VI. IKERGAZTE

NAZIOARTEKO IKERKETA EUSKARAZ

2025eko maiatzaren 28, 29 eta 30a
Bilbo, Euskal Herria

ANTOLATZAILEA:
Udako Euskal Unibertsitatea (UEU)

INGENIARITZA ETA ARKITEKTURA

Adimen Artifiziala Joko Patologikoaren Arriskua Aurresateko

Xabier Larrayoz, Arantza Casillas eta *Alicia Pérez*

13-18 or.
https://dx.doi.org/10.26876/ikergazte.vi.03.01

Adimen Artifiziala Joko Patologikoaren Arriskua Aurresateko

Xabier Larrayoz, Arantza Casillas, Alicia Pérez

HiTZ Center - Ixa, University of the Basque Country UPV/EHU 20080, Donostia, Spain.

xabier.larrayoz@ehu.eus

Laburpena

Lan honetan, ikasketa neuronal sakonean oinarritutako adimen artifizialeko teknikak proposatzen ditugu, sare sozialetan ludopatiaren arrisku-zantzuak goiz detektatzeko. Horretarako, erabiltzaileen mezuen sekuentziak aztertzen dira, testuetan latente dauden seinaleak identifikatzeko. Erronka nagusietako bat datuen klase-desoreka da, ludopatia duten erabiltzaileen multzoa kontrol multzoa baino 13 aldiz txikiagoa baita. Horri aurre egiteko, *Doitasuna* eta *Estaldura* orekatzeko lagin-aukeraketa estrategia bat garatu dugu. Gainera, denboran zehar mezuen bilakaera modelatzeko sistema bat diseinatu dugu, arrisku-zantzuak ahalik eta mezu gutxien bidez detektatzeko helburuarekin. Gure ekarpenak 'Hizkuntzaren Azterketa eta Prozesamendua' arloan kokatzen dira, medikuntzaren domeinurako aplikazioekin.

Hitz gakoak: Arrisku goiztiarraren iragarpena, Hizkuntzaren Azterketa eta Prozesamendua, Klaseen desoreka, Ikasketa sakona, Osasun mentala

Abstract

In this work, we propose artificial intelligence techniques based on deep neural learning to enable the early detection of gambling addiction risk indicators in social media. To achieve this, we analyze sequences of user posts to identify latent signals within the text. One of the main challenges is the class imbalance in the data, as the group of users with gambling addiction is 13 times smaller than the control group. To address this issue, we have developed a sample selection strategy aimed at balancing Precision and Recall. Additionally, we designed a system capable of modeling the evolution of risk indicators over time, with the goal of detecting them using as few messages as possible. Our contributions are framed within the field of 'Language Analysis and Processing,' with applications in the medical domain.

Keywords: *Early risk prediction, Natural Language Processing, Class imbalance, Deep learning, Mental health*

1 Sarrera eta motibazioa

Munduan zehar, gutxi gorabehera 450 milioi pertsonek mota desberdinetako osasun mentaleko nahasmenduak pairatzen dituzte, hau da, lau pertsonatik batek bere bizitzako uneren batean (Vos *et al.*, 2015; Sayers, 2001) pairatuko du. Nahasmendu hauen detekzio goiztiarra ezinbestekoa da pazienteen egoerak hobetzeko eta eguneroko bizitzan izan dezaketen eragina arintzeko (Keyes, 2007). Hala ere, diagnostiko tradizionalen metodoak, elkarrizketa klinikoetan edo galdetegiak betetzean oinarrituak, baliabide eta denbora aldetik garestiak dira (Lin *et al.*, 2017).

