

МРНТИ 27.01.43
УДК 51:37.019

<https://doi.org/10.51889/2020-2.1728-7901.37>

К.С. Иманбаев¹, Б.Д. Шарипова¹, Ж.Ж. Қожамқұлова¹, Е.Е. Аубакиров¹

Алматы Технологиялық Университеті, Алматы қ., Қазақстан

НЕЙРОНДЫҚ ЖЕЛІЛЕРДІ ТЕРЕҢДЕТІП ОҚЫТУДЫ ЗЕРТТЕУ

Аңдатпа

Қазіргі уақытта сөйлемді тану, табиғи тілді өңдеу, компьютерлік көруді зерттеу нейрондық желілерді терең оқытуда танымал тәсілдер мен әдістердің бірі болып табылады. Мақалада нейрондық желілерді терең оқытудың даму тарихы мен жалпы қазіргі жағдайы туралы шолу қарастырылған. Тереңдетіп оқыту үшін қолданылатын нейрондық желілерді оқыту алгоритмі кері таралу әдісінің көмегімен дәл баптауды таңдау және оны қолдану қарастырылған. Нейрондық желілерде жасырын қабаттардың саны көп, жиі кездесетін және жоғалып бара жатқан градиенттерді оқыту қиындығымен ерекшелінеді.

Қарастырылып отырған жұмыста көп қабаттар (жүзден астам) және жойылып бара жатқан градиенттер кезінде нейрондық желілерді оқытудың табысты іске асыратын әдістері қарастырылады.

Түйін сөздер: терең оқыту, нейрондық желілер, градиенттік әдіс, алгоритм, рекуррентті желі.

Аннотация

К.С. Иманбаев¹, Б.Д. Шарипова¹, Ж.Ж. Қожамқұлова¹, Е.Е. Аубакиров¹

¹ Алматының технологиялық университеті, Алматы, Қазақстан

ИССЛЕДОВАНИЕ УГЛУБЛЕННОГО ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ

В настоящее время при распознавании речи, обработка естественного языка, компьютерное зрение одним из популярных методов является глубокое обучение нейронных сетей. В статье выполнен обзор истории развития глубокого обучения нейронных сетей и современное состояние в целом. Рассматриваются алгоритмы для обучения нейронных сетей применяемый для глубокого обучения нейронных сетей, последующий тонкой настройки с помощью метода обратного распространения ошибки. Нейронные сети при больших количествах скрытых слоев, часто встречающихся и исчезающих градиентов очень трудно обучить. В данной работе рассматриваются методы, успешно реализующие обучение нейронных сетей при больших количествах слоев (более ста) и исчезающих градиентов.

Проводится обзор известных библиотек применяемых для успешного глубокого обучения нейронных сетей.

Ключевые слова: глубокое обучение, нейронные сети, градиентный метод, алгоритм, рекуррентная сеть.

Abstract

RESEARCH IN-DEPTH TRAINING OF NEURAL NETWORKS

Imanbaev K.S.¹, Sharipova B.D.¹, Kozhamkulova Zh.Zh.¹, Aubakirov E.E.¹

¹ Almaty technological University, Almaty, Kazakhstan

Currently, deep learning of neural networks is one of the most popular methods for speech recognition, natural language processing, and computer vision. The article reviews the history of deep learning of neural networks and the current state in General. We consider algorithms for training neural networks used for deep training of neural networks, followed by fine-tuning using the method of back propagation of errors. Neural networks with large numbers of hidden layers, frequently occurring and disappearing gradients are very difficult to train. In this paper, we consider methods that successfully implement training of neural networks with large numbers of layers (more than one hundred) and vanishing gradients.

A review of well-known libraries used for successful deep learning of neural networks is conducted.

Keywords: deep learning, neural networks, gradient method, algorithm, recurrent network.

