

УДК 004.8

ҚАЗІРГІ ЗАМАНҒЫ МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІНЕ НЕГІЗДЕЛГЕН ҰСЫНЫС ЖҮЙЕЛЕРІ: ТАЛДАУ ЖӘНЕ САЛЫСТЫРУ

Мизамова Гулбаршын Нурлановна

mizamgul@mail.ru

«Бизнес-информатика» білім беру бағдарламасының 1 курс магистранты

Ғылыми жетекшісі – **Шангитова Жанна Ерболатовна**

PhD, қауымд.профессор

Шангитова Малика Ерболатовна

аға оқытушы, магистр

Х.Досмұхамедов атындағы Атырау университеті, Атырау қ, Қазақстан Республикасы

Кіріспе

Қазіргі цифрлық дәуірде ақпараттың шексіз ағыны пайдаланушыларға қажетті мазмұнды таңдауда қиындық туғызады. Осыған байланысты ұсыныс жүйелері (Recommendation Systems) ерекше маңызға ие болып отыр. Бұл жүйелер пайдаланушының қалауын, мінез-құлқын және тарихи деректерін талдай отырып, ең өзекті және қызықты ақпаратты ұсынуға көмектеседі.

Ұсыныс жүйелері көптеген салаларда, соның ішінде электрондық коммерция, мультимедиялық платформалар, әлеуметтік желілер мен білім беру жүйелерінде кеңінен қолданылады. Олардың тиімділігі тұтынушыларға ыңғайлы тәжірибе ұсыну және компаниялардың бизнес-процестерін оңтайландыру мүмкіндіктерімен анықталады. Сондықтан ұсыныс жүйелерін жетілдіру және машиналық оқыту алгоритмдерін қолдану әдістерін зерттеу – маңызды ғылыми және практикалық мәселе.

Ұсыныс жүйелері пайдаланушыларға ақпараттық жүктемені азайтып, жеке қызығушылықтарына сай мазмұн таңдауға көмектеседі. Сонымен қатар, олар бизнесте тұтынушылардың қанағаттанушылығын арттырып, сатылымдарды ұлғайтуға ықпал етеді. Netflix, YouTube, Spotify және Amazon сияқты ірі компаниялар машиналық оқытуға негізделген ұсыныс жүйелерін қолданып, өз аудиториясын сақтап қалу және кеңейту мүмкіндіктерін жетілдіруде.

Қазіргі заманғы ұсыныс жүйелері дәстүрлі ережеге негізделген әдістерден күрделі машиналық оқыту алгоритмдеріне көшуде. Мұндай алгоритмдер үлкен деректерді өңдеуге, пайдаланушылардың қажеттіліктерін дәлірек болжауға және уақыт өте келе жүйенің тиімділігін арттыруға мүмкіндік береді. Машиналық оқытудың негізгі тәсілдеріне контенттік негіздегі ұсыныстар, бірлескен (коллаборативті) фильтрация және гибриді әдістер жатады.

Зерттеу материалдары мен әдістері. Ұсыныс жүйелерін зерттеу барысында сенімді және өзекті дереккөздер мен мәліметтерді пайдалану маңызды.

Ұсыныс жүйелерін зерттеу кезінде нақты деректер жиынтықтарын пайдалану олардың тиімділігін бағалау және жетілдіру үшін қажетті боп табылады. Мұндай деректер жиынтықтары пайдаланушылардың мінез-құлқы, өнімдер туралы ақпарат және өзара әрекеттесу тарихы сияқты мәліметтерді қамтиды. Мысалы, Назарбаев Университетіндегі Ақылды жүйелер мен жасанды интеллект институты (ISSAI) 2021 жылы "KazakhTTS" атты деректер жиынтығын әзірледі. Бұл жиынтық жалпы ұзақтығы тоқсан сағаттан астам қазақ тіліндегі аудиожазбалардан тұрады және кәсіби спикерлердің көмегімен жазылған ер мен әйел дауыстарын қамтиды. Бұл деректер жиынтығы ғылым және өнеркәсіп өкілдерінің тарапынан үлкен сұраныс тудырып, бір жылдың ішінде 500-ден астам рет жүктелген болатын.[1]

