



## Olatz Arregi Uriarte

Donostiako Informatika Fakultateko irakasle eta ikertzailea, UPV/EHU

# ChatGPTren tripetan

**B**at baino gehiago liluratu da azken asteotan ChatGPTren emaitzak ikusita, baina ez dut hemen sistemaren emaitzei buruz hitz egin nahi, baizik eta haren barnean, haren tripetan, dagoenari buruz; hau da, elkarrizketa-sistemaz edo txatbot horrek horrelako erantzun onak emateko behar duenaz.

Lau gertaera daude arrakasta horren atzean:

**1. Testu eta dokumentu asko eta asko dago eskuragarri formatu elektronikoan.** ChatGPTren oinarrian hizkuntza-eredua dago; alegia, hitz baten ondoren, hurrengoa zein izango den “iragartzen” duen sistema. Horretarako, testu askorekin entrenatzen da sistema, eta hortik sortutako eredua gai da hizkuntza naturalean esango genukeena imitatzeke.

ChatGPTren oinarrian dagoen hizkuntza-eredua (GPT-3) entrenatzeko wikipedia osoa, liburu elektronikoak, webguneak... 1.000 milioi hitz baino gehiago erabili dira, guztira, eta, horri esker, kalitate oneko erantzunak ematen ditu txatbotak hizkuntza nagusietan, horietatik ikasi baitu.

**2. Eskulan handia egin da hizkuntza-eredua aberastan.** ChatGPT ez da hizkuntza-eredua, elkarrizketa-sistema baizik; hau da, erabiltzaileari argibideak eta erantzunak modu lagungarrian ematea helburu duen sistema. Horretarako, *Reinforcement Learning from Human Feedback (RLHF)* izeneko teknika erabili da; alegia, hizkuntza-eredua eskuz aberastu da elkarrizketak bideratzeko gaitasuna izan dezan. Helburua ez da sistemari galdera guztien erantzunak ematea, baizik eta, batzuk emanda, orokortzen eta antzeko galderei erantzunak ematen erakustea. ChatGPTren kasuan, eskuzko parte-hartze zuzena hiru fasetan egin zen:

1. Hasieran pertsona talde batek galderak eta erantzun egokiak (txatbota balira bezala) sartu zizkion sistemari.
2. Gero, txatak ematen zituen erantzun posibleak ordenatzen eta sailkatzen zituzten eskuz, erantzunen egokitasuna

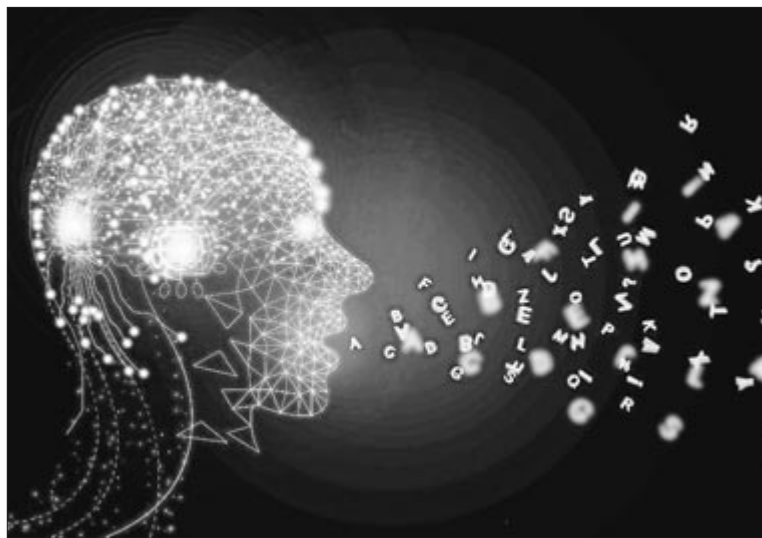
eta kalitatea kontuan izanik, eta berriz entrenatu zuten sistema.

3. Azkenik, sistemaren emaitza desagokiak alboratzeko, horiek eskuz etiketatu ziren. Erantzun desagoki gisa honako hauek hartzen ziren: gorroto-mezuak, arrazistak, sexistak edo abusuak bultzatzen zituztenak. Prozesu horren ondorioz, mezu toxikoak antzemateko sistema sortu eta elkarrizketa-sisteman integratu zuten.

**3. Ahalmen handiko konputagailuak, kalkulu-abiadura izugarria dutenak, erabili dira sistema entrenatzeko.** Hainbeste informazio duen sistema batek ondo erantzuten ikas dezan, konputagailu sendoak behar dira. Adibidez, ChatGPT entrenatzeko, OpenAI enpresak Microsoft-i eskatu zion superkonputagailu berezi bat egitea. Konputagailu horrek

285.000 prozesadore “arrunt” baino gehiago eta 10.000 GPU ditu. GPUak datu-multzo handiekin paraleloan lan egiteko diseinaturik dauden prozesadore espezializatuak dira.

**4. Teknika eta paradigma berriak erabili dira entrenamendurako.** Neurona-sareak duela urte batzuk hasi ziren arrakastatsu bihurtzen, eta gaur egun adimen artifizialeko aplikazio gehienetan erabiltzen dira. Gizakion garunaren egituraren eta funtzionamenduan oinarritutako eredu matemati-



Txatbotak badu zer hobetu euskararen eta antzeko hizkuntza txikien kasuan.

koa da neurona-sarea; hau da, ikasketa automatikoko teknika ahaltsua. ChatGPT sortzeko InstructGPT hizkuntza-eredua hobetuaz baliatu dira, eta eredu hori 175.000 milioi parametro erabiltzen dituen neurona-sare batekin entrenatu dute.

Txatbotak asko dauka hobetzeko oraindik, eta horretan ari dira buru-belarri. Egun hauetan, esaterako, GPT-4 hizkuntza-eredua iragarri dute, zeinak, testuez gain, ahotsa, irudia edo bideoak txertatuko dituen. Baina beste arlo bati erreparatuta, are eta berdeago daude gauzak. Esate baterako, zein leku dute sistema hauetan nagusiak ez diren edo baliabide urri dituzten euskara bezalako hizkuntzek? Jarraituko du... •