

**Н. А. ТОЙГАНБАЕВА^{1*}, Ж. Х. ЖУНУСОВА¹, А. ПРОВАТА²,
Г. А. ТЮЛЕПБЕРДИНОВА¹**

¹Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ., Қазақстан

²Наногылым және Нанотехнология институты, Афина қ., Греция

E-mail: bodinaz@mail.ru, tyulepberdinova@gmail.com,

a.provata@inn.demokritos.gr, zhunussova777@gmail.com

ТЕРЕҢ ОҚЫТУ МОДЕЛЬДЕРІ НЕГІЗІНДЕ ҚАЗАҚ-ОРЫС ТІЛІНДЕГІ ОФФЛАЙН ҚОЛЖАЗБА МӘТІНДЕРІН ТАҢУ

Қолжазба жазбаларының цифрланған мәтіні көптеген компаниялардың бизнес-процестерін автоматтандыруға мүмкіндік береді және адамның жұмысын жеңілдетеді. Мақалада тереңдетілген оқыту модельдерін қолдана отырып, орыс және қазақ тілдеріндегі жазбаша қолжазбаны таңу қарастырылады. Әр адамның қолжазбасы ерекше және қолжазба мәтінінің жалпы үлгілерін жасау мүмкіндігі жоқ болғандықтан мәтінді офлайн таңу есебі онлайн таңу есебіне қарағанда күрделі болып саналады. Жұмыста кириллица графикасына қатысты қазақ-орыс тілдеріндегі қолжазбаны таңу модельдерін жасауда терең оқытудың әртүрлі тәсілдері қолданылады. Терең оқытудың сәтті маңызды элементі – бұл нейрондық желі модельдерін құруды және оларды шығару үшін орындауды жеңілдететін деректердің, есептеулердің, бағдарламалық платформалардың және жұмыс уақытының қол жетімділігі. Оқытудың әртүрлі тәсілдері мен әдістерін талдау мен салыстырудың жылдам, сандық және әділ құралдарын ұсынатын, бәсекеге қабілетті, қарқынды дамып келе жатқан сала терең оқыту ұғымдары талқыланды. Abdallah, Bluche, Flor және PUIGCERVER сияқты танымал терең оқыту модельдеріне шолу жасалды және эксперименттер нәтижелері гистограмма негізінде талданды. Эксперименттер Kazakh Offline Handwritten Text Dataset (КОНТД) деп аталатын қазақ тіліндегі оффлайн қолжазба мәтіндерінің үлкен деректер қоры негізінде жасалды. Қазақ тілінде қолжазба деректерді жинау үшін арнайы құрылған телеграм бот құрылды. Осы телеграм-бот арқылы енгізілген ақпаратын тексеріп, жауаптарды түзету құралы нейрондық желі негізінде жасалды.

Түйін сөздер: қолжазба, терең оқыту, нейронды желі, КОНТД, деректер жиыны.

Кіріспе. Қолмен жазылған мәтінді таңу (ағылш. Handwriting recognition, HWR немесе Handwritten Text Recognition, HTR) – бұл қағаз құжаттар, суреттер және сенсорлық экрандар және басқа құралдар көздерден түсінікті қолжазба ақпаратын алу және таңу қабілеті. Қолмен жазылған мәтінді оффлайн таңу – кірісі екі өлшемді кескін, шығысы таңбалар тізбегі болып табылатын кескіндегі әріптерді немесе сөздерді сандық мәтінге түрлендіру әдісі. Қазіргі уақытта қол жетімді қолжазбаны таңу бағдарламалық шешімдерін онлайн және оффлайн таңу деп екі негізгі бөлікке бөлуге болады: Онлайн таңу әртүрлі құрылғылардың экранында арнайы құрылғылармен (стилустармен) жазылған мәтінді таңуды қамтиды. Онлайн таңу мәселесі Microsoft, Evernote, Google және т.б. сияқты көптеген өнімдерде сәтті шешілді және Windows Vista, Windows 7 және Windows Mobile операциялық жүйелеріне енгізілген. Заманауи компоненттер мен икемді құрылыммен үйлестірілген Google-дің қолжазбаны таңу жүйесі 97 тілді қолдайды және Android құрылғыларына енгізу әдісі ретінде

* E-mail корреспондирующего автора: bodinaz@mail.ru

қол жетімді [1]. Оффлайн қолжазбаны тану суреттегі мәтінді компьютерлік және мәтінді өңдеу қосымшаларында қолдануға болатын әріптік кодтарға автоматты түрде түрлендіруді қамтиды. Өз кезегінде, қолжазба мәтінін оффлайн тану мәселесі әлі толық шешілген жоқ. Соңғы 10 жылда орын алған баспа және қолжазба мәтіндерін тану саласындағы ілгерілеу байқалады. Есептеулердің күрделілігіне қарамастан, кейбір өте күрделі есептеулерді машиналық оқыту және терең нейрондық желілер үшін жасалған құрылымдар арқылы шешуге болады. Кириллица негіздегі қазақ-орыс тіліндегі қолжазбаларды тану мәселесі әлі де толық зерттелмеген. Сондықтан, қазақ тіліндегі оффлайн қолжазба мәтіндерін тану алгоритмдерін жасау өте маңызды және осы салада зерттеулер жүргізу қажет, яғни мақала тақырыбы өзекті.

