

Mikel Galar Idoate

ADIMENTSU NATURALA



KONTUZ IBILI BEHAR DUGU DATUAK EMATERAKOAN, BAINA PARANOIKO BIHURTU GABE



Giza adimena beharrezkoa duten ikaskuntza, arrazoiketa edota pertzepzioa bezalako zereginak egin ditzaketen sistemak sortzera bideratzen den informatikaren arloa da adimen artifiziala (AA). Horretan aditua eta aritua da Mikel Galar irakasle eta ikerlari iruindarra eta itxaropenerako mezu bat eman digu: adimen artifiziala, horren adimentsua bada, gai izanen da hizkuntza gutxituetan ere bikain aritzeko.

 REYES ILINTXETA  ZOE MARTIKORENA / ARGIA CC BY-SA

2011tik zara irakaslea NUPen. Zein da zehazki zure lana?

Irakaslea naiz Adimen Artifizial eta Konputazio sailean. Klasea ematen dut ingeniari informatikako graduan eta masterrean. Lanaren beste aldeak ikerkuntza da. Adimen Artifiziala eta Ikasketen Automatikoaren arloan aritzen naiz gehienbat.

Berriki Cambridgeko unibertsitateak argitara eman duen liburu bat plazaratu duzu beste ikerlari batekin elkarlanean. Nolako da?

Big datari buruzkoa da. Saiatzen da azaltzen nola egin dezakegun datu asko ditugun datu horiek prozesatzeko modu eraginkorrean. Hau ulertzeko esango

genuke guk normalean ordenagailu batekin lan egiten dugula eta erabiltzen dugun informazioa kudeatzeko ez dugula arazorik izaten, baina imajina dezagun Googlek zenbat datu dituen web orriak prozesatzeko, esate baterako. Ordenagailu bat ez da gai informazio hori guztia gordetzeko, beraz, kasu horietan egiten dena da ordenagailu asko aldi berean erabili, datu horiek banatzeko, baina aldi berean lotuta mantenduz informazioa elkarrekin prozesatu ahal izateko. Liburuan azaltzen dugu hori nola egin Python programazioa lengoaiarekin eta Spark izeneko erremintarekin. Ikasketen liburuak da, gradua amaitzen edo masterrean egiten ari direnen ikasleei zuzendua bereziki.

Eta zer da sortu duzuen Neuraptic IA enpresa?

Unibertsitatearen arlo bat teknologiaren transferentzia da. Hemen egiten ditugun ikerketen emaitzak merkaturatzean datza hori. Modu honetara gizarterako onuragarria den zerbait lortzen da eta, aldi berean, ikerkuntzarako dirua ere eskuratzen da. Orain dela zenbait urte, doktore-tza egiten ari ziren bi ikasle eta hirurok hasi ginen enpresentzako hainbat produktu egiten adimen artifiziala erabiliz, eta ikasketak amaitu zituztenean *spin off* hori sortu genuen, hau da, unibertsitatean sortzen den enpresa bat kanpo lanak egiteko. Unibertsitateak etekinaren zatitxo bat hartzen du eta enpresei haien teknologia garatzen laguntzen die.



Mikel Galar Idoate

IRUÑEA, 1986



Ingeniari informatikoa eta NUPeko doktorea da. Konputazioaren Zientzien eta Adimen Artifizialaren arloko irakasle titularra da NUPeko Estatistika, Informatika eta Matematika sailean. Smart Cities Institutuko (ISC) ikertzailea, Konputazio Zientzien eta Adimen Artifizialaren Tracasa katedraren zuzendaria, Neuraptic AI enpresaren sortzaileetako bat eta Emotional Films proiektua garatzen ari den ikerlari taldeko kidea da. 56 ikerketa-proiektutan lan egin du, ehun argitalpen baino gehiago ditu idatziak, Euskalerrria Irratiko kolaboratzailea da eta 2020an, irakasleen bikaintasun saria jaso zuen NUPen. Cambridge University Pressek azaroan argitaratu zuen *Large-Scale Data Analytics with Python and Spark* liburua idatzi du Granadako Unibertsitateko Isaac Triguero Velázquez ikerlariarekin batera.

Iniciate saria eman zizuen Nafarroako Gobernuak 2020an horregatik. Punta-puntako enpresa da Neuraptic IA?

Bai. Enpresa sortzen ari ginenean Nafarroako Gobernuak CEINen bidez sari bat eman zigun. Horri esker negozio plana egin ahal izan genuen dena abiarazteko.

