

Dirbtinio intelekto (DI), biometrinių technologijų ir natūralios kalbos apdorojimo sprendimų vystytoja „Neurotechnology“ išleido balso sintezės modelį, kuris tekstą paverčia garso įrašu lietuvių kalba.

Naujasis sprendimas leidžia bet kokią įvestą tekstą paversti realistišku garso įrašu – su natūraliomis intonacijomis, pauzėmis ir sklandžiu žodžių jungimu. Priešingai nei daugelis rinkoje esančių įrankių, kurie tiesiog atkuria iš anksto įrašytus balsus, ši technologija geba atpažinti ir taisyklingai tarti įvairias žodžių formas realiuoju laiku, prisitaikydama prie sudėtingos lietuvių kalbos fonetikos. Tai – reikšmingas žingsnis į priekį, suteikiantis plačiai pritaikomą kalbinę naująją Lietuvos rinkai.

Technologiją jau dabar galima išbandyti viešai: www.netgeist.ai/lt/tts. Demonstruojamoje versijoje yra 4 moteriški ir 3 vyriški balsai. Naudotojai gali pasirinkti bet kurį balsą, įterpti tekstą ir vos per kelias sekundes išgirsti skaitmenizuotą skaitymą lietuviškai.

Reaguoja į kalbos iššūkius

Pasak kalbos technologijų specialistų, balso sintezė tampa ypač svarbi kontekste, kur didžioji dalis turinio yra anglų kalba. Tai didina riziką, kad mažiau vartojamos kalbos skaitmeninėje erdvėje taps antraeilėmis.

„Lietuvių kalba pasižymi sudėtinga fonetine ir morfologine sandara – kintančiais kirčiais, gausiais garsiniais junginiais bei išplėtotą linksnių ir laikų sistema. Mūsų tikslas buvo sukurti modelį, kuris ne tik atkurtų balsą, bet ir tiksliai perteiktų prozodiją: kada turi būti daroma pauzė, kur krinta kirtis ir kaip tai lemia sakinio intonaciją“, – sako Rokas Šidlauskas, „Neurotechnology“ natūralios kalbos apdorojimo inžinierius.

Anot jo, dauguma pasaulinių sintezės sistemų lietuvių kalbą apdoroja kaip „egzotinę“ – t. y., su menku mokymo duomenų kiekiu. „Tai reiškia, kad vartotojas gali girdėti netaisyklingus kirčius, per ilgas pauzes ar net sulietus žodžius. Mūsų modelis mokytas specialiai lietuviškai, todėl geba taisyklingai interpretuoti net ir sudėtingas sakinio konstrukcijas bei gali būti pritaikytas skirtingoms kalbinėms situacijoms – nuo žinių skaitymo tono iki rišlaus pokalbio“, – priduria R. Šidlauskas.

Pritaikymas: nuo kasdienybės iki prieinamumo

Technologija sukurta taip, kad būtų prieinama plačiam naudotojų ratui: jai nereikalinga speciali įranga, ji veikia tiek debesijoje, tiek lokaliuose sistemose. Todėl šį sprendimą gali

lengvai integruoti tiek startuoliai ar verslai, tiek bibliotekos, mokyklos ar viešosios įstaigos.

Viena svarbiausių galimų technologijos taikymo krypčių – socialinis prieinamumas. Pavyzdžiui, regos negalią turintiems asmenims balso sintezė leidžia klausytis informacijos, el. knygų, naujienų ar atlikti užklausas balsu. Taip pat tai gali padėti žmonėms, turintiems mokymosi ar rašytinės kalbos suvokimo sunkumų.

Kartu – tai sprendimas automatizuotai komunikacijai verslo aplinkoje. Skambučių centrai, savitarnos sistemos, tekstų įgarsinimas žiniasklaidoje ar e. paslaugose – tai tik keletas sričių, kuriose šis modelis gali veikti visą parą, nepriklausomai nuo darbuotojų skaičiaus.

Apie įmonę:

„Neurotechnology“ yra didelio tikslumo algoritmų ir programinės įrangos, pagrįstos giliaisiais neuroniniais tinklais ir kitomis dirbtinio intelekto technologijomis, kūrėja. 1990 m. Vilniuje įsteigta bendrovė siūlo dirbtiniu intelektu paremtus sprendimus įvairiose srityse, įskaitant biometriją, natūralios kalbos apdorojimą, kompiuterinę regą, smegenų ir kompiuterio sąsają, taip pat ultragarso technologijas. Bendrovės kuriami NKA sprendimai skirti viešajam ir privačiam sektoriams, sukuriant naujas galimybes automatizuoti kalbos apdorojimo užduotis.

Įmonė pelnė „2024 Metų inovacijos“ apdovanojimą už didžiojo kalbos modelio lietuvių kalbai sukūrimą, o 2023-ųjų pabaigoje buvo tituluota „Lietuvos metų dirbtinio intelekto (AI) įmone“.

