

El contenido sexista en las redes sociales suele ser más implícito y sutil que violento, revela EXIST 2021

El contenido sexista en las redes sociales suele ser más implícito y sutil que violento, revela EXIST 2021

Profesores y doctorandos de la UNED participan en EXIST 2021, una campaña para detectar mensajes sexistas en Internet. EXIST 2021 propone a la comunidad científica la tarea de detectar y categorizar automáticamente mensajes sexistas en redes sociales, clasificándolos en sexistas o no sexistas, y categorizando a los sexistas de acuerdo con la faceta de la mujer a la que atacan. Para ello, han creado un *dataset* de más de 11.000 mensajes cortos provenientes de Twitter y Gab.com que contienen más de 200 expresiones sexistas detectadas por los organizadores. EXIST ha recibido una gran acogida con la participación de más de 30 equipos de 11 países.

¿Cómo suena el sexismo en las redes sociales? No siempre igual, no siempre tan obvio y, a veces, puede parecer hasta amable o “gracioso”. Veamos: “Las mujeres deben ser amadas y respetadas, trátalas siempre como un vidrio frágil” ó “Hay que amar a las mujeres... solo eso... nunca las entenderás”. Ambos ejemplos los aporta el profesor Jorge Carrillo de Albornoz, profesor de la ETSI de la UNED. “En el primer caso se esconde la idea de que las mujeres son más débiles que los hombres; en el segundo, puede parecer una broma y es un chiste típico de humor sexista. En los casos más extremos, el sexismo puede sonar ofensivo e incluso incitar al odio. Por eso es complicado detectar automáticamente el sexismo *online*, porque se expresa de maneras muy diferentes.

“Hemos encontrado una gran cantidad de conversaciones en español, en este caso concreto procedentes de Twitter, de contenido misógino. Dado que Twitter es una red social con mecanismos de censura y políticas de uso bastante restrictivas, esperábamos encontrar pocos tweets para la categoría de misoginia. Del mismo modo, hemos encontrado un alto porcentaje de tweets en inglés para la categoría “violencia sexual”, cuenta Carrillo, y añade que, “el porcentaje de contenido sexista que podríamos considerar implícito o sutil en el *dataset* es considerablemente más alto que el explícitamente violento, lo que corrobora que es imprescindible avanzar en la detección automática de este tipo de sexismo”.

“La campaña EXIST pretende impulsar el desarrollo de algoritmos automáticos capaces de detectar el sexismo en un sentido amplio, desde la misoginia explícita hasta otras expresiones sutiles que involucran comportamientos sexistas implícitos. Se trata de una tarea muy compleja debido al uso de vocabulario neutro y la dependencia con el contexto en que se produce la conversación”, explica el profesor.

La campaña de evaluación EXIST se organiza como parte del proyecto MISIMIS (Misinformation and Miscommunication in Social Networks) en el que participan investigadores de la UNED, la Universidad Politécnica de Valencia y la Universidad de Barcelona, como un instrumento para potenciar la investigación en detección automática de sexismo en la red, fundamentalmente en conversaciones en redes sociales. “Investigaciones anteriores han estudiado la presencia de contenidos misóginos y violentos hacia la mujer en Internet, pero nunca antes se han tratado el sexismo en su globalidad, que incluye formas sutiles, amables e incluso encubiertas”, señala Carrillo.

El equipo de Trinidad Donoso, de la Universidad de Barcelona, referente nacional en el estudio del sexismo en la red, ha sido el encargado de identificar un conjunto de términos que, con frecuencia, se presentan en conversaciones sexistas, pero que, dependiendo del contexto de la conversación, pueden también presentarse en conversaciones sin ninguna connotación sexista. También han elaborado una categorización de los tipos de sexismo que con mayor frecuencia encontramos en redes sociales, tomando como referencia las distintas facetas de la mujer que son atacadas o menoscabadas. A partir de estos términos, el equipo de la UNED ha recopilado un *dataset* de conversaciones que los contienen y que han sido anotados manualmente como sexistas o no sexistas, y se asigna así a cada conversación sexista una categoría conforme a la categorización elaborada por las expertas.

“Durante la competición EXIST se ha pedido a los equipos participantes que desarrollen sistemas automáticos capaces de clasificar automáticamente



los textos como sexistas o no sexistas y, en caso de ser sexistas, clasificarlos según el tipo de sexismo: misoginia, violencia sexual, objetivación de la mujer, de estereotipo y dominancia, e ideológico”, señala Carrillo.

El profesor explica que el objetivo no es identificar términos sexistas sino detectar conversaciones con connotaciones sexistas. “De hecho, frecuentemente, y máxime cuando se trata de un sexismo sutil, las conversaciones no presentan términos que, *per sé*, puedan considerarse sexistas, pero, sin hacer uso de un vocabulario específico u ofensivo, encierran mensajes que menoscaban o discriminan a la mujer”, advierte.

La detección automática de conversaciones sexistas en la red tiene muchas aplicaciones prácticas, todas ellas de gran impacto social. Tal y como explica el profesor Carrillo, “por un lado, proporciona valiosísima información estadística para conocer qué tipos de sexismo predominan en la red, a qué colectivos afecta, en qué zonas geográficas de nuestro país se concentra, qué características sociodemográficas tienen las víctimas, cómo se propagan los contenidos sexistas, etc. Toda esta información redundará en una mejor toma de decisiones y permitirá a las instituciones y empresas un mejor desarrollo de políticas de igualdad. Por otro lado, puede utilizarse para el desarrollo de filtros y alertas automáticas que permitirá detectar las formas violentas y potencialmente peligrosas para detener comportamientos que puedan incluso amenazar la vida o poner en riesgo la salud física y mental de las mujeres, especialmente, adolescentes y mujeres jóvenes”.