Қазіргі уақытта нейрондық желілерді тереңдетіп оқыту сөйлемді тану, табиғи тілді өңдеу, компьютерлік жүйелеуді көрсету және т.б. сияқты жасанды интеллект жүйесін құрудың ең танымал тәсілдерінің бірі болып табылады. Мақалада терең нейрондық желілерді оқыту әдістерінің қазіргі жағдайы мен даму тарихына шолу жасалған. Нейрондық желілер архитектурасының дамуы сипатталады: неокогнитрон, автокодирлеушілер, жинақталған нейрондық желілер, Больцманның шектелген машинасы, тереңдетілген сенімділік желілері, ұзақ-қысқа мерзімді жады желілері, басқарылатын рекурренттік нейрондық желілер және қалдықты оқыту желілері. Жасырын қабаттары көп тереңдетілген нейрондық желілерді жоғалып бара жатқан градиент проблемасына байланысты оқыту қиынға соғады. Мақалада жүз қабаттан астам тереңдетілген нейрондық желілерді табысты оқытуға мүмкіндік беретін осы проблеманы шешу әдістері қарастырылады. Нейрондық желілерді

терең оқытудың танымал кітапханаларына шолу жасалады, олар осы технологияны кеңінен практикалық тәсілде қолдануға мүмкіндік береді. Қазіргі уақытта компьютерлік жүйелерді көру міндеті үшін жинақталған нейрондық желілер, ал тізбектер үшін, оның ішінде табиғи тілді өңдеу үшін рекурренттік нейрондық желілер, ең алдымен ұзақ – қысқа мерзімді жады желілері және басқарылатын рекурренттік нейрондық желілер қолданылады.

Тереңдетілген нейрондық желілер қазіргі уақытта машиналық оқытудың ең танымал әдістерінің бірі болып табылады. Олар сөйлемді тану, табиғи тілді өңдеу, компьютерлік жүйелеуді көрсету [1], медициналық информатика және т.б. салалардағы баламалы әдістермен салыстырғанда жақсы нәтижелер көрсетеді. Тереңдетілген нейрондық желілердің табысты қолданылуының себептерінің бірі, желі автоматты түрде маңызды белгілерін шешу үшін қажетті міндеттерін бөледі. Машинамен оқытудың балама алгоритмдерінде белгілер адамдармен бөлінуі тиіс, олар арнайы зерттеу бағыты бар — ерекшелік инженериясы (feature engineering). Үлкен көлемді деректерді өңдеу кезінде нейрондық желі белгілердің бөлінуін адамға қарағанда әлдеқайда жақсы орындайды.

Мақалада терең нейрондық желілер архитектурасының дамуын және оларды оқыту тәсілдеріне тарихи шолу қарастырылған. Терең нейрондық желілердің нұсқалары өте көп ұсынылған және терминология уақыт өте өзгеріп отырады. Жасанды нейрондық желілер моделі 1943 жылы ұсынылған [2], ал терең оқыту термині (deep learning) 2006 жылдан бастап кеңінен қолданыла бастады [3]. Бұған дейін терминдер тереңдетілген желілерін жүктеу (loading deep networks), және оқытудағы терең жад деп қолданылған (learning deep memories).

Соңғы бірнеше жылда болған терең нейрондық желілердің танымалдығының өсуін үш фактормен түсіндіруге болады. Біріншіден, GPU (Graphics Processing Unit) есептеуіштерінің жылдамдатқыштарының өнімділігі айтарлықтай артуы болды, бұл желінің тереңдетілген нейрондық желілерін анағұрлым жылдам әрі жоғары дәлдікпен оқытуға мүмкіндік берді. Бұрын есептеуіш қуаттары бар практикалық міндеттерді шешу үшін қандай да бір жарамды күрделі желіні оқыту үшін жетіспеді. Екіншіден, терең нейрондық желілерді оқыту үшін қажетті деректердің үлкен көлемі жинақталған. Үшіншіден, жүздеген және одан да көп қабаттардан тұратын желілерді жылдам және сапалы оқытуға мүмкіндік беретін нейрондық желілерді оқыту әдістері әзірленген, бұл бұрын жоғалып бара жатқан градиент және қайта оқыту проблемаларынан қарастыру мүмкін емес еді. Үш факторлардың үйлесуі терең нейрондық желілерді оқытуда және оларды практикалық пайдалануда айтарлықтай прогреске қол жеткізді және бұл терең нейрондық желілерге машиналық оқыту әдістері арасында жетекші орынға ие болуға мүмкіндік берді [4].

