

ІНСТРУМЕНТЫ ШТУЧНАГА ІНТЭЛЕКТУ ДЛЯ АГУЧКІ ТЭКСТАЎ

Калеснік Н.В.

*Беларускі дзяржаўны ўніверсітэт інфарматыкі і радыёэлектронікі
г. Мінск, Рэспубліка Беларусь*

Пятрова Н.Я. – к.філал.н.,дацэнт

Артыкул прысвечаны разгляду розных тэхналогій сінтэзу маўлення, такіх як нейронныя сеткі і алгарытмы машыннага навучання. Вызначаюцца прыклады выкарыстання інструментаў штучнага інтэлекту для агучкі тэксту ў розных галінах, звяртаецца ўвага на актуальныя платформы і дадаткі для агучкі тэкстаў.

У сучасным грамадстве інтэрнэт паспеў укараніцца ў многіх галінах: адукацыя, праца, быт, забавы і інш. Рэалізацыя сябе чалавекам у гэтых сферах стала магчыма побач са штодзённым ростам колькасці разнастайных сервісаў, але не заўсёды магчыма знайсці патрэбны сайт сярод тысяч падобных. Праблему вырашаюць нейрасеткі – ім трэба толькі адправіць запыт, і яны згенеруюць прыдатныя для вас рэсурсы, крыніцы інфармацыі, навіны і інш. Увогуле, штучны інтэлект значна эканоміць час і павялічыць прадуктыўнасць працы. Паводле стандартнага вызначэння, штучны інтэлект (ШІ) – гэта тэорыя і распрацоўка камп'ютарных сістэм, здольных выконваць задачы, якія патрабуюць чалавечага розуму, такія як візуальнае ўспрыманне, распазнаванне мовы, прыняцце рашэнняў, пераклад і т. п. [1].

Развіццё і распаўсюджванне сучасных лічбавых тэхналогій ва ўсіх галінах чалавечага жыцця прыводзіць да пастаяннага пошуку аптымальных форм узаемадзеяння паміж чалавекам і камп'ютарам. У 1990-х гадах былі зроблены першыя спробы масавага ўкаранення ШІ ў сферу адукацыі або адпачынку. Гэта дапамагло людзям пазнаёміцца з рознымі тыпамі прадуктаў на аснове яго тэхналогій: у выглядзе тамагочы, iPod Touch, чат-ботаў і інш. Такія кампаніі, як Mattel, ствараюць асартымент цацак з падтрымкай ШІ для дзяцей ва ўзросце ад трох гадоў.

Выкарыстоўваючы запатэнтаваныя сістэмы і сродкі распазнання мовы, яны могуць разумець чалавечае маўленне, даваць інтэлектуальныя адказы і хутка вучыцца.

ШІ таксама актыўна выкарыстоўваецца ў індустрыі гульні. Напрыклад, у відэагульнях пашыраны боты, якія прызначаны для таго, каб гуляць ролю праціўнікаў, дзе людзі недаступны або непажаданы. У 2018 годзе даследчыкі з Карнельскага ўніверсітэта ўкаранілі нейронныя сеткі і навучылі іх ствараць новыя ўзроўні гульні Doom. Людзі атрымалі магчымасць генераваць новыя ўзроўні гульні без дапамогі чалавека [2]. Гэта гаворыць аб тым, што ў апошнія гады ШІ стаў важным інструментам у розных галінах навукі і тэхнікі, і мовазнаўства не стала выключэннем.

Развіццё тэхналогій апрацоўкі натуральнай мовы дазволіла ШІ аналізаваць і мадэляваць моўныя структуры з вялікай дакладнасцю. У сваю чаргу, аналіз моўных структур з'яўляецца адной з асноўных задач у лінгвістыцы, а таксама ў дадатках ШІ для апрацоўкі натуральнай мовы. Менавіта таму ў цэнтры ўвагі лінгвістыкі аказваюцца спосабы адаптацыі мовы для кантролю і кіравання самымі разнастайнымі прыладамі, якія спрашчаюць дзейнасць чалавека. Тэхналогіі сінтэзу і распазнання мовы шырока ўжываюць у самых разнастайных сферах жыцця – медыцыне, ваеннай, транспартнай, у тэлекамунацыях, сферы забаў і ў працэсе навучання. Гэты працэс уключае ў сябе дэталёвае даследаванне слоў, сказаў і тэксту ў цэлым, каб зразумець, як яны функцыянуюць і як перадаюць сэнс[3]. ШІ мае важную ролю ў аўтаматызацыі і ўдасканаленні гэтага аналізу, дазваляючы даследчыкам і распрацоўшчыкам апрацоўваць велізарныя аб'ёмы даных з высокай дакладнасцю і хуткасцю. Раней для гэтага патрабаваўся ўдзел чалавека. Цяпер неабавязкова шукаць прафесійнага дыктара або запісваць свой голас, калі трэба нешта агучыць. На фоне развіцця Text-to-Speech тэхналогій агульнадаступнымі сталі сінтэзатары гаворкі[4].