Жұмыстың мақсаты орыс және қазақ тілдеріндегі қолжазбаларды тану және өңдеу міндетін шешу үшін терең оқытудың заманауи әдістерін пайдалана отырып бағдарламалық қамтамасыз етуді әзірлеу болып табылады. Терең және машиналық оқыту арқылы қолжазба мәтінін танумен айналысатын зерттеушілердің эксперименттеріне негіз ретінде қолжазба деректер жиыны (датасет) қажет. Ағылшын тілінде IAM, EMNIST [2,3], қытай тілінде HIT-MW [4], испан тілінде RODRIGO [5], орыс тілінде НКР [6] қолжазба деректер жиыны бар. Жақында қазақ тіліндегі оффлайн қолжазба мәтіндерінің деректер қоры Kazakh Offline Handwritten Text Dataset (КОНТД) құрылды [7], құрамында шамамен 922010 таңба және 140335 сегменттелген кескін бар. Осы КОНТД қолжазба деректер жиыны негізінде нейрондық желісінің архитектурасын пайдаланып, қазақ тілінде қолжазбаны тану қарастырылады.

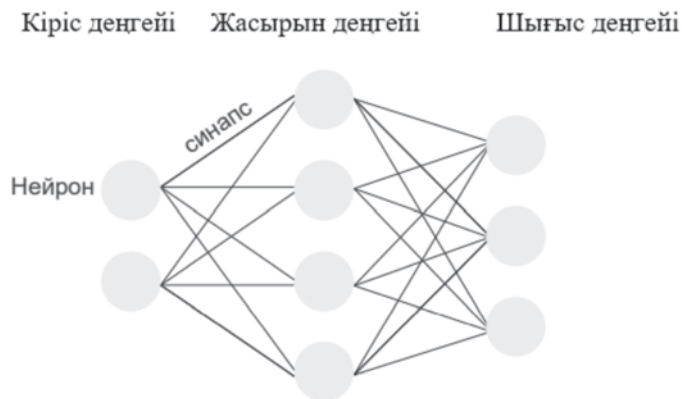
Зерттеу әдістері. Терең конволюциялық нейрондық желілердің өркендеуімен қолжазбаны тану мәселесінің жаңа шешімдері ұсынылды. Терең нейрондық желілер тілдік таңбалардың/әріптердің белгілі бір ерекшеліктерін автоматты түрде зерттей алады, осылайша қолжазбаларды тану дәлдігін жоғары деңгейге көтеруге болады.

Терең оқыту – бұл «терең» нейрондық желілерді қолданатын машиналық оқыту әдістерінің жалпы термині. Бүгінгі таңда терең оқыту машиналық оқытудың ең көрнекті бағыттарының бірі болып табылуына компьютерлік көру, тілді өңдеу, ойын ойнау, шешім қабылдау және модельдеу сияқты салалардағы жетістіктері дәлел болады. Жасанды интеллекттің бір саласы – машиналық оқыту ІТ саласында маңызды орын алады. Біз күнделікті өмірде пайдаланатын сандық деректердің көлемі күн сайын артып келеді, осыған байланысты технологиялық прогресті одан әрі дамыту үшін смарт-деректерді автоматты талдау қажеттілігі туындады. Машиналық оқытудың белсенді дамуы жасанды нейрондық желілердің кең қолданысқа ие болуына әкелді. Адам миының биологиялық құрылымына негізделген нейронды желілер өзінің жоғарғы деңгейде есептеу мүмкіндігімен басқа оқыту алгоритмдерінен бірнеше есе тиімді болып есептеледі [8].

Нейронды желілерді қолданудың басты артықшылығы көлемді немесе айқын емес мәліметтерден ақпараттарды терең оқыту әдістері арқылы алу. Оқытылған нейронды желі дайындалған деректер санаты үшін «сарапшы» рөлін атқара алады. Жасанды нейрондық желілердің ең жарқын және маңызды түрлерінің бірі – үйірткілі (сверточная) нейронды желі (Convolution Neural Networks CNN). Үйірткілі нейронды желілерді қолданатын алгоритмдер әртүрлі бұрмалауға, атап айтсақ, камераны бұру, суреттегі жарықтың біркелкі болмауы, көлденең немесе тік ығысулар және т.б. инвариантты

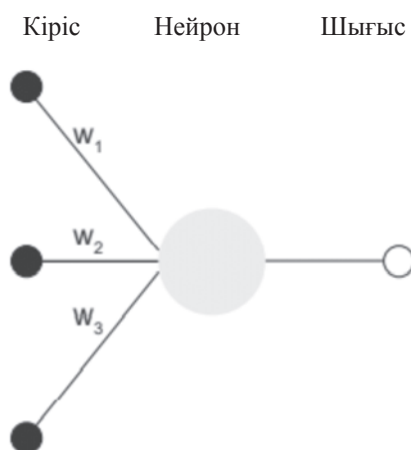
болып келеді. Үйірткілі нейронды желілер жадыда көп орынды қажет етпейді, сонымен қатар, қолданылатын көрсеткіштердің санын азайту арқылы оқу жылдамдығын арттыруға болады. Үйірткілі нейронды желілерін тану есептеріне қолданғанда басқа нейрондық желілердің өнімділігінен бірнеше есе артады. Үйірткілі нейронды желілер оптикалық таңбаларды тану (OCR – Optical Character Recognition) жүйелерінде көптеген жетістіктерге қол жеткізуге мүмкіндік берді. Суреттегі таңбаны тану – кіріс кескініндегі әр таңбаға алдын-ала анықталған класты сәйкес қоятын үдеріс. Қазіргі таңда әртүрлі тану қосымшаларына сұраныс артқанына байланысты компьютерді «оқуға» үйретудің шұғыл қажеттілігі туындады.

