

Elena Lazkano Ortega eta Olatz Arbelaiz Gallego

Donostiako UPV/EHUko Informatika Fakultateko irakasle eta ikertzaileak

Ba al du adimen artifizialak genero joerarik?

Adimen artifiziala edo makina adimena, makinek erakusten duten adimena da; jendarteak eta animaliek aldiz, adimen naturala dute. Gaur egun, adimen artifizialaren helburu nagusiak arrazonaimendua, ezagutzaren adierazpena, ikastea, lengoaia naturalaren prozesamendua eta objektuen identifikazio eta manipulazioa dira besteak beste; eta haren azpitik dagoena informatika, estatistika, matematika, psikologia, linguistika, filosofia eta beste hainbat jakintza arlo.

Konpainia asko dira dagoeneko ikasketa automatikoa eta irudi errekonozimendua erabiltzen dutenak datuei etekina ateratzeko. Hala nola, mailegu-eskaeren sinesgarritasuna balioztatzeke, minbizia duten pertsonentzat tratamendu egokia hautatzeko, iruzurra detektatzeko, eta abar.

Ikasketa automatikoko algoritmoek datuak behar dituzte ikasteko, datu-base erraldoiak miatuz ezaugarriak eratortzen dituzte gero informazio berria sailkatu ahal izateko. Algoritmoek ikasten dutena, datuetan, matematikan eta estatistikan oinarritzen da batik bat. Ondorioz, ikasteko erabiltzen diren datuek eta egiturek joera bereziren bat badute, ikasitakoan islatuko da hori. Horregatik, maiz, gure inguruan ditugun sistemek genero eta arraza joerak izaten dituzte.

Horrelakoen lehen adibide nabarmenak aspaldi agertu ziren, ahotsa errekonozitzeko softwarearekin batera. Duela hamabost urte, emakumeek zailtasunak zituzten autoak gidatzeko ahotsean oinarritutako sistemak funtzionarazteko, haien ahotsaren doinua entrenamenduetarako erabilitakoaren desberdina zelako.

Ordutik hona sistemak asko hobetu diren arren, haien joerak oraindik arazo dira. Adibidez, bilaketa-tresna informatikoei sarri jotzen dute Wikipedia-ra, baina, Wikipedian bertan, sarrera biografikoen %18 baino ez dira emakumezkoenak. Bilaketan emaitzak, beraz, ezin dira genero aldetik parekatuta egon.

Lengoaia naturalaren prozesamenduan hitzen arteko erlazioak adierazteko erabiltzen diren bektore distribuzionalak (*word-embeddings*) aztertuz dituzte Bostongo hainbat ikertzailek, eta konturatu dira bektoreek dituzten desbiderapenak direla medio, bektoreotan oinarritutako sistemek, adibidez, Jon izeneko pertsonak Maria izenekoak baino ordenagailu programa hobek egiten dituela ondorioztatzeke joera dutela.

Softwareak argazkien edukia hobeto uler dezan erabiltzen diren datu-baseetan ere antzeman dira genero joerak. Washingtongo Unibertsitatean sortutako ImSitu eta hasiera batean Microsoft eta gaur egun Facebook eta Microsoft *AI*. *startup*-a atzean dituen COCO datu-baseek, adibidez, argazki gehiago

dituzte gizonenak emakumeenak baino, eta genero bakoitzeko pertsonekin batera ageri diren irudi eta aktibitateek genero joera handia dute. Izan ere, sukaldeko objektuak emakumeekin lotuta ageri dira, eta, aire libreko kirol-materiala, aldiz, gizonekin.

Ahots laguntzaileak ere ez dira joera horretatik at geratzen. Hain zehatzak diren Sirik, Alexak eta Cortanak, emakume izena eta ahots lehenetsia dituzte; eta berriagoa den Autodesk enpresaren AVA alegiazko laguntzaileak, berriz, ahotsaz gain, emakume itxura ere badu. Horrela, jendartearen hain iltzatuta dagoen emakume laguntzaile eta zaintzailearen irudia errepikatu eta hedatzen dute.

Egokia da asumitzea horrelako sistemak sortzeko erabiltzen ditugun datu guztietan desbiderapena dagoela. Desbiderapenak identifikatu eta zuzentzen saia gaitezke klase desberdinen sailkapenerako algoritmo desberdinak erabiliz edota datuetatik problematikoa diren adibideak zein aldagaiak ezabatuz. Baina, oker diseinatutakoa konpondu baino, zergatik ez sortu errotik genero joerarik gabeko sistema adituak? Erraza ez den arren, entrenamendurako genero joerarik gabeko datuak erabiliz gero, haiekin sortutako sistemek modu aniztiotsuan (jendarte idealarekin lerrotatuta) funtzionatzea lortu daiteke, portaera zuzena zabaltzen lagunduz. •