Gure egunerokoan sare sozialen presentzia gero eta handiagoa denez, erabiltzaileek beren pentsamenduak eta emozioak askatasunez adierazten dituzten plataforma garrantzitsu bihurtu dira. Fenomeno honek testu-datu ugari sortzen ditu, eta era egokian aztertuz gero, osasun mentalaren adierazle baliotsuei buruzko informazioa eskain dezake (Rahman *et al.*, 2018). Hala ere, sare sozialetako testuak prozesatzeak erronka bereziak dakartza, hala nola hizkuntza informalaren erabilera, akats ortografikoak, esaldi osatugabeak eta gramatika-arau koherenterik eza. Zailtasun horiek gorabehera, testu horiek informazio aberatsa eskaintzen dute, ohiko iturri tradizionaletan (adibidez, artikulu edo txosten klinikoetan) aurkitu ohi ez dena.

Biztanleriaren %26k jokoetan parte hartzen du (Thomas, 2022), eta Espainiako nerabeen erdiek dagoeneko probatu dute (Chóliz Montanés eta Marcos Moliner, 2021). Sare sozialetan oinarritutako osasun mentalaren inguruko ikerketak depresioaren eta suizidioaren detekzioan zentratu dira batez ere (Zhang *et al.*, 2022), baina premiazkoa

da beste arazo kritiko batzuk ere jorratzea, hala nola ludopatia.

Gure lana, ObserMenH[1] deritzon ikerketa proiektuaren testuinguruan kokatzen da Hizkuntzaren Azterketa eta Prozesamendua (HAP) teknologia aurreratuetan oinarritutako irtenbide berritzaileak eskainiz. ObserMenH proiektuaren helburua sare sozialetan aurkitutako arrisku psiko-linguistikoen zantzuak aztertzea da, gaztelanian arreta berezia jarriz (aurrekarien ikerketa ingeleserako garatzen baita gehien bat).

Artikulu honetan, ObserMenH proiektuaren baitan, joko patologiko goiztiarra detektatzeko proposatutako metodologia eta emaitza esperimentalak aurkeztu ditugu, ekarpen zientifikoak azpimarratu. Lan honekin, osasun mentalaren kudeaketan hizkuntzaren azterketa eta prozesamenduak aro digitalean izan dezakeen aplikazio potentziala agerian uzten da.

2 Arloko egoera eta ikerketaren helburuak

Lan honen markoan, aurreko ikerketek joko patologikoaren detekziorako erabilitako teknikak, hala nola analisi estatistikoa edo sare neuronal sinpleak, ez zuten denboran zeharreko portaeraren eboluzioa kontuan hartzen. Gure proposamena, aldiz, sekuentzia-denboralak modelatzen dituen FFNN batean oinarritzen da, arriskuaren bilakaera aztertzeko aukera emanez.

Lan honen helburua da sare sozialetan joko patologikoarekin lotutako arriskuak goiz detektatzeko metodo berritzailea garatzea, erabiltzaileen portaera linguistikoaren bilakaera denboran zehar kontuan hartuz. Arriskuzantzuak identifikatzeaz gain, helburua da hori ahalik eta azkarren egitea, erabaki fidagarriak hartzeko beharrezko mezu kopurua ahalik eta txikiena izanik.

Proposamenak sistema bat aurkezten du, eta sistema honek arrisku-faktoreak esleitzen dizkie denboran kronologikoki prozesatutako mezu bakoitzari. Hurbilketa horrek erabiltzailearen arriskuen bilakaera aztertzea ahalbidetzen du, erorketak edo portaera problematikoen jarraitutasuna adieraz ditzaketen ereduak detektatuz, argitalpen-sekuentzia osoa aztertu beharrik gabe. Ezaugarri hau funtsezkoa da detekzio goiztiarrerako, esku-hartzeko behar den denbora minimizatzen baitu.

Beste helburu garrantzitsu bat da datuetan dagoen desoreka berezkoa lantzea, izan ere, arrisku-portaerak erakusten dituzten erabiltzaileen proportzioa askoz txikiagoa da populazio orokorraren aldean. Horretarako, kostu-funtzio pertsonalizatu bat ezartzen da, kasu kritikoak lehenesten dituen eta datu desorekatuekin errendimendua hobetzen duen eredua garatzeko.