Терең оқыту – бұл оқыту үшін нейрондық желілерді қолданатын машиналық оқытудың бір түрі. Жасанды нейрондық желі (немесе жай нейрондық желі) - бұл біздің миымыз жұмыс істейтін принциптер бойынша жұмыс істейтін бағдарламалық жасақтама. Біздің биологиялық жүйке жүйесін басқару синапстар деп аталатын арналар арқылы бір-бірімен өзара әрекеттесетін нейрондар арқылы жүзеге асырылады. Нейрондық желіде өзара байланысты жасанды нейрондар желіні оқыту үшін адам миының нейрондарының жұмысын модельдейді. Оқу процесінде қажетті нәтижеге жету үшін белгілі бір нейрондар арасындағы байланыс күшейтіледі. Келесі диаграммада үш деңгейлі нейрондық желі көрсетілген (Сурет 1). Әр шеңбер нейронды білдіреді, ал олардың арасындағы сызықтар синапстарды модельдейді.



Сурет 1 – Үш деңгейлі нейрондық желі

Осы суретте толық байланысқан желі көрсетілген – берілген деңгейдегі әрбір нейрон келесі деңгей нейрондарымен байланысады. Нейрондық желілерді оқыту итеративті принцип бойынша жүреді. Әрбір итерация дәуір (epoch) деп аталады. Оқу жиынтығындағы үлгі бір рет өңделеді және «дұрыс» дәуірлер саны жоқ. Бұл гиперкөрсеткішті оқу деректеріне және моделіне байланысты баптау қажет болды. Желінің кіріс мәліметтері ретінде оқу үлгілерінің белгілері алынады. Бір деңгейлер алдыңғы деңгейлердің шығысы бойынша жаңа белгілерді зерттесе, басқалары осы белгілерді болжам жасау үшін қолданады. Келесі суретте барлық нейрондарға (немесе олардың бөліктеріне) берілетін үш кірістен, бір шығыстан тұратын нейрон бейнеленген (Сурет 2):



Сурет 2 – Үш кірістен, бір шығыстан тұратын нейрон

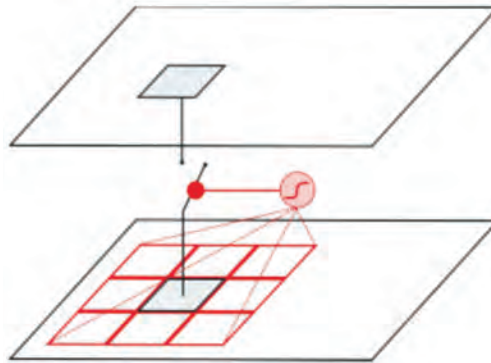
Мұнда w_1 , w_2 және w_3 мәндері салмақ деп аталады. Жаңадан оқылатын модельдерде бұл мәндер модель арқылы кездейсоқ қосылады. Оқыту барысында желі болжамды модель белгілері мен нақты үлгі белгілері арасындағы қателіктердің үлесін азайтуға тырысады. Қателіктердің үлесі шығындар деп аталады, ал шығындар мөлшерін анықтау формуласы шығындар функциясы деп аталады. Оқу процесінде желі әр нейронның жалпы шығындарға әкелетін мөлшерін анықтайды, содан кейін ол деңгейлерден кері бағытта өтеді және шығындарды азайтуға тырысып, салмақты реттейді. Бұл әдіс кері таралу деп аталады. Салмақтарды оңтайландыру градиентті түсу деп аталатын үдерісте біртіндеп іске асады [9].

Нейрондық желілер қолжазба мен сөйлеуді тануда заманауи жүйелерінің негізгі құрамдас бөлігі болып отыр. Дербес жағдайда, қолжазба мен сөйлеуді тануда ең тиімді болып танылған рекурентті нейронды желілер.

Қолжазба адамдарға ақпарат тарату үшін қолданылады. Әрі қарай талдау үшін осы ақпаратқа қол жеткізу үшін бетті оптикалық сканерлеп, машина танитын пішінге түрлендіру керек. Осы орайда, зерттеушілер терең оқытуға негізделген объектілерді анықтау архитектурасын қолдана отырып, ағылшын тіліндегі қолжазба сөздерін тану моделін ұсынған [10]. Бұл модель авторлардың жазу стиліне тәуелді болмай жақсы жұмыс істейді.