Egia esan, hemen Nafarroan adimen artifizialean oinarritutako enpresa asko ez daude. Aurpegi-ezagutza egiten duten Das-nano edo Veridas agian ezagunenak dira, baina horiek AA erabiltzen dute soilik kasu zehatz horretan. Enpresa gehienak ez daude prest eta horregatik gurea bezalako ekimenak beharrezkoak dira beste enpresei laguntzeko.

Orain berriko Navarra Artificial Intelligence Research Center jarri dute abian. Nola doa?

Hasi da martxan eta oso positiboa da, baina hori unibertsitateetik bultzatu beharko litzateke eta horretarako jende falta dago. Oso irakasle eta ikerlari gutxi gara konputazio eta AA sailean eta horregatik kanpoko jendea ekarri behar da zentro horretara.

Bestalde, oso zaila da jendea unibertsitatean gelditzea, enpresa munduan eskaintzen diren soldatak hementengoen bikoitza direlako. Soldata horiek ezin dugu lehiatu.

Gaur egun informatikako ikasleak ateratzen direnean ez zaie lana falta. Badira ikasleak hirugarren mailan jada lanean hasten direnak, ez praktiketan baizik eta kontratu onekin. Lan asko dago, ez solik big dataekin lotuta, baita software garatzaile gisa ere.

Zeure burua nola ikusten duzu etorkizunean?

Ni hemen gustura nago eta hurrengo urteotan hemen jarraitzeko asmoa dut. Klaseak, ikerketa eta dibulgazio lana ere oso garrantzitsuak direla iruditzen zait.

Emotional Films ekimenean ere bazabiltza. Zer da?

Ikuslearen sentimenduen arabera aldatzen diren pelikulak egin ahal izateko sistema bat sortzea da honen helburua. Horrela, adibidez, eszena batean barre egiten baduzu, sistemak interpretatzen du gustatu zaizula eta bide jakin batetik abiatuko da. Edo Iruñean ari bazara pelikula hori ikusten, ba

agian, protagonistak kalera atera eta Gazteluko plazan daudela ikus daiteke eta Madrilen badaude, Puerta del Solen, adibidez. Erabiltzen dugu daukagun ikuslearen informazio guztia pelikula pertsonalizatu bat sortzeko, momentuan eta denbora errealean.

Proiektua joan den urtean bukatu zen. Teknologia prest dago eta orain enpresa finantzazio bila ari da film oso bat egiteko. Demostratzaile batzuk ditugu hori erakusteko. Walt Disney eta Carl Sagan ziberdizipuluak sortu ditugu. Zuk haiekin hitz egin dezakezu bere ezagupen arloari buruz eta haiek zure emozioak erabiliko dituzte modu batera edo bestera erantzuteko. Beste demostratzaile bat dugu ipuinak kontatzeko. Zuk eskatzen ahal diozu nahi duzunari buruzko ipuin bat, adibidez neska baten istorioa, Iruñean gertatzen dena eta dinosauroak agertzen direna, eta hark AAren bidez asmatzen du, momentuan irudiak denbora errealean emanez.

AA oso aurrerapen garrantzitsua da. Sua, idazketa eta Internetaren pareko aurrerakuntza da?

Hori esaten da, bai. Noraino iritsiko gara? Ikusiko dugu. Aldaketarik handiena iritsi da Chat GPTekin. Izan ere, erabiltzen ez dutenak jada atzean gelditzen ari dira. Gauzak azkarrago egiteko aukera ematen dizu. Baina egia da mugak ere badituela. Oso espektatiba altua daukagu eta segur aski orain joango da hozten pixka bat, ikusiko dugulako oraindik ezin dugula egin hainbeste gauza bat-batean. Hala ere, lan batzuk jada aldatzen hasiak dira.

Zeintzuk izanen dira datozen urteetako berrikuntza nagusiak?

Denok gaude zain ea zer egiten duten enpresa handiek, Googlek, eta Chat GPT sortu zuen Open AI-k, bereziki. Esaten dute aurten aterako dela GPT 5. Orain GPT 4 dago eta denok doan erabiltzen duguna GPT 3.5 da. Google ere hor dago lehian eta duela hilabete pare bat aurkeztu zuen Google Gemini, AA mugikorretan eta ahots bidez erabiltzeko sistema, baina oraindik ez du GPT 4ren maila lortu. Chat GPT oso berrikuntza handia izan zen. Jendeak ez zekien ezer eta bat-batean agertu zen erreminta hau, orain ari dira aurrerapen inkrementalak egiten, baina, printzipioz, nire ustez denbora laburrean beste salto handi bat ikustea zaila izango da.