Нейрондық желілерді оқыту деп-нейрондар арасындағы қосылыстардың таразыларын анықтау процесінде желі берілген дәлдікпен қажетті функцияны жақындататындай етіп түсіндіріледі. Нейрондық желілерді оқытуда үш тәсілдің бары белгілі, олар: мұғаліммен оқыту (supervised learning), мұғалімсіз оқыту (unsupervised learning) және бекіте оқыту (reinforcement learning). Мұғаліммен бірге оқыту кезінде желіге кіру сигналдары (объектілер) беріледі, олар үшін алдын ала дұрыс жауаптары белгілі болады (оқыту жиыны). Шығу сигналы салмаққа байланысты және желіге дұрыс берілгеніне байланысты белгілі бір ережелер бойынша өзгереді. Мұғалімсіз оқыту кезінде желінің кіре берісіне дұрыс шығу сигналында алдын ала белгісіз объектілер беріледі. Бекіте оқыту кезінде желі өзара әрекеттесетін сыртқы ортаның болуын болжайды. Оқыту процесі осы ортадан алынған сигналдардың негізінде жүргізіледі. Маккаллок-Питтс нейрондық желілерінің негізгі кемшілігі – нейрондардың барлық сигналдарының кірісі үшін олардың салмаққа сәйкес нәтижелері алдын ала берілуі тиіс [5].

Қазіргі уақытта нейрондық желілерді, оның ішінде тереңдетілген желілерді оқыту үшін градиенттік төмен түсу әдісіне негізделген қатенің кері таралу алгоритмі (error backpropagation algorithm) пайдаланылады. Бұл алгоритмді нейрондық желілерде оқыту үшін алғаш рет 1981 жылы шыққан жұмыста қолданылып, сипатталған. Осыдан кейін бұл тақырыпқа бірнеше жұмыстар пайда болды. Қатені кері тарату алгоритмінде мұғаліммен оқыту тәсілі пайдаланады, ол үшін алдын ала белгілі дұрыс жауаптары бар оқыту жиыны қажет.

Тереңдетілген нейрондық желіде бірнеше жасырын қабаттары бар қатені есептеу жүргізіледі, ол бір қабаттан екіншісіне беріледі. Бірінші кезеңде біз білетін дұрыс жауаптарды нейрондық желінің шығу қателігінің мәнін есептейді. Содан кейін жасырын қабаттың шығуында қате ретінде пайдаланылатын желінің шығыс қабатына кіріс қатесі есептеледі. Осылайша, есептеудегі кіріс қабатындағы қате белгілі сәтке дейін жалғасады. Сондықтан алгоритм қателіктің кері таралуы деген атқа ие болады.

1992 жылы ұсынылған зерттеуде өте терең білімді ғалым жоғалып бара жатқан градиент мәселесін шешуді және рекуррентті нейрондық желілер иерархиясын мұғалімсіз алдын ала оқытуды қолдану есебінен жүздеген қабатқа дейін тереңдікті оқытуды қамтамасыз ететіндігін дәлелдеді. Әрбір

рекурренттік желі келесі мәнді болжау үшін бөлек оқытылады. Оқытудан кейін тек қате болжамды мәндер желінің жоғары деңгейіне беріледі. Бұл желі баяу уақыт шкаласында жұмыс істейді, соның есебінен ақпарат толық қанды берілмейді және сигналдардың әрбір реттілігі желінің барлық терең деңгейлерінде азғана артық кодтау жиынтығына сәйкес келеді. Нейрондық желі архитектурасының басқа атауы – тарих компрессоры (History Compressor), ол деректерді кеңістікте де, уақытта да қысуы мүмкін, яғни ақпарат алуға мүмкіншілік аз болады. Сондай-ақ, тереңдетілген нейрожүйеде тарих компрессорының үздіксіз нұсқасы бар.