Қазақ тіліндегі қолжазбаларды тану үшін Theodore Bluche and Ronaldo Messina ұсынған [11] екі бағытты LSTM (long short-term memory) декодеріне негізделген модельді қолданамыз. Модель авторлары IAM және Rimes дерекқорларындағы жолдар мен абзацтар деңгейіндегі ең заманауи тану нәтижелерін алған. Конволюциялық нейрондық желілер компьютерлік көруде кеңінен қолданылатынын және екі бағытты LSTM тілдік қолданбалар үшін өте танымал екенін ескере отырып, кіріс кескінінің кодері, интерфейс және декаодер деп аталатын үш негізгі бөліктен тұратын терең нейрондық желі құрылады. Кіріс кескінінің кодері конволюциялық қабаттардан жасалған. Ол екі өлшемді көріністерді өңдейді және объектілердің 2D карталарын ұсынады. Бұл бөлікте модельдің шамамен 20% бос параметрлері бар, бірақ архитектураның

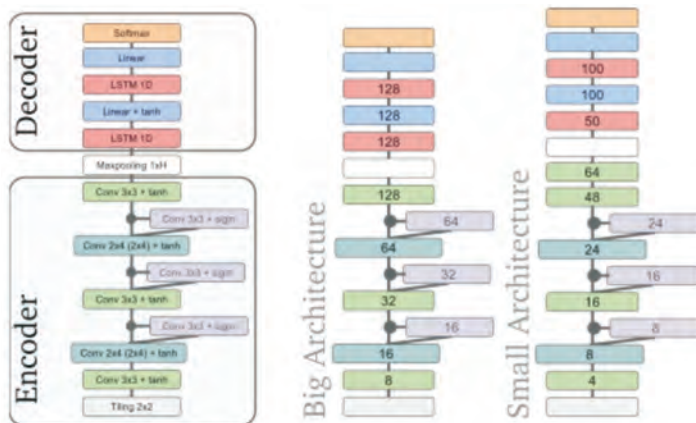
ең баяу компоненті (есептеу 80%). Мақсат – осы бөлікті қайта пайдалану үшін мүмкіндігінше әмбебап ету (мысалы барлық тілдер немесе әртүрлі тапсырмалар үшін өңдеу). Интерфейс 2D кескініне ұқсас көріністі күтілетін 1D көрінісіне түрлендіреді (біз таңбалар тізбегін болжаймыз). Декодер – таңбалар тізбегін болжау үшін белгілер тізбегін өңдейтін екі бағытты LSTM RNN(Recurrent Neural Network). Ол желінің сыйымдылығының көп бөлігін алады (модель өлшемінің 80%), бірақ жылдам өңдеуді қамтамасыз етеді (есептеу уақытының 20%). Ұсынылған модельде қай контексте есептеу белгісі релевантты екенін жүйе тани алады (Сурет 3).



Сурет 3 – Конволюциялық элемент (қызыл түспен белгіленген)

Бұл жерде кодерде қолданылатын конволюциялық элементтер сипатталған. Басқару элементі объектінің келесі қабатқа таралуын басқарады. Негізінде, шлюз берілген позициядағы объектінің мәнін және іргелес мәндерді қарастырады және сол нысанды сол күйде сақтау немесе одан бас тарту туралы шешім қабылдайды. Бұл бүкіл суреттегі жалпы белгілерді есептеуге және контекстке сәйкес бұл белгілер маңызды болған кезде сүзуге мүмкіндік береді.

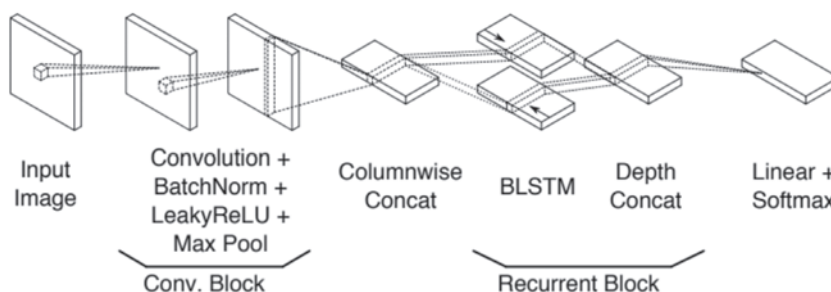
Әр түрлі тапсырмаларды (дәлдік, жылдамдық және өлшем) оңтайландыру үшін нейрондық желілердің архитектурасы құрылды (Сурет 4).



Сурет 4 – Ұсынылған нейрондық желілердің архитектурасы

Theodore Bluche and Ronaldo Messina ұсынған модель қолжазбаны танудың ең заманауи нәтижелерін беретін нейрондық желі архитектурасынан тұрады. Ол қолжазба мәтінінің жалпы белгілерін шығаратын конволюциялық кодерден және таңбалар тізбегін болжайтын әр тілге бейімделген LSTM декодерінен тұрады. Кодтаушының маңызды аспектісі – тиісті функцияларды таңдауға және қалғандарын блоктауға қабілетті конволюциялық қабаттар түрінде жүзеге асырылатын басқару элементтері.

Оффлайн қолжазбаны танудың заманауи тәсілдері LSTM ұзақ қысқа мерзімді жақтың көп өлшемді желілеріне негізделген. LSTM (ағылш. Long short-term memory) – ұзақ қысқа мерзімді жады желісі, бұл рекурентті нейронды желілердің архитектурасының бір түрі. Алайда, бұл архитектуралар өте жоғары есептеу шығындарына ие. Joan Puigserver осы архитектураға балама ретінде айтарлықтай жылдамырақ жұмыс істейтін тек конволюциялық және бір өлшемді қайталанатын қабаттарға сүйенетін модельді ұсынды [12]. Ұсынған модельдің архитектурасы үш блоктан тұрады: Конволюциялық блоктар рекурентті блоктар және сызықты блоктар (Сурет 5).