Gainera, lan honetan hizkuntza naturalaren prozesamendurako (*HAP*) teknika aurreratuak esploratu nahi dira, testuen eta joko patologikoarekin lotutako arriskuen arteko harreman konplexuak harrapatzeko. Sare neuronal sakonen bidez, informazio latentea ateratzen da, eta horrek ahalbidetzen du egileak sailkatzea (arrisku positiboa edo negatiboa) eta portaera linguistiko ebolutiboen detekzioa egitea.

Azkenik, lan honek *HAP* arloan bi aurrerapen gako ekartzen ditu:

- Erabiltzaileak denboran zehar monitorizatzeko gai den sistema baten inplementazioa, normalean sarrera independenteak aurkezten dituzten dataset-en mugak gaindituz (Komati, 2021; Namdari, 2022).
- Entrenamendu-esparru berritzaile bat, klaseen arteko desoreka erronka kritikoa den beste testuinguru batzuetara ere estrapola daitekeena.

Laburbilduz, lan honen helburuak osasun mentalaren arriskuen detekziorako tresna zehatzagoak eta eraginkorragoak garatzeko beharrarekin bat datoz, eta horrek esku-hartze goiztiarra ahalbidetzeko gaitasuna indartzen du, non erabaki bakoitzak eragin handia izan dezakeen.

3 Ikerketaren muina

Erabiltzaile batek l mezu jarraian (*posts*) idatzi dituela emanda, $\mathbf{t}_k = (t_k^1, t_k^2, \ldots, t_k^l)$ bezala adierazten dena, non t_k^j hurrenkerako j-garren mezua den, helburu nagusia erabiltzaile-mailako sailkapen-etiketa ($\widehat{u}_k$) bat estimatzea da. Etiketa honek jokalari patologikoak eta kontrol-erabiltzaileak bereizten ditu. Horretarako, t_k^j mezu bakoitza prozesatzen da, konfiantza-maila jakin batekin mezu-mailako etiketa (c_k^j) bat kalkulatzeko. Sortutako informazioa ondoren erabiltzailearen etiketa globala estimatzeko erabiltzen da.

[1] ObserMenH: `http://nlp.uned.es/obser-menh-project/index.html`

Ezarritako eredua sare neuronal *Feed Forward* (FFNN) batean oinarritzen da, eta PyTorch-en garatua dago (Paszke *et al.*, 2019). t_k^j mezu bakoitza aldez aurretik tamaina finkoko bektore numeriko batean bihurtzen da, $\mathbf{x}_k^j = (x_{k1}^j, x_{k2}^j, \ldots, x_{kN}^j) \in \mathbb{R}^N$, eta honek testuaren ezaugarriak jasotzen ditu, FFNN-rako sarrera gisa balioko duena. Kasu honetan, $t_k^j \in \Sigma^*$, Σ izanik zereginaren hiztegia. Sareak erabiltzen duen informazio guztia $\mathbf{x}_k^j$-n kodetuta dago, eta horrek errepresentazio hau ereduaren osagai kritikoa bihurtzen du.

Errepresentazio bektorial horiek sortzeko, bi *encoder* aztertu ditugu: SBERT (Reimers eta Gurevych, 2019) eta DAN (Cer *et al.*, 2018). SBERT-ek $N = 384$ tamainako bektoreak sortzen ditu, eta DAN-ek berriz $N = 512$. *Encoder*-aren hautaketak eragin zuzena du ereduaren errendimenduan eta testuen arteko harreman semantikoak harrapatzeko gaitasunean.