Жойылып бара жатқан градиент проблемасы рекурренттік нейрондық желінің басқа архитектурасын – ұзақ-қысқа мерзімді жады желісін шешуге мүмкіндік береді (Long Short-Term Memory). Мұндай желілерде ұзақ мерзімге мәндерді есте сақтауға мүмкіндік беретін арнайы типтегі тораптар бар.

Қазіргі уақытта тереңдетілген нейрондық желілерді оқыту үшін көптеген бағдарламалық жүйелер құрылған [6]. Олардың ішінде ең танымалдары – Caffe, Theano, TensorFlow, Torch және CNTK. Олардың негізгі сипаттамалары төменгі 1- кестеде келтірілген.

Кесте 1. Тереңдетілген нейрондық желілерді оқытудың бағдарламалық жүйелері

Қасиеті	Caffe	Theano	TensorFlow	Torch	CNTK
Негізгі тіл	C++	Python	C++	Lua	C++
API	C++ Python	Python	C++ Python	Lua Python	C++ C# Python
Көп ядролы CPU	+	+	+	+	+
CPU	+	+	+	+	+
Xeon Phi	+	+	-	-	-
Бөлінген оқыту	+	-	+	+	+
Әзірлеуші	Компьютерлік жүйелеуді көрсету орталығы және Беркли оқыту жүйесі	Монреаль Университеті	Google компаниясы	Ронан Коллаберт	Microsoft
Ашық кодтар	+	+	+	+	+
Желіні оқыту	+	-	+	+	+

Caffe кітапханасы-тереңдетіп оқытудың ең танымал жүйелерінің бірі. Оны компьютерлік жүйелеуді көрсетуде қолданды және Беркли оқыту орталығында (Berkeley Vision and Learning Center) әзірледі, бастапқы кодтары 2014 жылы ашылды. Caffe алдын ала оқытылған модельдерді пайдалануға арналған дайын ең көп әдістер санын қамтиды. Theano жүйесі Канададағы Монреаль университетінде құрылған. Theano – Python бағдарламасында әзірленген, бірақ Python бағдарламасы автоматты түрде c++ бағдарламасына айналады, себебі ол алдымен бағдарламаға сәйкес құрастырылуы керек және содан кейін жүзеге асып орындалады. TensorFlow 2015 жылы Google компаниясымен құрылған және тензорлармен тиімді жұмыс істеуде, сонымен қатар бағанда деректердің ағынды өңдеу жүйесін қамтиды. Torch кітапханасы Lua тілінде негізделген және MATLAB сияқты машиналық оқыту бағдарламаларын жасау үшін ыңғайлы жоғары деңгейлі интерфейсті ұсынады. Жоғары өнімділік Theano сияқты бағдарламалық жүйеде қамтамасыз етіледі және ол тілді біріктіру арқылы жүзеге асады. Torch авторлары C және Lua интеграциясының қарапайымдылығына байланысты Python орнына Lua бағдарламалық жасақтамасын қолдануды жөн көрді. Microsoft компаниясы CNTK (Cognitive Toolkit) жүйесін жасап, 2016 жылы оның бастапқы кодтарын ашты.

Нейрондық желілерді терең оқытудың барлық аталған жүйелері көп ядролы процессорлар, сондай-ақ GPU есептеу үдеткіштері (cuDNN оңтайландырылған кітапханасын қоса алғанда) оқытуды жеделдету үшін пайдалана алады. Сонымен қатар, бағдарламаны қайта жасаудың қажеттілігі жоқ, себебі CPU және GPU-ге параллельдеу жүйесі автоматты түрде орындалады. Caffe және Theano жүйелері Intel Xeon Phi үдеткіштерін қосымша түрде қолдайды, сондай-ақ тереңдетілген нейрондық желілердің оқу уақытын айтарлықтай қысқартуға көмектеседі. Theano-дан басқа барлық дерлік жүйелер есептеуіш кластерлерде нейрондық желілерді тарату үшін пайдалануға болады. Жоғарыда сипатталған жүйелерге қосымша ретінде, терең нейрондық желілерді оқыту үшін ыңғайлы және қарапайым бағдарламалық интерфейсті ұсынатын Keras кітапханасын атап өтуге болады. Keras

дербес жүйе емес, алайда Theano, TensorFlow немесе CNTK секілді бағдарламалық жүйелерде жұмыс істейді. 2016 жылы Keras TensorFlow-дың құрамына кірді.