Сурет 5 – 1D-LSTM екі бағытты қолданатын нейрондық желі архитектурасы

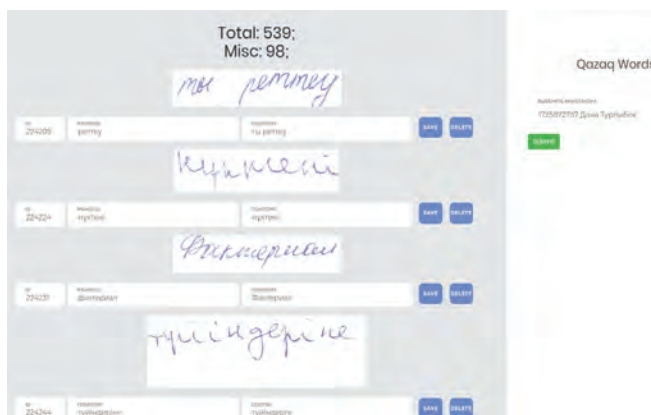
Joan Puigserver зерттеулерінде қолжазаны тану дәлдігіне қол жеткізу үшін көп өлшемді қайталанатын қабаттар қажет болмауы мүмкін екендігі туралы көптеген дәлелдер келтірген. Сонымен қатар, эксперименттік нәтижелердің статистикалық негізделген талдауы кеңінен қолданылатын IAM, Rimes екі деректер жиынтығы үшін бұл мәлімдемені растайды.

Қолжазба мәтінін тану конволюциялық қайталанатын нейрондық желілерге (ағылш. Convolutional Recurrent Neural Networks, CRNNs) негізделген. Тану үрдісінде конволюциялық қайталанатын нейрондық желілер, рекуренттік нейрондық желілерді пайдаланғанда құжаттарды сканерлеген кезде градиенттің жоғалуы, үлкен көлемдегі деректер мен есептеу ресурстарын қажет ететін миллиондаған параметрлер есептеулерді баялатуы мүмкін. Жақында зерттеушілер CRNN әдістерін толықтыратын жаңа класының архитектурасының модельдеп, жабық конволюциялық қайталанатын нейрондық желіні (ағылш. Gated Convolutional Neural Networks, Gated-CNN) ұсынды [13]. Gated-CRNN жаңа архитектурасы машиналық оқытудың соңғы тәсілдері Dauphin ұсынған Gated механизмi [14] мен екі бағытты басқарылатын қайталанатын модуліні (ағылш. Bidirectional Gated Recurrent Units, BGRU) [15] біріктіреді. Ұсынған Gated-CRNN моделі бірнеше параметрлерді (мыңдаған) қамтиды және мәтінді тану процесінде (жол деңгейінде және сегментациясыз) қателерді азайтады. Сонымен қатар, оқытылатын параметрлердің санын (мыңдаған) азайта отырып, есептеу шығындарын да кемітуге болады.

Жақында кириллица негіздегі қазақ және орыс тілдерінде қолжазбаны тану модельдерін өңдеу бойынша алғашқы рет [16] зерттеу жұмытарының нәтижесі жарияланды. Ұсынылған модель қолжазба мәтіннің жалпы сипаттамаларын бөліп шығаратын толығымен жабық конволюциялық кодерден, BGRU декодерден және таңбалар тізбегінің болжау үшін CTC (Connectionist temporal classification, CTC) моделінен тұрады. Бұл әдістің тиімділігі үш тілде: ағылшын, орыс және қазақ тілдерінде қолжазба мәтіндерінің дерекқорларын пайдалана отырып бағаланды. Ол басқа танымал модельдерге қарағанда қолжазба қазақ және орыс тілдерінде HKR [6] деректер жиынында үздік нәтижелерді көрсетті.

Нәтижелер және оларды талқылау. Барлық модельдер Python ортасында Tensorflow, Keras, Matplotlib кітапханаларының көмегімен жүзеге асырылды. TensorFlow – Google Brain командасы жасаған сандық есептеулер мен ауқымды машиналық оқытуға арналған Ашық бастапқы кітапхана. TensorFlow Машиналық оқыту мен терең оқытудың көптеген модельдері мен алгоритмдерін біріктіреді және оларды жалпы метафора ретінде пайдалы етеді. Keras – TensorFlow Машиналық оқыту платформасының үстінде жұмыс істейтін Python тілінде жазылған терең оқыту API. Keras-тың негізгі авторы-Google инженері Francois Chollet [17]. Keras iOS және Android смартфондарында терең оқыту үлгілерін енгізуге ниетті пайдаланушылар үшін мобильді платформаларды әзірлеуді ұсынады. Matplotlib-бұл деректерді визуализациялауға арналған пакет. Ол графиктердің кең ауқымын құруға арналған құралдарды ұсынады және жоғары деңгейдегі баптауларды қамтамасыз етеді.