Ahozko ohiko komunikazioa makinekin izanen da hurrengo erronka nagusia?

Hori badago. Gauza batzuk egiteko teknologia prest dago, baina oraindik ez dute hori dena izateko egin behar den puzzlea muntatu. Zuk instalatzen duzu GPT telefonoan, adibidez, eta hitz egin ahal diozu. Berak testu bihurtuko du eta testu moduan erantzungo dizu.

Windows 11k baditu Windows Copilot eta Office co-pilot. Eskatzen ahal diozu Word testu batean duzun informazioarekin Power Point bat egiteko irudiekin eta guzti. Agian ez da perfektua izango, baina lehenengo lan hori aurreratzen dizu eta zuk bakarrik zuzenketak egin beharko dituzu.

Sistema horiek gure eguneroko lanean ikusiko ditugu eta hori izanen da aldaketarik handiena. Baina argi izan behar dugu beti pertsonak behar direla makinak gauzak ongi egiten ari direla egiaztatzeko. Zuk galdetzen badiozu GP-Tri, adibidez, Pedro Sánchezi buruzko informazioa bitan, agian jaiotze data desberdinak emango dizkizu. Dena zailtzan jarri behar da, hori da gaur egun sistema horien arazorik handiena.

Hori zergatik gertatzen da?

Ulertzea zaila izan liteke. Sistema horiek ez dute memorizatzen Interneten dagoen informazio guztia. Hori Google modukoak izango litzateke. Sistema horiek informazio guztia hartu eta hitz batzuen atzean beste batzuk jartzen dituzte, hor ager-

tzeko probabilitateak kontuan hartuta. Horrela, datu koherente bat emango dizu beti, baina ez dago bermatuta zuzena denik. Zenbat eta datu eta informazio gehiago eskura izan, orduan eta zehatzagoa izango da; hala ere, oraindik askotan ikusten dira zentzua duten erantzunak, baina egiazkoak ez direnak. Horregatik ikasi behar dugu erreminta hauekin oso kritiko izaten.

Itzultzaile automatikoekin bezala, beti erne ibili behar dugu?

Horixe. Pertsona bat behar da beti dena kontrolatzen. Itzultzaileek asko hobetu dute eta normalean ez dute sortzen jadanik existitzen ez den zerbait. Agian ematen dizun ordaina itzulpen literala da eta hizkuntzaren aldetik ez da oso ona izango, baina ez du deus berririk asmatzen. Beste sistema hauekin, aldiz, GP-Tri eskatzen badiozu zerbait itzultzea zerbait itzuliko du, baina gerta daiteke hor erdian beste zerbait sartzea, bere kabuz. Gauzak itzultzeko, printzipioz, hobe da itzultzaile automatikoak erabiltzea.

Zinemaren eta ikus-entzunezkoen alorrean gauza harrigarriak ikusiko ditugu?

Bai, eta jada aktore batzuk hasi dira haien irudi eskubideak saltzen pelikuletan AA bidez moldatutako irudiak erabiltzeko. Aktore batzuk betirako izango ditugu.



Eskatzen badiozu adimen artifizialari erizain baten irudia sortzeko egingen dizu, kasuen %95ean, neska zuri bat"

Ez da gauza bera izango, baina aurpegia eta ahotsa berdinak izango dira.

Aldi berean teknologia hau erabil daiteke *deep-fake*-ak egiteko. Egin deza-kegu edozein politikariren bideoa edozein gauza esaten eta hori oso arriskutsua da. Eta animazioan ere gauza harrigarriak egin daitezke.

Eguneroko bizimoduan ere, hildakoak "berpizteko" aukera izanen dugu?

Halakoak entzuten dira. Lortzen baduzu zure amonaren eguneroko bat, berak idatzitako informazio pila bat, argazkiak, bideoak... Gauza horiek eta irudiak sortzeko sistema bat GPT moduko lengoia sistema batekin konbinatuta, sor dezakezu bere itxura duen pertsonaia bat eta berearekin solas egin. Horrelakoak egiteko teknologia prest dago, baina beste kontu bat da alderdi etikoa. Zer datu erabil dezakezue, norik eman behar duen baimena... Chat GPT entrenatuta dago Interneten dagoen informazio guztia hartzeko, baina agian etikoki dena ezin da erabili. Haiek erabiltzen dute baimena izan ala ez izan. Eskubideen inguruan ere oraindik ez dago oso garbi nola kudeatuko den eta oso gai delikatua da.

