



EL SENTIR DE LOS 'TUIITS'. LA UPV/EHU Y ELHUYAR DESARROLLAN UN SISTEMA PARA CONOCER LA OPINIÓN DE LOS CIUDADANOS EN LAS REDES SOCIALES // P8



EL SENTIR DE LOS TUIITS

La UPV/EHU y Elhuyar desarrollan un sistema para conocer la opinión de los ciudadanos en las redes sociales

VITORIA – La UPV/EHU y Elhuyar han desarrollado un sistema para conocer la opinión vertida por la ciudadanía vasca en las redes sociales a través de una herramienta que analiza los sentimientos expresados en tuits en euskera y que en la actualidad, todavía está en fase de trabajo, presenta una tasa de acierto cercana al 75%. El grupo IXA de la UPV/EHU y Elhuyar han desarrollado un sistema para extraer y clasificar las opiniones sobre temas concretos publicados en Twitter, superando dificultades como la escritura no estándar o la mezcla de distintas lenguas, según ha informado la universidad vasca en un comunicado.

El principal autor de este trabajo,

Iñaki San Vicente Roncal, subraya que internet y, especialmente, las redes sociales, son "una generosa fuente de contenidos generados por los propios usuarios", una información que tiene una importancia vital para organizaciones y empresas, dado que les permite conocer la opinión de los ciudadanos; entre ellos, sus potenciales usuarios o consumidores. San Vicente recuerda que tradicionalmente esta información se obtenía mediante encuestas o cuestionarios que presentaban el problema de que las muestras eran pequeñas. Actualmente, sin embargo, se generan ingentes cantidades de datos en las redes sociales, y, en general, en internet. En esta situación, el reto consiste en extraer

y clasificar correctamente la información de interés. El área de investigación conocida como *Análisis de sentimientos* busca métodos automáticos para determinar si un texto dado expresa alguna opinión o sentimiento positivo o negativo. El sistema desarrollado ahora es fruto de la colaboración entre el grupo IXA de la Facultad de Informática de la UPV/EHU y la Unidad de Lengua y Tecnología de la Fundación Elhuyar.

Los autores han trabajado los mensajes escritos en euskera, tanto los exclusivamente escritos en esa lengua como los que la incluían entre otras, principalmente de la red social Twitter. El primer paso para crear el sistema de análisis de sentimiento en

DATOS

75%

PORCENTAJE DE ACIERTO. La herramienta, aún en fase de trabajo, presenta en la actualidad una tasa de acierto cercana al 75%

2016

ESTRENO. La primera aplicación del sistema fue *Behagunea*, para el análisis de sentimiento de las declaraciones vertidas en Twitter relacionadas con Donostia capital cultural.

textos en euskera fue generar la polaridad léxica, es decir, crear listas con las palabras que de por sí tienen connotación positiva o negativa. En la creación de estas listas hay que tener en cuenta, no obstante, el tema o contexto en el que estás trabajando, ya que algunas palabras pueden tener una polaridad contraria dependiendo del contexto.

Para resolver esos posibles problemas se desarrolló un programa informático que extrajera la información de los textos. Además, el grupo de investigación tuvo que tener en cuenta las particularidades que se dan en las redes sociales como Twitter en las que el lenguaje que se utiliza es singular, muy parecido al

lenguaje hablado.

Toda la información recopilada fue empleada para *entrenar* los sistemas de aprendizaje automático. De esa forma se generaron miles de ejemplos, debidamente clasificados, con los que se alimentó el sistema con ellos, al que también se definió cuáles son las características en las que ha de reparar. La primera aplicación del sistema desarrollado fue *Behagunea*, mediante la cual hicieron el análisis de sentimiento de las declaraciones vertidas en Twitter relacionadas con la capitalidad europea de la cultura de Donostia en 2016.

También se desarrolló un proyecto con el diario *Berria*, con el que se hizo el seguimiento de las elecciones autonómicas vascas de 2016. El grupo de investigación también ha trabajado con el instituto vasco de criminología estudiando la percepción sobre las víctimas del terrorismo en redes sociales. El objetivo ahora es mejorar la tasa de acierto, que ronda el 75%. - E.P.