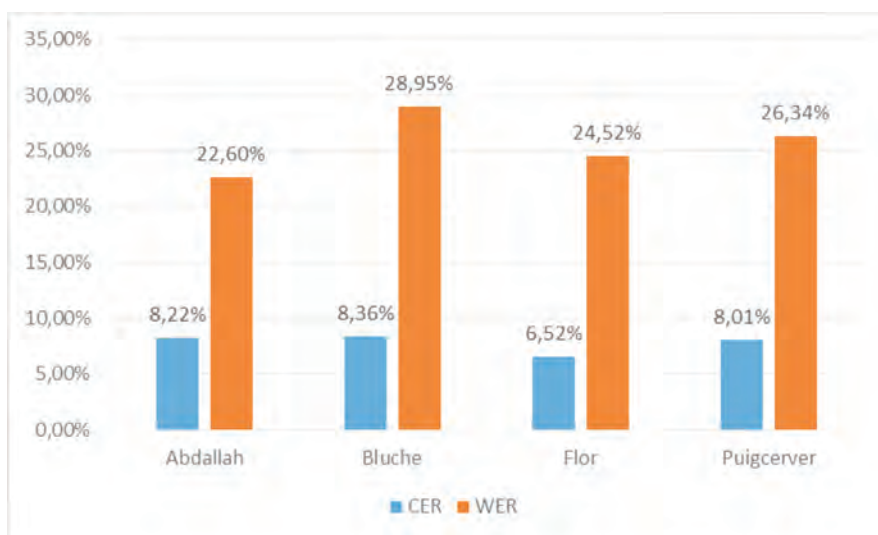
Қазақ тілінде қолжазба деректерді жинау үшін арнайы Телеграм бот жасалды. Телеграм ботқа студенттердің қазақ тілінде 6000-ға жуық А4 форматта емтихан парағы сканерленді. Бұл алынған суреттер генетикалық алгоритм негізінде сөздерге сегменттелген. Осы сегменттелген сөздер ерікті пайдаланушыларға Телеграм боттың сілтемесі жіберілді. Ерікті пайдаланушылар Телеграм бот ұсынған сөзді енгізеді. Аннотацияларды жинау процесін жеңілдету үшін бір суретті екі кездейсоқ пайдаланушыға жіберу туралы шешім қабылданды. Егер екі пайдаланушыдан жіберілген аннотациялар абсолютті бірдей болса, біз бұл нәтижені сенімді деп есептеді. Деректер жиынтығын толық тексерілді. Сәйкес келмейтін сөздер үшін нейрондық желілер негізінде жауаптарды түзету құралы да жасалды (Сурет 7).



Сурет 7 – Нейрондық желілер негізінде жауаптарды түзету құралы

Жалпы қолжазба деректерін жинау барысында ерікті пайдаланушылар аннотацияларды өте ұқыпсыз толтырғаны, кейде тіпті дұрыс емес аннотациялар жазғаны белгілі болды. Осы жағдайларды болдырмау үшін біз нейрондық желіні бұрыннан бар толық емес деректер жиынтығында қолжазбаны тануға үйретілді. Нәтижесінде, деректер жиынтығында нейрондық желі субъективті түрде тамаша нәтижелер берді.

Барлық модельдер КОНТД қолжазба деректер жиынында оқытылды. Оқыту, валидация және тестілеу үшін деректер жиынтығын дайындағаннан кейін модельдер оқытылды және салыстырмалы бағалау бойынша бірқатар эксперименттер жүргізілді. Модельдердің таңбалардағы (CER) және сөздердегі (WER) қателер жиілігі көрсеткіштері бойынша бағаланды (сурет 6).



Сурет 6 – Модельдердің таңбалардағы (CER) және сөздердегі (WER) қателер жиілігі

Бұл зерттеу жұмысында Abdallah, Bluche, Flor және PUIGCERVER сияқты танымал терең оқыту модельдері пайдалана отырып, қазақ тіліндегі қолжазба мәтіндерін тану мәселесін шешуге әрекет жасалды.

Қорытынды. Бұл мақалада жасанды нейрондық желі негізінде кириллица негізіндегі қазақ тілінде қолмен жазылған мәтінді тану моделдері қарастырылған. Бұл модельлер нейрондық желінің архитектурасы конволюциялық нейрондық желі қабаттарынан тұрады және рекурренттік нейрондық желі, сондай-ақ коннекциялық уақыт жіктемесінен тұрады. Терең оқытудың әдістеріне шолу жасалып, эксперименттер жүргізілді. Эксперименттерге нәтижесіне сәйкес, Floop моделі қазақ тілінде қолмен жазылған мәтінді тануда ең жоғары жылдамдығын көрсетті.

ӘДЕБИЕТ

1 Keysers D, Deselaers T, Rowley H. A., Wang L. -L., Carbune V. Multi-Language Online Handwriting Recognition. (2017). IEEE Transactions on Pattern Analysis and Machine Intelligence. <https://doi.org/10.1007/s11042-018-6764-0>

2 Marti U.-V., Bunke H. (2002). The IAM-database: an English sentence database for offline handwriting recognition. *International Journal on Document Analysis and Recognition*. <https://doi.org/10.1007/s100320200071>

3 Cohen G, Afshar S, Tapson J, van Schaik A. (2017). EMNIST: Extending MNIST to handwritten letters. 2017 International Joint Conference on Neural Networks (IJCNN). <https://doi.org/10.1109/ijcnn.2017.7966217>

4 Su T., Zhang T., Guan D. (2007). Corpus. based HIT-MW database for offline recognition of general-purpose Chinese handwritten text. <https://doi.org/10.1007/s10032-006-0037-6>