Жақында құрылған, соған қарамастан танымалдыққа ие жаңа тереңдетілген оқыту кітапханалары да назар аударуға тұрарлық. PaddlePaddle жүйесі (Baidu компаниясымен құрылған) және MXNet жүйесі таратылған кластерлерде терең нейрондық желілерді оқыту үшін арнайы әзірленген. Neon кітапханасы Nervana компаниясымен даярланған. Intel компаниясы Nervana сатып алғаннан кейін Neon – GPU және Intel Xeon Phi үдеткіштерін қолдайтын ең жылдам дамып келе жатқан кітапханалардың біріне айналды, сондай-ақ алдын ала оқытылған нейрон желілерінің көп санына айналды. MXNet және Neon өнімділік сынақтарында жақсы нәтижелер көрсетіп келеді [7].

Нейрондық желі мен терең оқытудың басты айырмашылығы - әртүрлі есептеу тапсырмаларын тез орындау үшін нейрондық желі адамның миындағы нейрондарға ұқсас жұмыс істейді, ал терең оқыту - бұл адамдардың білім алу үшін қолданатын оқу тәсіліне еліктейтін машиналық оқытудың ерекше түрі.

Нейрондық желі күрделі мәселелерді шешу үшін болжамды модельдерді құруға көмектеседі. Екінші жағынан, терең оқыту - машиналық оқытудың бір бөлігі. Бұл сөйлеуді, суретті тануды, табиғи тілді өңдеуді, ұсыну жүйелерін, биоинформатиканы және тағы басқаларды дамытуға көмектеседі. Нейрондық желі - терең оқытуды жүзеге асыру әдісі. Терең оқудан бұрын машинамен оқытуды талқылау керек. Бұл компьютерге нақты бағдарламаланбай-ақ үйренуге мүмкіндік береді. Басқаша айтқанда, бұл деректерді талдау және шешім қабылдау үшін заңдылықтарды тану үшін өздігінен білім алгоритмдерін құруға көмектеседі. Бірақ жалпы шектеулер бар, ол - машинаны жалпы оқыту. Біріншіден, үлкен көлемді деректермен немесе кірістер мен шығулардың үлкен жиынтығымен жұмыс істеу қиын. Мүмкіндіктерді алу қиын болуы мүмкін.

Терең оқыту бұл мәселелерді шешеді. Бұл машиналық оқытудың ерекше түрі. Бұл адамның миына ұқсас жұмыс істей алатын оқыту алгоритмдерін құруға көмектеседі. Терең нейрондық желілер мен қайталанатын нейрондық желілер - бұл тереңдетілген сәулет өнері. Терең нейрондық желі дегеніміз - бірнеше жасырын қабаттары бар нейрондық желі. Қайталанатын нейрондық желілер кірістер тізбегін өңдеу үшін жадты пайдаланады.

Нейрондық желі - бұл әр түрлі есептеу тапсырмаларын тезірек орындау үшін адам миындағы нейрондарға ұқсас жұмыс жасайтын жүйе. Ал терең оқыту дегеніміз - адамның білім алу үшін қолданатын оқыту әдісіне еліктейтін машиналық оқытудың ерекше түрі. Нейрондық желі - терең білім алуға жету әдісі болып табылады. Екінші жағынан, терең оқыту (Deep Learning) - бұл машинаға сүйенудің ерекше түрі. Бұл нейрондық желі мен терең оқытудың басты айырмашылығы. Ал терең оқыту мен машиналық оқытудың басты айырмашылығы - мәліметтер көлемі артқан сайын оның орындалуы. Терең оқу алгоритмдеріне көптеген мәліметтер қажет, сондықтан деректер аз болған кезде алгоритмдер жақсы нәтиже бермейді. Екінші жағынан, машинада оқу алгоритмдері олардың жоғары сапалық принциптерімен женеді.