Erabiltzaile baten l mezuk osatutako sekuentzia emanda orden kronologikoan, $(t_k^1, t_k^2, \ldots, t_k^l)$, ereduak konfiantza-balioen sekuentzia bat kalkulatzen du, $\widehat{\mathbf{y}}_k = (y_k^1, y_k^2, \ldots, y_k^l) \in \mathbb{R}^l$. Bertan, y_k^j balioak, k erabiltzailearen j-garren mezuak (t_k^j) arriskuarekin lotuta egoteko probabilitatea irudikatzen du. Jarraian, $\widehat{\mathbf{y}}_k$-tik abiatuta, mezu-mailako etiketak sortzen dira, $\widehat{\mathbf{c}}_k$, hain zuzen ere, (1) adierazpenean deskribatutako eran. Hau da, y_k^j balioa positiboa bada, mezuari arrisku-etiketa esleitzen zaio ($\widehat{c}_k^j = 1$), eta bestela, kontrol-klasera sailkatzen da ($\widehat{c}_k^j = 0$). Metodo honek iragarpenak errazten ditu atalase zehatzik ezarri gabe, seinale sinbolikoa soilik erabiliz arriskuaren presentzia adierazteko.

$$\widehat{c}_k^j = \text{sign}(\widehat{y}_k^j) \tag{1}$$

Amaitzeko, mezu-mailan lortutako arrisku zantzuetatik abiatuta, erabiltzaile-mailan dagoen arriskua estimatzeko proposatu dugun prozedura definituko dugu. Doiketa-etapan, erabiltzaile-mailako sailkapena ($\widehat{u}_k$) mezu bakoitzaren etiketen arabera zehazten da (2) adierazpenari jarraituz. Beste hitzetan, doiketa etapan, mezu bat arriskutsutzat sailkatzen bada ($c_k^i = 1$), erabiltzailea patologikotzat jotzen da ($\widehat{u}_k = 1$).

$$\widehat{u}_k = \begin{cases} 1, & \text{baldin } \exists i \; 1 \leq i \leq l : c_k^i = 1 \\ 0, & \text{bestela} \end{cases} \tag{2}$$

Estimatutako konfiantza-balioen sekuentzia, $\widehat{\mathbf{y}}_k$, mezu-mailako etiketa errealekin alderatzen da, $\mathbf{y'}_k$. Konparaketa hori entropia gurutzatuaren kostu-funtzioaren bidez kuantifikatzen da, $H(\mathbf{y'}_k, \widehat{\mathbf{y}}_k)$, PyTorch-en ezarrita (Paszke *et al.*, 2019). Gainera, erabiltzailearen araberako haztapen-faktore bat gehitzen zaio, (3) adierazpenean azaltzen den moduan.

$$Loss_k = W_{user_k} \cdot H(\mathbf{y'}_k, \widehat{\mathbf{y}}_k) \tag{3}$$

Haztapen-faktoreak, W_{user_k} terminoak, sistemaren akatsak zigortzeko balio du. Hain zuzen ere, penalizazio-faktore honen bidez negatibo faltsuei eta positibo faltsuei haztapen-faktore ezberdina esleitu ahal zaie. Gure kasuan, patologiak dituzten erabiltzaileaz ez ohartzeko arriskua minimizatzearren, negatibo faltsuei haztapen faktore handiagoa ezarri diegu positibo faltsuei baino.Praktikan, W_{user_k} penalizazio balioak, sentikortasun-analisi baten bidez ezarri ziren, datu-desoreka baldintzetan ereduaren sendotasuna maximizatzeko irizpideaz.

4 Ondorioak

Lan honek agerian uzten du hizkuntza naturalaren prozesamenduaren (*HAP*) gaitasuna sare sozialetako testuetatik informazio baliotsua ateratzeko, bai modu esplizituan, bai modu inplizituan. Sare neuronal sakonak tresna eraginkorrak dira testuetako informazio semantiko latentea harrapatzeko eta informazio hori bektore numeriko gisa adierazteko. Kalkulu bektorial numerikoan datza sare neuronal sakonen potentziala, azterketa eraginkorrak eta zehatzak egitea ahalbidetzen baitira. Testuinguru honetan, garatutako ereduak erabiltzaileen ezaugarri sotilak azaleratu ditu, hala nola joko patologikoaren arriskua.