Сінтэзатары гаворкі могуць з'яўляцца эфектыўным сродкам навучання, які можна выкарыстоўваць пры вывучэнні замежных моў, а таксама могуць прысутнічаць у вытворчасці аўдыёкніг і рэкламных ролікаў. Інфакіёскі, галасавыя і сігналізацыйныя сістэмы апавяшчэння, камп'ютары, якія гавораць для інвалідаў па зроку і слыху, навучальныя сістэмы – усё гэта можа стварацца з дапамогай сінтэзатараў гаворкі. Увогуле, агучка тэкстаў не толькі паляпшае даступнасць інфармацыі, але і робіць яе больш зручнай і прывабнай для людзей. Прычым сучасныя інструменты дазваляюць гэта зрабіць так, каб гаворка ўспрымалася рознымі галасамі[5]. Для гэтага можна выкарыстоўваць такія сервісы, як Oddcast, YandexSpeechKit, Steosvoice, Robivox і інш. Многія з іх маюць выбар розных моў, галасоў, хуткасці чытання, даўжыні паўз. Часам існуе опцыя па стварэнні пэўнага настрою, эмацыйнай афарбоўкі, напрыклад, радасць, злосць, разгубленасць і г. д. Наяўнасць гэтых і многіх іншых характарыстык дапамагае меркаваць аб узроўні якасці сервісу і выбраць самы аптымальны. Параўнаем некаторыя з іх.

Разгледзім наступную табліцу:

Табліца 1 – Нагляднае параўнанне сервісаў для агучкі тэкстаў.

Сервіс	Text-to-Speech	Oddcast	YandexSpeechKit	Steosvoice	NaturalReader	Robivox
Тып	Анлайн-сервіс	Анлайн-сервіс	Анлайн-сервіс	Сервіс-бот у Telegram.	Анлайн-сервіс, мабільны дадатак	Анлайн-сервіс
Запампоўкаў аўдыёфармаце	Даступна	Не даступна	Не даступна	Даступна	Даступна ў платнай падпісцы	Даступна
Колькасць даступных моў	15	32	15	20	100	100
Натуральнасць маўлення	Галасы гучаць рабатызавана, ненатуральна, сустракаецца няправільная пастаноўка націску.	Гаворка пазбаўлена інтанацыі, няправільнае вымаўленне слоў сустракаецца нячаста.	Згенераваная гаворка мае натуральнае і інтанацыйнае гучанне.	Гаворка максімальна набліжана да чалавечай.	Мноства інструментаў для акцэнтаў, выражэння эмоцый, перадачы голасам узросту дыктара.	Шмат даступных фарматаў, галасы не пазбаўлены інтанацыі.

Разнастайнась галасоў	Ёсць як мінімум 1 жаночы голас. Радзей сустракаецца мужчынскі, часцей пара з дзвюх жаночых.	Галасы дарослых людзей (некалькі варыянтаў жаночых і мужчынскіх амаль на кожную мову), дзіцячыя і нават нестандартныя галасы	Больш за 10	Больш за 800	50	10
Максімальны аб'ём знакаў у дзень	500	600	500	2000	2000	100

Усе разгледжаныя сервісы даюць магчымасць карыстацца імі бясплатна, але не ва ўсіх іх ёсць пашырэнныя настройкі. Напрыклад, сервіс Text-to-Speech хоць і выконвае асноўную задачу – агучка тэксту, але свабоды выбару інтанацыі і настрою не падае, у той час як сервіс Robivox насычаны рознымі настройкамі і мае ў сваёй бібліятэцы большую колькасць галасоў у параўнанні з іншымі сервісамі. Але гэты прадукт абмяжоўвае колькасць даступных знакаў для агучкі: усяго 100 за раз, калі іншыя сервісы прапануюць іншыя значэнні.

