

HIZKUNTZA-ULERMENARI EKARPENAK: N-GRAMEN ARTEKO ATENTZIO ETA LERROKATZEAK ANTZEKOTASUN ETA INFERENTZIA INTERPRETAGARRIRAKO

Tesiaren egilea: Iñigo Lopez-Gazpio

Unibertsitatea: Euskal Herriko Unibertsitatea (UPV/EHU)

Saila: Lengoaia eta Sistema Informatikoak saila (LSI)

Tesi-zuzendaria: Eneko Agirre eta Montse Maritxalar

Tesiaren laburpena:

Hizkuntzaren Prozesamenduaren bitartez hezkuntzaren alorreko sistema adimendunak hobetzea posible da, ikasleen eta irakasleen lan-karga nabarmenki arinduz. Tesi honetan esaldi-mailako hizkuntza-ulermena aztertu eta proposamen berrien bitartez sistema adimendunen hizkuntza-ulermena areagotzen dugu, sistemai erabiltzailearen esaldiak modu zehatzagoan interpretatzeko gaitasuna emanez, esaldiak modu finean interpretatzeko gaitasunak feedbacka modu automatikoan sortzeko aukera ematen baitu.

Tesi hau garatzeko hizkuntza-ulermenean sakondu dugu, antzekotasun semantikoari eta inferentzia logikoari dagokien ezaugarriak eta sistemak aztertuz. Berezi, esaldi barneko hitzak multzotan egituratuz eta lerrokatuz, esaldiak hobeto modelatu daitezkeela erakutsi dugu. Horretarako, hitz solteak lerrokatzen dituen aurrekariaren egoerako neurona-sare sistema bat inplementatu eta n-grama arbitrarioak lerrokatzeko moldaketak egin ditugu. Hitzen arteko lerrokatzea aspalditik ezaguna bada ere, tesi honek, lehen aldiz, n-grama arbitrarioak atentzio-mekanismo baten bitartez lerrokatzeko proposamenak plazaratzen ditu.

Gainera, esaldien arteko antzekotasunak eta desberdintasunak modu zehatzean identifikatzeko, esaldien interpretagarritasuna areagotzeko eta ikasleei feedback zehatza emateko geruza berri bat sortu dugu: iSTS. Antzekotasun semantikoa eta inferentzia logikoa biltzen dituen geruza horrekin chunkak lerrokatu ditugu, eta ikasleei feedback zehatza emateko gai izan garela frogatu dugu hezkuntzaren testuinguruko bi ebaluazio-eszenarioran.

Tesi honekin batera, hainbat sistema eta datu-multzo argitaratu dira, etorkizunean komunitate zientifikoak ikertzen jarrai dezan.