Atal esperimentala garatzeko, ikerketa alorrean erreferentziazko datu-sorta erabili da, hain zuzen ere, 2023-eRisk datu-sorta (Parapar *et al.*, 2023). Trebatze-multzoan 164 jokalari konpultsiboen (1. klasea) eta 2.184 kontrol-erabiltzaileen (0. klasea) argitalpenek osatu dute. Klaseen desoreka ikasketa automatikorako erronka nabarmena

da. Klase positiboa (1. klasea) lagin osoaren %7a baino ez da. Desoreka horrek azpimarratzen du eredua inplementatzeko laginaren aukeraketa egokia egin behar izango dela, laginean oinarritutako ikasketatik sortutako ereduaren errendimendu egokia bermatzeko.

Emaitza esperimentalei dagokienean, hizkuntza eredua handiak tartean direnean, konputazio errekurtso bereziak behar izaten dira, kasu honetan, ordea, eredu txikiak erabili ditugu eta entrenamenduak ez du baliabide handirik eskatu. Zehazki, Titan XP grafiko-txartel bakarrarekin lan egin dugu eta entrenamendu-prozesua pare bat ordutan burutu da. 1. taulan aurkezten dira bai oinarrizko kostu-funtzioarekin (B bidez adierazita) bai kostu-funtzio aldatuarekin (M bidez adierazita) lortutako emaitzak. B eta M kostu-funtzioek desberdintasun nabarmenak dituzte. Oinarrizko kostu-funtzioak (B) entropia gurutzatua erabiltzen du penalizazio berezirik gabe, eta aldatuak (M) berriz, (3) adierazpenean deskribatutako pisu-sistema ezartzen du, positibo faltsuak eta negatibo faltsuak desberdin penalizatuz arrisku-zantzuen detekzioan eraginkortasuna hobetzeko. Emaitzek adierazten dute (3) adierazpenean deskribatutako kostu-funtzioak eragin nabarmenak dakartzala.

1. taula aztertuta, ikus daiteke positibo faltsuak eta negatibo faltsuak desberdin penalizatzeak hobekuntza nabarmenak dakartzala ludopatia duten erabiltzaileen (1. klasea) detekzioan. Hau funtsezkoa da gure helbururako, 1. klaseak joko patologikoaren arriskua duten erabiltzaileak baititu eta detekzio goiztiarra erabakigarria baita eskuhartze azkarrerako. Adibidez, *F1-neurria* 71etik 78ra igo da eta *Doitasuna* 58tik 83ra. Hobekuntza hauek positibo faltsuen kontrol hobea erakusten dute, nahiz eta horrek 1. klaseko *Estaldura*-ren apur bat jaistea ekarri duen.

Garrantzitsua da nabarmentzea ebaluazio hauek 1.998 kontrol-klaseko (0. klasea) lagin eta 81 ludopatia klaseko (1. klasea) lagin erabiliz egin direla. Honen bidez azpimarratu nahi dugu proposatutako laginketa eraginkorra izan dela, maizen agertzen den klaserako alborapenari aurre egin zaiola proposatutako metodologiaz. Proposatutako hurbilpena eraginkor suertatu da desorekatutako klasea duen datuetatik ikasteko testuinguruan ezarri eta garaile atera baita. Nahiz eta 1. taulan kostu-funtzio aldatuarekin *Estaldura* apalagoa agertu, arriskuan dauden erabiltzaileen detekzio goiztiarra lehenesteko helburua bete da. Izan ere, erabiltzaile bakoitzeko sistemak, batezbestekoan, 8 mezu baino ez ditu aztertu behar izan. Honenbestez, aurrekariekiko iragarpen lehiakorrak eta azkarrenak lortu baitira. Detekzio goiztiarraren hobe beharrez.

1. Taula: Kostu-funtzioaren (Loss) eragina iragarpen gaitasunean. Oinarrizko kostu-funtzioa (B) eta funtzio aldatua (M) alderatzen dira. Iragarpen gaitasuna ebaluatzeko erabilitako metrikak: doitasuna (Pr), estaldura (R) eta F1-neurria (F1). Emaitzak klaseka xehetuta eman dira (0: kontrolekoa, 1:patologikoa) baita klaseen batezbesteko haztatua (W.Avg) eta makro-batezbestekoa (M.Avg) erabiliz.