Мы лічым, што адным з лепшых для агучкі тэкстаў з'яўляецца сервіс NaturalReader, бо яго прымушае вылучацца сярод канкурэнтаў не толькі аб'ёмная бібліятэка галасоў, якія можна рэдагаваць па сваім жаданні, але і наяўнасць зручнага мабільнага дадатку, што шмат у чым спрашчае працу з сервісам. Адзначым, што сервіс Steosvoice таксама ўяўляецца нам вельмі прыдатным. Ён працуе на аснове месенджара Telegram, таксама мае ў сваім запасе мноства моў, прасты інтэрфейс, што значна палягчае карыстанне гэтым сервісам.

Каля паловы з разгледжаных намі сервісаў пашыраны на англійскай мове, і, нягледзячы на велізарную разнастайнасць выбару ўбудаваных перакладчыкаў для нашых браўзераў, усё яшчэ застаецца цяжка арыентавацца ў свеце, дзе дамінуе англійская мова, людзям, у якіх англійская не з'яўляецца роднай мовай. Напісанне патрабуе высокага навыку ў пытаннях граматыкі, структуравання сказаў, выбару слоў і пошуку інфармацыі. З-за непрафесійных метадаў напісання і нізкай хуткасці пісьменнікі звяртаюцца па дапамогу да чат-ботаў. Хоць яны нядрэнна спраўляюцца са сваёй задачай, ім часта не хапае яснасці і чалавечага інтэлекту. Менавіта з такой мэтай пісьменнікі могуць гуманізаваць тэксты ШІ для атрымання адноснага і натуральнага зместу. Напрыклад, сервіс CudekAI хутка і бясплатна ачалавечвае тэксты для аўтараў, чыя англійская мова яшчэ не цалкам развіта. Гэты і іншыя падобныя сервісы з'яўляюцца эфектыўнымі памочнікамі ў напісанні ў розных сектарах, дапамагаюць карыстальнікам удаканалваць памылкі, каб падтрымліваць стандарт дакладнасці, асабліва на роднай мове.

Такім чынам, агучка тэкстаў з дапамогай штучнага інтэлекту з'яўляецца важным напрамкам у развіцці не толькі нейронных сетак, але і мноства іншых сфер. Рост даступнасці бясплатных платформ для агучкі тэксту дапамагае ў распаўсюджванні ўнікальнага кантэнту, спрашчае працэс навучання, ствараючы новыя інтэрактыўныя прыёмы, і спрыяе атрымання інфармацыі ў аўдыёфармаце людзям з абмежаванымі магчымасцямі.

Спіс выкарыстаных крыніц:

1. Морхат, П. М. К вопросу об определении понятия искусственного интеллекта / П. М. Морхат // *Право и государство: теория и практика*. – 2017. – №12. – С. 25–32.
2. Искусственный интеллект [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/%D0%98%D1%81%D0%BA%D1%83%D1%81%D1%82%D0%B2%D0%B5%D0%BD%D0%BD%D1%8B%D0%B9_%D0%B8%D0%BD%D1%82%D0%B5%D0%BB%D0%BB%D0%B5%D0%BA%D1%82. – Дата доступа: 11.03.2025.
3. Левчук, А. Штучний інтелект: лінгвістичні аспекти / А. Левчук // *Науковий вісник Східноєвропейського національного університету імені Лесі Українки*. – 2015. – С. 205–209.
4. ИИ и Языкознание: как искусственный интеллект изменяет изучение языков и перевод [Электронный ресурс]. – Режим доступа: <https://dzen.ru/a/ZlcCg00nSm24QU3d>. – Дата доступа: 11.03.2025.
5. Пятрова, Н. Я. Штучны інтэлект у інклюзіўнай адукацыі (на прыкладзе выкладання дысцыпліны «Беларуская мова» у ВНУ) / Н. Я. Пятрова // *Непрерывное профессиональное образование лиц с особыми потребностями: сборник статей V Международной научно-практической конференции, Минск, 14 декабря 2023 / Белорусский государственный университет информатики и радиоэлектроники; редкол.: А. А. Охрименко [и др.]. – Минск, 2023. – С. 239–243.*