



Deia. Noticias de Bizkaia - Noticias de última hora de Bizkaia, Euskadi e internacionales.

Twitter sare sozialean

UPV/EHUK eta Elhuyarrek sare sozialetan emandako iritzia jakiteko sistema abiarazi dute

Datu-kantitate erraldoiak sortzen dira Interneten, eta hortik interesa duen informazioa behar bezala erauzi eta sailkatzea da xedea

Deia.eus - 2019(e)ko irailaren 14a, larunbata - 13:13etan eguneratua



Iñaki San Vicente, IXA taldeko ikerlaria (UPV/EHU)

BILBO. Erabiltzaileak sare sozialetan duten jarrera eta iritzia jakiteko einean, Euskal Herriko Unibertsitateko IXA taldeak eta Elhuyarrek sistema berri bat abiarazi dute. Twitter izango da sistemak aztertuko duen sare sozial nagusia, gutxienez hasiera batean.

Internet, eta, bereziki, sare sozialak, erabiltzaileek sortutakc [Preferencias de privacidad](#)

dira; besteak beste, zernahiri buruzko iritziak ematen dituzte erabiltzaileek, nonahi eta noiznahi, doan. Informazio horrek "berebiziko garrantzia du erakunde, enpresa, eragile eta abarrentzat, beti izan baitute interesa jakiteko herritarrek, erabiltzaileek edo bezeroek zer iritzi duten haiei buruz. Tradizionalki, kostu handia duten inkesta edo galdera-sorta bidez eskuratu izan dute informazioa, baina laginak beti txikiak izan dira", dio Iñaki San Vicente Roncal UPV/EHUko [IXA Taldean](#) zuzendutako lanaren egile nagusiak.

Orain, berriz, datu-kantitate erraldoiak sortzen dira sare sozialetan, eta, oro har, Interneten. Erronka ordea, hortik interesa duen informazioa behar bezala erauzi eta sailkatzea da. Sentimenduen Analisia deritzon ikerketa-arloak (Ingeleseaz, sentiment analysis) edonolako testu batean iritzi edo sentimendu positibo edo negatiboren bat adierazten ote den ebazteko metodo automatikoak bilatzen ditu. "Gu 2011. urtean hasi ginen euskararako lantzen sentimenduen analisia egiteko teknikak", ekarri du gogora San Vicente doktoreak.

UPV/EHUko Informatika Fakultateko IXA Taldeak eta Elhuyar Fundazioko Hizkuntza eta Teknologia Unitateak aspalditik duten lankidetzaren emaitza izan da garatutako sistema. Bereziki, Twitter sare sozialean euskara hutsean idatzitako txioak, edo tartean euskara dutenak, izan dituzte langai. Bidean ez dituzte falta izan erronkak eta zailtasunak, hutsetik abiatu baitzuten lana.

Euskarazko testuetan sentimenduen analisia egiteko sistema sortzeko lehengo pausoa, oinarrikoa, polaritate lexikoak sortzea izan zen, hau da, berez kutsu positiboa edo negatiboa duten hitzen zerrendak osatzea. Adibidez, "txarra" edo "gaiztoa" hitzak beti negatiboak dira, eta "ona" eta "maitagarria", berriz, positiboak. "Zerrenda horiek sortzean, ordea, oso kontuan izan behar da zer gai edo testuingurutarako ari zaren lanean, hitz batzuek kontrako polaritatea izan dezaketelako testuinguruaren arabera. Esate baterako, "gora egin" edo "behera egin" aditzek ez dute kutsu bera eskailerak igo edo jaisteaz ari bagara, burtsako akzioez ari bagara edo langabezia ari bagara", azaldu du San Vicentek.

Lexikoaz gain, hizkuntza orotan gertatzen diren fenomeno linguistikoak ere kontuan hartu behar izan zituzten: "Esate baterako, ezezko esaldiek kontrako zentzua ematen diete hitzei ("hau ez da batere ona"), edo ironiaz esandako adierazpenak antzematea ere garrantzitsua da. Sor daitezkeen arazoak ebazteko, programa informatiko bat garatu genuen, testuetako informazioa erauzi, hitzak lematizatu eta bestelako azterketak egiteko".

Hizkuntzen berezko konplexutasuna gutxi ez, eta Twitterren berezitasunak ere kontuan hartu behar izan zituzten ikertzaileek, hizkera berezia baitago sare sozialetan, ahozko hizkerara asko hurbiltzen dena. "Gramatika ez-estandarra izaten da askotan, eta horrek lana zailtzen die hizkuntza-azterketarako tresnei. Gainera, oso maiz gertatzen dira hizkuntzen arteko nahasketak ("Kaixo. Acabo de hacer la azterketa de gizarte. Fatal atera zait!"), eta halako fenomenoak tratatu egin behar dira", dio San Vicentek.

Zailtasunekin batera, dena den, lagungarri batzuk ere badituzte sare sozialetako adierazpenetan: "Enfasia emateko, hizki larriak erabiltzen dira, edo hitzak luzatzen ("osooo onaaa"), edo harridura-ikur asko jartzen dira; emotikonoen bidez emozioak adierazten dira... Horrek guztiak informazioa ematen digu", gaineratu du.

Aurreko informazio guztia ikasketa automatikoko sistemak entrenatzeko erabili zuten ondoren, San Vicentek azaldu duenez: "Milaka adibide sortu genituen, behar bezala sailkatuta, eta sistemari eman genizkion, baita zer ezaugarri erreparatu behar dion erakutsi ere, hortik aurrera bere kabuz egin zezan adierazpenen azterketa".

[Behagunea](#) izan zen garatutako sistemaren lehenengo aplikazioa, 2016. urteko Donostia Europako Kultur Hiriburutzari lotuta Twitterren esan zirenen sentimenduen analisia egin zuten. "Oso ongi funtzionatu zuen; une jakinetan ekitaldi polemikoak egon ziren, eta horiek denak jaso zituen sistemak. Hiriburutzaren barruan gauzatutako proiektuen ebaluazioetan ere erabili ziren bildutako datuak", deritza San Vicentek. Beste proiektu bat Berria egunkariarekin egin zuten, 2016. urteko Euskal Autonomia I Preferencias de privacidad

autonomikoen kanpainaren segimendua egiteko, “eta hor ere ez ziren falta izan polemikak”. Elhuyarreko ikertzaileak UPV/EHUKo Kriminologiaren Euskal Institutuarekin (IVAC-KREI) ere aritu dira lanean, sare sozialetan terrorismoaren biktimei buruzko pertzepzioa nolakoa den aztertzeke. “Guretzat oso garrantzitsuak dira proiektu horiek, tesian oinarritzko ikerketatik industria-aplikazio errealerara bitarteko ibilbidea osorik egin dela erakusten dutelako. —dio San Vicentek—. Egungo sistema gai da euskarazko testuak ez ezik, gaztelania, frantsesa eta ingelesa ere aztertzeke, hankamotz gelditzen baita euskaraz egindako adierazpenen analisisa bakarrik egitea”.

Garatutako sistemen emaitzaren balorazioa ona bada ere, oraindik badago hobetzeko tarterik, eta ikertzaileek lanean dihardute. “Ohiko eredu estatistikoetatik algoritmo neuronalekin lan egitera pasatu gara, eta oso emaitza onak ematen ari da. Helburua da adierazpenen sailkapenean arrakasta- edo asmatze-tasa hobetzea, une honetan % 75 ingurukoa baita”, dio, bukatzeke.

- © Deia - Noticias de Bizkaia
-
-
-
-
-
-