con la intención de aprovechar el valor intrínseco de subjetividad contenido en las expresiones irónicas, presentamos un modelo para detectar la presencia de ironía denominado emotIDM, el cual está basado en una amplia variedad de rasgos afectivos. Para caracterizar instancias irónicas, utilizamos un amplio conjunto de recursos que cubren diferentes ámbitos afectivos: desde *sentimientos* (positivos o negativos) hasta *emociones* específicas definidas con una granularidad fina. Abordamos la detección de ironía interpretándola como un problema de clasificación binaria. Para evaluar nuestro modelo, recopilamos un conjunto de datos de Twitter previamente utilizados en trabajos relacionados a la detección de ironía teniendo en cuenta dos objetivos: por un lado, comparar el desempeño de nuestro modelo con otros enfoques en el estado del arte, y por el otro, evaluar su robustez considerando varios aspectos diferentes, los cuales están relacionados con las características de los conjuntos de datos como modalidad de obtención, tamaño y grado de desbalance entre clases. Los resultados obtenidos muestran que emotIDM tiene un desempeño competitivo en los experimentos realizados, validando la efectividad del enfoque propuesto. En la mayoría de los casos, nuestros resultados superaron el estado del arte confirmando que la información relativa al afecto y las emociones ayuda para distinguir entre tweets irónicos y no irónicos.

Otro objetivo de la tesis es investigar las diferencias entre tweets etiquetados con #irony y #sarcasm. Nuestra finalidad es contribuir a un tema menos investigado en lingüística computacional: la separación entre el uso de ironía y sarcasmo en redes sociales, con especial énfasis en rasgos afectivos. Además, estudiamos un hashtag que ha sido menos analizado:

#not, el cual ha sido utilizado en investigaciones relacionadas a la detección de contenidos irónicos para recopilar datos que expresan una intención sarcástica. Nuestros resultados parecen evidenciar que existen diferencias entre los tweets que contienen dichos hashtags, sugiriendo que son utilizados para hacer referencia de diferentes recursos lingüísticos. Identificamos un conjunto de características basadas en diferentes fenómenos afectivos que parecen ser útiles para discriminar entre diferentes tipos de recursos lingüísticos, además nuestros resultados superaron aquellos en el estado del arte. Adicionalmente analizamos la reversión de polaridad en tweets que contienen hashtags irónicos, observamos que el impacto de dicho fenómeno es diferente en cada uno de ellos. En el caso de los tweets que están etiquetados con el hashtag #sarcasm, a menudo hay una reversión total (que va desde una polaridad a su opuesta, casi siempre de un sentimiento positivo a uno negativo), mientras que en el caso de los tweets etiquetados con el hashtag #irony se produce una atenuación de la polaridad (principalmente de negativo a neutral).

Detectar ironía en contenido generado por usuarios podría tener un amplio rango de aplicaciones. Sin duda alguna, una de las áreas que puede resultar más beneficiada por la detección de ironía es el análisis de sentimientos. Llevamos a cabo un estudio del impacto de la ironía y el sarcasmo en el análisis de sentimientos, observamos una disminución en el rendimiento de los sistemas de PLN desarrollados para dicha tarea cuando la ironía está presente. Por consiguiente, exploramos la posibilidad de utilizar nuestros resultados en detección de ironía para el desarrollo de un sistema de análisis de sentimientos que considere de la presencia de ironía, suponiendo que la detección de contenido irónico podría ayudar a mejorar la correcta identificación del sentimiento expresado en un texto dado. Con este objetivo, incorporamos emotIDM como la primera fase en un sistema de análisis de sentimientos para determinar la polaridad de mensajes en Twitter. Comparamos nuestros resultados con el estado del arte establecido en la tarea de evaluación `Semeval-2015 Task 11: Sentiment Analysis of Figurative Language in Twitter', demostrando la importancia de utilizar información afectiva en conjunto con características que alertan de la presencia de la ironía para desempeñar análisis de sentimientos en textos con lenguaje figurado que provienen de redes sociales. En resumen, demostramos la utilidad de aprovechar diferentes aspectos de información relativa al afecto y las emociones para tratar cuestiones relativas a la presencia de la ironía en Twitter.