5 Serrano N, Castro F, Juan A.(2010). The Rodrigo database. *Proceedings of the Seventh International Conference on Language Resources and Evaluation*. Valletta, Malta

6 Nurseitov D, Bostanbekov K, Alimova A, Abdallah A, Kurmankhojayev D. Tolegenov R. (2021). Handwritten Kazakh and Russian (HKR) database for text recognition. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-021-11399-6>

7 Toiganbayeva N., Kasem M, Abdimanap G., Bostanbekov K., Abdallah A., Alimova A., Nurseitov D. (2022). KOHTD: Kazakh Offline Handwritten Text Dataset. *Signal Processing: Image Communication*. Elsevier. doi: <https://doi.org/10.1016/j.image.2022.116827>

8 Shrivastava V., Sharma N. (2012). Artificial Neural Networks Based Optical Character Recognition. *Signal & Image Processing: An International Journal (SIPIJ)*. <https://doi.org/10.48550/arXiv.1211.4385>

9 Дейтел Пол, Дейтел Харви. (2020). Python: Искусственный интеллект, большие данные и облачные вычисления. СПб.: Питер

10 Mondal R., Malakar S., Barney Smith, E.H. et al. (2022). Handwritten English word recognition using a deep learning based object detection architecture. *Multimed Tools Appl.* vol. 81, pp. 975–1000 (2022). <https://doi.org/10.1007/s11042-021-11425-7>

11 Bluche T., Messina R. (2017). Gated Convolutional Recurrent Neural Networks for Multilingual Handwriting Recognition. 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). Kyoto, Japan. <https://doi.org/10.1109/icdar.2017.111>

12 Puigcerver J. (2017). Are Multidimensional Recurrent Layers Really Necessary for Handwritten Text Recognition? 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan. <https://doi.org/10.1109/ICDAR.2017.20>

13 A. F. de Sousa Neto, B. L. D. Bezerra, A. H. Toselli and E. B. Lima. (2020). HTR-Flor: A Deep Learning System for Offline Handwritten Text Recognition .33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Porto de Galinhas, Brazil. pp. 54-61, doi: <https://doi.org/10.1109/SIBGRAPI51738.2020.00016>

14 Dauphin Y. N., Fan A., Auli M., Grangier D. (2017) Language modeling with gated convolutional networks. *Proceedings of the 34th International Conference on Machine Learning*. Vol. 70, ser. ICML'17. JMLR.org, 2017, p. 933–941

15 Cho K, van Merriënboer B., Bahdanau D., Bougares, Fethi S., Bengio Y. (2020). Learning phrase representations using RNN encoder–decoder for statistical machine translation. 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, Doha, Qatar. <https://doi.org/10.3115/v1/D14-1179>

16 Abdallah Abdelrahman, Mohamed Hamada, Daniyar B. Nurseitov. (2020). Attention-Based Fully Gated CNN-BGRU for Russian Handwritten Text. *Journal of Imaging*. <https://api.semanticscholar.org/CorpusID:221150485>

17 Chollet F.(2017) Deep Learning with Python. United States of America

REFERECES

- 1 Keysers D, Deselaers T, Rowley H. A., Wang L. -L., Carbune V. Multi-Language Online Handwriting Recognition. (2017). IEEE Transactions on Pattern Analysis and Machine Intelligence. <https://doi.org/10.1007/s11042-018-6764-0>
- 2 Marti U.-V., Bunke H. (2002). The IAM-database: an English sentence database for offline handwriting recognition. International Journal on Document Analysis and Recognition. <https://doi.org/10.1007/s100320200071>
- 3 Cohen G, Afshar S, Tapson J, van Schaik A. (2017). EMNIST: Extending MNIST to handwritten letters. 2017 International Joint Conference on Neural Networks (IJCNN). <https://doi.org/10.1109/ijcnn.2017.7966217>
- 4 Su T., Zhang T., Guan D. (2007). Corpus. based HIT-MW database for offline recognition of general-purpose Chinese handwritten text. <https://doi.org/10.1007/s10032-006-0037-6>
- 5 Serrano N, Castro F, Juan A.(2010). The Rodrigo database. Proceedings of the Seventh International Conference on Language Resources and Evaluation. Valletta, Malta.
- 6 Nurseitov D, Bostanbekov K, Alimova A, Abdallah A, Kurmankhojayev D. Tolegenov R. (2021). Handwritten Kazakh and Russian (HKR) database for text recognition. Multimedia Tools and Applications. <https://doi.org/10.1007/s11042-021-11399-6>
- 7 Toiganbayeva N., Kasem M, Abdimanap G., Bostanbekov K., Abdallah A., Alimova A., Nurseitov D. (2022). KOHTD: Kazakh Offline Handwritten Text Dataset. Signal Processing: Image Communication. Elsevier. doi: <https://doi.org/10.1016/j.image.2022.116827>
- 8 Shrivastava V., Sharma N. (2012). Artificial Neural Networks Based Optical Character Recognition. Signal & Image Processing: An International Journal (SIPIJ). <https://doi.org/10.48550/arXiv.1211.4385>
- 9 Dejtjel Pol, Dejtjel Harvi. (2020). Python: Iskusstvennyj intellekt, bol'shie dannye i oblachnye vychisleniya. SPb.: Piter
- 10 Mondal R., Malakar S., Barney Smith, E.H. et al. (2022).Handwritten English word recognition using a deep learning based object detection architecture. Multimed Tools Appl. vol. 81, pp. 975–1000 (2022). <https://doi.org/10.1007/s11042-021-11425-7>
- 11 Bluche T., Messina R. (2017). Gated Convolutional Recurrent Neural Networks for Multilingual Handwriting Recognition. 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). Kyoto, Japan. <https://doi.org/10.1109/icdar.2017.111>
- 12 Puigcerver J. (2017). Are Multidimensional Recurrent Layers Really Necessary for Handwritten Text Recognition? 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan. <https://doi.org/10.1109/ICDAR.2017.20>
- 13 A. F. de Sousa Neto, B. L. D. Bezerra, A. H. Toselli and E. B. Lima. (2020). HTR-Flor: A Deep Learning System for Offline Handwritten Text Recognition .33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Porto de Galinhas, Brazil. pp. 54-61, doi: <https://doi.org/10.1109/SIBGRAPI51738.2020.00016>
- 14 Dauphin Y. N., Fan A., Auli M., Grangier D. (2017) Language modeling with gated convolutional networks. Proceedings of the 34th International Conference on Machine Learning. Vol. 70, ser. ICML'17. JMLR.org, 2017, p. 933–941
- 15 Cho K, van Merriënboer B., Bahdanau D., Bougares, Fethi S., Bengio Y. (2020). Learning phrase representations using RNN encoder–decoder for statistical machine translation. 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics, Doha, Qatar. <https://doi.org/10.3115/v1/D14-1179>
- 16 Abdallah Abdelrahman, Mohamed Hamada, Daniyar B. Nurseitov. (2020). Attention-Based Fully Gated CNN-BGRU for Russian Handwritten Text. Journal of Imaging. <https://api.semanticscholar.org/CorpusID:221150485>
- 17 Chollet F.(2017) Deep Learning with Python. United States of America