Терең оқыту машинада жылдам дамып келе жатқан және өркендеген әдіс. Машинаны үйреншікті күйге келтіру кезінде ең қиын мәселелердің бірі - инжиниринг. Үлгілік инженерия модельге берілуі мүмкін қолайлы мүмкіндіктерді шығарумен айналысады. Егер мүмкіндіктер толық емес немесе аз болса, модельде кемшіліктер бар (жоғары жағымсыздық) және егер мүмкіндіктер өте көп болса және олардың барлығы модельдің шығуына ықпал етпесе, онда модель қайтадан кемшілікке ие (жоғары дисперсия). Егер бізде мүмкіндіктер тым көп болса, онда модельде кемшіліктер бар екенін білу үшін өте үлкен мәліметтер қоры қажет. Машиналық оқуда «көрнекілік оқыту» деп аталатын кіші бөлім бар, ол «инженерлерді оқыту» деп аталады, оның көмегімен адам инженерлерінің қолды таңдау мүмкіндіктері қарапайым емес суреттер сияқты мүмкіндіктерден мәліметтерді алуға бағытталған.

Терең оқыту өкілді оқытуға негізделген. Іске асырудың тиімді әдісі ол нейрондық желілердің көптеген қабаттарынан тұрады (қабат саны неғұрлым жоғары болса, модель тереңірек болады), онда әрбір қабат алдыңғы қабаттан кіріп, оны келесі қабатқа өткізеді. Бастапқы қабаттар неғұрлым жалпы және дөрекі сипаттамаларға ие болады және желі тереңдеген сайын мәліметтер жиынтығынан ең жақсы мәліметтерді үйренуге болады, ақырында белгілі бір сенімділік коэффициентімен нәтиже береді. Функционалдылық, олар сүтқоректілердің нейрондарының қалай жұмыс жасайтынынан шабыт алады. Нейрондар кіріс химиялық сигналды қабылдау арқылы жұмыс істейді және белгілі бір шекті мәнге сүйене отырып, сигнал беріледі немесе өшіріледі. Мұндай мінез-құлық нейрондық желілерді жүзеге асыру кезінде әр түрлі математикалық функцияларды (ең кең таралған сигма тәрізді функцияларды) қолдана отырып эмуляцияланады. Терең оқуға негізделген модельдер көбейіп келе жатқанына қарамастан, олар әрқашан жауап бере бермейді. Оларды жақсы орындау үшін жаттығулар туралы көптеген мәліметтер қажет.

Терең нейрондық желілер альтернативті тәсілдерге қарағанда қол жетімді деректердің барлық жиынтығымен жұмыс істей алады. Оқу үрдісінде нейрондық желінің өзі деректердегі қандай белгілер маңызды және қандай белгілер керек болмайтынын анықтап отырады. Жасанды нейрондық желілер адамдардың болжай алмайтын белгілерін болжай алады. Сондықтан терең нейрондық желілердің көмегімен машиналық оқытудың дәстүрлі алгоритмдері орындай алмайтын міндеттерін шеше алады. Дегенмен терең нейрондық желілерді оқыту үлкен есептеу ресурстарын талап етіледі, себебі деректердің үлкен көлемімен жұмыс істеуге тура келеді.

Нейрондық желілер оқыту үлгісінде оқытылады. Оқудың мәні – градиентті түсу әдісімен оңтайландыру мәселесін шешуде әртүрлі қосылыстардың салмағын түзету. Нейрондық желіні меңгеру барысында автоматты түрде экстракция, негізгі ерекшеліктер арасындағы өзара байланыстың маңыздылығын және құрылысын анықтау. Оқытылған жүйке желісі жалпылама қабілеттерге байланысты белгісіз бейнелер бойынша жаттығу кезінде алынған тәжірибені қолдануға мүмкіндік береді деп болжанады.