Gaur egun mugarik ba al da?

Europar Batasuna da une honetan arau gehien jartzen ari dena eta laster aterako

da AAre inguruko legea. Jarri dituzte muga batzuk: adibidez, ezin duzu AA erabili jendea kontrolatzeko, jakiteko noiz iritsi den lanera eta zer egin duen bertan, edo aurpegi-ezagutza edozertarako. Legeek oraingo ez dute mugatzen sistemak entrenatzeko datuak nondik atera, baizik eta zertarako erabiltzen diren.

Zer ikusi, hura ikasi. Algoritmoak matxistak eta arrazistak dira?

Bai. Joera horiek sesgoak edo hesiak deitzen dira. AAK eta GPTk Interneten dauden datuetatik ikasten dute eta Internet, gehienbat, zuria da. Stable Difussion edo Dall-E bezalako sistemak testua irudi bihurtzeko balio dute eta zuk eskatzen badiezu erizain baten irudia sortzeko, kasuen %95ean neska zuri bat egingen dizute. Eskatzen badiozu MacDonaldek s-eko langile baten irudia, pertsona emigrante baten irudia emango ditu. Inge- niaria, aldiz, zuria eta gizona izango da. Estereotipo horiek munduan daude eta AAK erreproduzitzen eta handiagotzen ditu. Ikertu behar da nola saihestu dai- tekeen hori. Etika kontuetan ere ikusten da nola Open AIk baduen bere araudi propioa, baina hori ere oso kontu arris- kutsua da.

Euskara nola dago itsaso zabal honetan? Atzean gelditzen ari da?

Beste hizkuntza gutxiuak bezala bere arazoak ditu, noski. Chat GPTri euska- raz hitz egiten badiozu zerbait egingo du, baina ingelesez edo gaztelaniaz askoz hobeto moldatuko da, informazio gehiago ikusi duelako entrenatzeko orduan. Gaztelaniaz ere ez du ingelesez bezain ongi funtzionatzen. Dena den, onena da Chat GPT eta horrelako sistemak jada hasi direla hizkuntzen artean loturak egiten, nolabait. Sistemaren burmuinak hizkuntzarik kontuan hartu gabe pentsa- tzen du. Lortu behar da Chat GPTrentzat hizkuntza interfaze bat bezalako ia izatea soilik. Hau da, edozein hizkuntzatan hitz eginda, berak bere gaitasun guztia era- biltzea lan egiteko eta gero zuk aukera- tzen duzun hizkuntzan emaitza ematea. Google Translate eta itzulpen sistema horiek, izan ere, sistema hori erabiltzen dute jada. Esate baterako: zuk testu bat eraman nahi baduzu euskaratik swahili- ra, segur aski ez dago testu asko bi hiz- kuntza horietan idatzita, baina sistemak

“

Lortu behar da GPTri edozein

hizkuntzatan hitz eginda, berak bere gaitasun guztia erabiltzea lan egiteko eta gero zuk aukeratzen duzun hizkuntzan emaitza ematea”

ikasi du informazioa eramaten toki neutro batera, nolabait esateagatik, eta hortik gero edozein hizkuntzara pasatzen. Hau itzultzaileekin egiten bada, zergatik ez beste teknologia horiekin?

Itxaropena da arazo hau bera dagoela hizkuntza as- korekin eta Chat GPTk eta Google- lek badute interesa beste hizkun- tza guztiak erabiliak izan daitezen.

Gaitu proiektua abian da euskarazko ahotsa biltzeko AA entrenatzeko. Datu asko ematea ez da arriskutsua izanen?

Letra txikia irakurri behar, gauza guztie- tan bezala, baina egia da batzuetan bel- dur asko dugula eta gero ez dela hainbes- terako. Gehienok mugikorra daukagu eta informazio pila ematen diogu Googlek. Ez dugu daturik eman nahi, baina aldi berean aprobetxatzen gara datu horien bidez ematen dizkiguten zerbitzuez. Gaur egun gure datuak, irudiak eta ahotsak toki askotan daude. Kontuz ibili behar dugu datuak ematerakoan, baina para- noiko bihurtu gabe. ●



EROSI ZURE SARRERAK ONLINE

www.gustoko.eu

KOLABORATZAILE NAGUSIAK



ANTOLATZAILEA



EXPOSSIBLE!