**Н. А. ТОЙГАНБАЕВА¹, Ж. Х ЖУНУСОВА¹, А. ПРОВАТА²,
Г. А. ТЮЛЕПБЕРДИНОВА¹**

¹Казахский национальный университет имени аль-Фараби, г. Алматы, Казахстан

²Институт нанонауки и нанотехнологий, г. Афины, Греция

РАСПОЗНАВАНИЕ ОФЛАЙН-РУКОПИСНЫХ ТЕКСТОВ НА КАЗАХСКО-РУССКОМ ЯЗЫКЕ НА ОСНОВЕ МОДЕЛЕЙ ГЛУБОКОГО ОБУЧЕНИЯ

Оцифрованный текст рукописных заметок позволяет автоматизировать бизнес-процессы многих компаний и упрощает работу человека. В статье рассматривается распознавание письменной рукописи на русском и казахском языках с использованием моделей углубленного обучения. Поскольку почерк каждого человека уникален и нет возможности создавать общие образцы рукописного текста, отчет о распознавании текста в автономном режиме считается более сложным, чем отчет о распознавании в Интернете. В работе используются различные подходы к глубокому обучению при разработке моделей распознавания рукописного ввода на казахско-русском языке, касающихся кириллической графики. Важным элементом успешного глубокого обучения является доступность данных, вычислений, программных платформ и времени выполнения, что упрощает создание и выполнение моделей нейронных сетей для их создания. Обсуждались концепции глубокого обучения, конкурентоспособной, динамично развивающейся отрасли, предлагающей быстрые, количественные и справедливые инструменты для анализа и сравнения различных подходов и методов обучения. Были проанализированы популярные модели глубокого обучения, такие как Abdallah, Bluche, Flor и PUIGSERVER, а результаты экспериментов были проанализированы на основе гистограммы. Эксперименты проводились на основе большой базы данных офлайн-рукописных текстов на казахском языке под названием *Kazakh Offline Handwritten Text Dataset (КОНТД)*. Создан телеграм-бот, созданный специально для сбора рукописных данных на казахском языке. Инструмент для проверки информации и исправления ответов, введенный этим телеграм-ботом, был создан на основе нейронной сети.

Ключевые слова: почерк, глубокое обучение, нейронная сеть, КОНТД, наборы данных.

**N. A TOIGANBAYEVA¹, ZH. ZHUNUSSOVA¹,
A. PROVATA², G. A. TYULEPBERDINOVA¹**

¹Al-Farabi Kazakh National University, Almaty, Kazakhstan

²Institute of Nanoscience & Nanotechnology, Athens, Greece

RECOGNITION OF OFFLINE HANDWRITTEN TEXTS IN KAZAKH-RUSSIAN BASED ON DEEP LEARNING MODELS

The digitized text of handwritten notes allows you to automate the business processes of many companies and facilitates the work of a person. The article discusses the recognition of written handwriting in Russian and Kazakh languages using in-depth learning models. Due to the fact that each person's handwriting is unique and there is no way to create general samples of handwritten text, the offline text recognition report is considered more complex than the online recognition report. The work uses various approaches to deep learning in the development of models of handwriting recognition in the Kazakh-Russian languages related to Cyrillic graphics. An important element of deep learning success is the availability of data, calculations, software platforms, and runtime, which makes it easier to build neural network models