Loss	Klasea	Pr	R	F1
B	0	100	97	98
	1	58	94	71
	M.Avg	79	96	85
	W.Avg	98	97	97
M	0	99	99	99
	1	83	73	78
	M.Avg	91	86	88
	W.Avg	98	98	98

Emaitza hauek azpimarratzen dute garatutako hurbilketaren potentziala osasun mentalaren azterketako arazo konplexuei aurre egiteko sare sozialetako datuen bidez. Konputazio metodologiei dagokionean, bereziki, desoreka baldintzetan sendotasuna agerian geratu da, eszenatoki praktikoetan, eraginkortasuna erakutsiz detekzio goiztiarrean. Detekzio goiztiarrerako gaitasuna ez da joan emandako doitasunaren kaltean, aitzitik, azpimarratzekoa da biak hobetu direla. Laburbilduz, ikasketarako laginak aukeratzeko egindako ekarpenaren bidez, ikasketa fasean alborapenen aurrean sistema sendoa lortu dugu, F1 metrikan oinarrituz zehaztasun altuko iragarpenak ematen dituena eta oso azkarra dena, erabiltzaileen mezu gutxi baino ez ditu analizatu behar alerta azaleratzeko.

5 Etorkizunerako planteatzen den norabidea

Garatutako sistemak potentzial handia du eta, zer esanik ez, hobetzeko abaguneak. Jarraian, ikerketarako norabide interesgarri batzuk azpimarratzen dira:

- **Denbora-dimentsioaren integrazioa:** Erabiltzaile-mailako etiketatze estrategia, (2) adierazpenean deskri-

batutakoa, eraginkorra eta konputazionalki ekonomikoa da klaseen desorekari aurre egiteko. Hala ere, interesgarria izango litzateke mezuek dituzten denbora-markak txertatzea, arriskuaren bilakaera denboran zehar zehatzago modelatzeko. Hurbilketa honek aukera emango luke erabiltzailearen portaeran denboran gertatzen diren aldaketekin lotutako eredu dinamikoak harrapatzeko.

- **Kostu-funtzioaren optimizazioa:** Kostu-funtzioak eta erabiltzen diren penalizazio-pisuek errendimendu sendoa erakutsi duten arren, haien azterketa sakontzea komenigarria litzateke, zehaztasunaren eta *estaldura*-ren arteko oreka hobea lortzeko. Bereziki, interesgarria izango litzateke pisu-faktoreak automatikoki doitzen dituzten metodoak esploratzea entrenamendu-prozesuan zehar, horrek ereduaren orokortze-gaitasuna hobetzeko datu-multzo ezberdinetan.
- **Datu-multzoaren hedapena:** Datu gehiago biltzea ala sortzea eredu generatiboak medio; bereziki, arriskuan dauden erabiltzaileen klasearentzako, funtsezkoa litzateke datuak izatea desoreka beste eratara kudeatzearren. Gainera, interesgarria izango litzateke beste iturrietako edo hizkuntzetako datuak txertatzea, ereduaren orokortzea testuinguru multikulturaletan ebaluatzeko.
- **Aplikazio praktikoak:** Praktikotasunari dagokionez, posiblea litzateke ereduaren iragarpenak osasun-profesionalentzat ulergarri bihurtzen dituzten tresna bisual interaktiboak garatzea. Tresna hauek arriskuaren bilakaera denboran irudikatzen duten grafikoak edo alerta automatikoak barne har ditzakete, ingurune klinikoetan integrazioa erraztuz.

Ondorioz, gure hurbilpenak sare sozialetako osasun mentalaren azterketan aurrerapen esanguratsua ekarri ez ezik, hizkuntzaren azterketa eta prozesamendua osasun digitalaren arloetako etorkizuneko ikerketetarako aplikazio posibleak ere agerian uzten ditu. Hala ere, horrelako sistemek etikoki arduratsua izan behar dute, diagnostiko okerrek edo sailkapen automatikoek ekar ditzaketen ondorioak kontuan hartuz. Beraz, beharrezkoa da pribatutasuna eta adierazpen-askatasuna bermatuko dituzten protokoloak ezartzea, baita detekzio goiztiarreko sistemek erabiltzaileengan izan dezaketen eragina aztertzea ere.