Қорытынды

Нейрондық желілер пайда болған сәттен бастап олардың архитектурасы мен оқыту әдістерінде көптеген өзгерістер болды. Қазіргі уақытта архитектураның екі түрі басым болып табылады: компьютерлік жүйелеуді көрсету міндеті үшін табысты қолданылатын жинақталған желілері және табиғи тілді өңдеу міндеті үшін белсенді пайдаланылатын рекурренттік желілер. Ерте жинақтау желілері оқытушымен және мұғалімсіз автокодирлеуші мен тереңдетілген сенім желілерін пайдалану жолы арқылы білім алды. Қалдықты оқыту сияқты заманауи әдістер тек мұғаліммен оқытуды қолдануға және алдын ала оқытудан бас тартуға мүмкіндік береді, бұл оқу процесін тездетеді және жеңілдетеді. Сондай-ақ, жинақталған нейрондық желілерді дамытудағы маңызды бағыт–оқытуды жетілдіру (transfer learning). Бұл әдіс басқа есептерді шешу үшін бір деректерде оқытылған нейрондық желілерді пайдалануды көздейді. Бұл ретте желіні оңай күйге келтіру және бізді қызықтыратын тапсырмалардан деректерді жете үйрену қолданылады. Нәтижесінде оқу уақыты қысқарады және алдын ала оқытылған нейрондық желілердің қолдану саласы кеңейтіледі. Сондай-ақ, жинақталған және рекуррентті нейрондық желілерді бірігіп пайдалану перспективті болып табылады.

Табиғи тілді өңдеу және тізбектерді өңдеудің жалпы жағдайы үшін қазіргі уақытта рекуррентті нейрондық желілер қолданылады. Олардың ішінде ең тиімді болып ұзақ-қысқа мерзімді жады желілері және басқарылатын рекуррентті нейрондық желілер болып табылады, өйткені олар қызықтырған оқиғаларды ұзақ уақытқа есте сақтауға мүмкіндік береді. Рекуррентті желілердің қосымша артықшылығы-мұғалімсіз және алдын ала белгіленген деректер жиынтығынсыз оқыту мүмкіндігі ие бола алады. Нейрондық желілерді практикалық қолданудың кең таралуы терең нейрондық желілерді оқыту үшін дайын шешімдердің көп санының, соның ішінде қазіргі заманғы көп ядролы процессорларды, GPU және Intel Xeon Phi есептеуіштерінің үдеткіштерін, сондай-ақ таратылған жады бар есептеуіш кластерлерді пайдалану мүмкіндігінің арқасында көптеген нәтижелерге жету болып табылады.

Пайданылған әдебиеттер тізімі:

- 1 Le Cun Y., Bengio Y., Hinton G. (2015) *Deep Learning Nature*. 436–444. DOI: 10.1038/nature14539.
- 2 Schmidhuber J. *Deep Learning in Neural Networks: an Overview Neural Networks*. 2015. Vol. 1. P. 85–117, DOI: 10.1016/j.neunet.2014.09.003.
- 3 Hinton G.E., Osindero S., Teh Y.-W. (2006) *A Fast Learning Algorithm for Deep Belief Nets Neural Computing*. 1527–1554. DOI: 10.1162/neco.2006.18.7.1527.
- 4 Ciresan D.C., Meier U., Gambardella L.M., Schmidhuber J. *Deep, (2010) Big, Simple Neural Nets for Handwritten Digit Recognition Neural Computation..* 3207–3220. DOI: 10.1162/NECO_a_00052.
- 5 He K., Zhang X., Ren S., et al. (2016) *Deep Residual Learning for Image Recognition IEEE Conference on Computer Vision and Pattern Recognition (Las Vegas, NV, USA, 27–30 June 2016)*, 770–778. DOI: 10.1109/CVPR.2016.90.
- 6 Jia Y., Shelhamer E., Donahue J., et al. (2014) *Caffe: Convolutional Architecture for Fast Feature Embedding Proceedings of the 22nd ACM International Conference on Multimedia (Orlando, FL, USA, November 03–07, 2014)*, 675–678. DOI: 10.1145/2647868.2654889
- 7 Bahrampour S., Ramakrishnan N., Schott L., et al. *Comparative Study of Deep Learning Software Frameworks*. URL: <https://arxiv.org/abs/1511.06435> (дата обращения: 02.07.2017).