Erreferentziak

Cer, Daniel, Yinfei Yang, Sheng yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Yun-Hsuan Sung, Brian Strope, eta Ray Kurzweil, 2018. Universal sentence encoder.

Chóliz Montanés, Mariano, eta Marta Marcos Moliner. 2021. La adicción al juego de azar en la adolescencia: apuestas, tecnologías y consumo de drogas.

Keyes, Corey L. 2007. Promoting and protecting mental health as flourishing: a complementary strategy for improving national mental health. *Am Psychol* 62.95–108.

Komati, Nikhileswar, 2021. Suicide and depression detection. `https://www.kaggle.com/datasets/nikhileswarkomati/suicide-watch`.

Lin, Huijie, Jia Jia, Jiezhong Qiu, Yongfeng Zhang, Guangyao Shen, Lexing Xie, Jie Tang, Ling Feng, eta Tat-Seng Chua. 2017. Detecting stress based on social interactions in social networks. *IEEE Transactions on Knowledge and Data Engineering* PP.1–1.

Namdari, Reihaneh, 2022. Mental health corpus. `https://www.kaggle.com/datasets/reihanenamdari/mental-health-corpus`.

Parapar, Javier, Patricia Martín-Rodilla, David E. Losada, eta Fabio Crestani. 2023. Overview of erisk 2023: Early risk prediction on the internet. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction*, ed. by Avi Arampatzis, Evangelos Kanoulas, Theodora Tsikrika, Stefanos Vrochidis, Anastasia Giachanou, Dan Li, Mohammad Aliannejadi, Michalis Vlachos, Guglielmo Faggioli, eta Nicola Ferro, 294–315, Cham. Springer Nature Switzerland.

Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, eta Soumith Chintala. 2019. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, 8024–8035. Curran Associates, Inc.

Rahman, Rohizah Abd, Khairuddin Omar, Shahrul Azman Mohd Noah, eta Mohd Shahrul Nizam Mohd Danuri. 2018. A survey on mental health detection in online social network. *International Journal on Advanced Science, Engineering and Information Technology* 8.1431–1436.

Reimers, Nils, eta Iryna Gurevych, 2019. Sentence-bert: Sentence embeddings using siamese bert-networks.

Sayers, Janet. 2001. The world health report 2001 — mental health: new understanding, new hope. *Bull World Health Organ* 79.1085.

Thomas, Michelle. 2022. Gambling addiction statistics worldwide.

Vos, Theo, Ryan M Barber, Brad Bell, Amelia Bertozzi-Villa, Stan Biryukov, Ian Bolliger, Fiona Charlson, Adrian Davis, Louisa Degenhardt, Daniel Dicker, eta others. 2015. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the global burden of disease study 2013. *The Lancet* 386.743–800.

Zhang, Tianlin, Annika Schoene, Shaoxiong Ji, eta Sophia Ananiadou. 2022. Natural language processing applied to mental illness detection: a narrative review. *NPJ Digital Medicine* 5.

6 Eskerrak eta oharrak

Lan hau Espainiako Zientzia eta Berrikuntza Ministerioak finantzatu du partzialki (EDHIA PID2022-136522OB-C22) eta Eusko Jaurlaritzak (IXA IT-1570-22). Horrez gain, lan hau LOTU (TED2021-130398B-C22) proiektuaren esparruan garatu da, MCIN/AEI/10.13039/501100011033, Europako Batzordea (FEDER), eta Europar Batasunak "NextGenerationEU"/PRTR programaren bidez finantzatuta. Lehenengo egileak Espainiako Hezkuntza Ministerioaren *Formación de Profesorado Universitario* (FPU) 2024 programaren beka jaso du.