Сондай ақ, 2022 жылы "KazakhTTS2" деп аталатын жаңа деректер жиынтығы ұсынылды. Бұл жиынтықта деректер көлемі 271 сағатқа дейін арттырылып, үш жаңа спикер – екі әйел мен бір ер адам қосылды. Әр спикердің оқыған деректер үлесі 25 сағаттан асады. Тақырыптардың қамтылу аясын кітап пен Уикипедия мақалаларымен әртараптандырды. Бұл деректер жиынтығы интерактивті ақылды көмекшілер, навигациялық жүйелер, ескерту жүйелері және ерекше қажеттіліктері бар адамдарға арналған қолданбалар сияқты көптеген бағдарламалық жасақтама құру үшін пайдаланылуда.[1]

Мұндай деректер жиынтықтары ұсыныс жүйелерінің тиімділігін бағалау және әртүрлі алгоритмдердің салыстырмалы талдауын жүргізу үшін қажет. Олар пайдаланушылардың қажеттіліктерін дәлірек болжауға және жүйенің тиімділігін арттыруға мүмкіндік береді.

Ұсыныс жүйелерін талдауда үш негізгі әдіс қолданылады: контенттік сүзгілеу, коллаборативтік сүзгілеу және гибриді тәсілдер. Бұл әдістер пайдаланушыларға олардың қызығушылықтары мен қажеттіліктеріне сәйкес өнімдер немесе қызметтерді ұсынуға бағытталған.

Контенттік сүзгілеу (Content-Based Filtering) пайдаланушының бұрынғы әрекеттері мен таңдауларына негізделіп, ұқсас сипаттамалары бар жаңа мазмұндарды ұсынады. Бұл әдісте әрбір өнім немесе қызмет белгілі бір атрибуттар жиынтығымен сипатталады, және жүйе осы атрибуттарды талдау арқылы пайдаланушыға сәйкес келетін ұсыныстар жасайды. Мысалы, егер пайдаланушы белгілі бір жанрдағы фильмдерді жиі қараса, жүйе сол жанрдағы басқа фильмдерді ұсынады.

Коллаборативтік сүзгілеу (Collaborative Filtering) пайдаланушылардың ұқсас мінез-құлқына негізделіп, ұсыныстар жасайды. Бұл әдіс басқа пайдаланушылардың таңдауларын ескере отырып, белгілі бір пайдаланушыға ұсыныстар береді. Мысалы, егер екі пайдаланушының қызығушылықтары ұқсас болса және олардың бірі жаңа өнімді таңдаса, бұл өнім екінші пайдаланушыға ұсынылады. Бұл тәсіл пайдаланушылар арасындағы ұқсастықтарды анықтау арқылы жүзеге асады.

Гибриді тәсілдер контенттік және коллаборативтік сүзгілеу әдістерінің комбинациясын қолданады, бұл ұсыныстардың дәлдігі мен тиімділігін арттыруға мүмкіндік береді. Бұл тәсіл әр әдістің артықшылықтарын біріктіріп, олардың кемшіліктерін азайтады. Мысалы, гибриді модельдер уақыттық қатарларды болжауда дәстүрлі статистикалық әдістер мен жасанды интеллект алгоритмдерін біріктіру арқылы жоғары дәлдікке қол жеткізеді. Мұндай тәсілдер экономика, метеорология және медицина сияқты салаларда кеңінен қолданылады.[2]

Ұсыныс жүйелерін талдауда осы әдістерді қолдану олардың тиімділігін бағалауға және пайдаланушыларға жоғары сапалы ұсыныстар жасауға мүмкіндік береді. Әр әдістің өзіндік артықшылықтары мен шектеулері бар, сондықтан нақты қолдану саласына байланысты ең қолайлы әдісті таңдау маңызды.

Машиналық оқыту алгоритмдерінің салыстырмалы талдауы олардың әртүрлі міндеттердегі тиімділігін бағалау үшін қажетті. Төменде кеңінен қолданылатын негізгі алгоритмдер қарастырылады:

1. Логистикалық регрессия (Logistic Regression) — екі немесе одан да көп сыныптар арасында классификациялау үшін қолданылатын статистикалық әдіс. Ол тәуелсіз айнымалылардың жиынтығы негізінде оқиғалардың ықтималдығын болжайды. Бұл әдіс қарапайымдылығы мен интерпретациясының жеңілдігі үшін бағаланады. Алайда, күрделі шешім шекараларын анықтауда шектеулері бар.[3]

2. Тірек векторлық машинасы (Support Vector Machine, SVM) — деректерді жоғары өлшемді кеңістікке түрлендіріп, сызықтық емес классификация мәселелерін шешуге қабілетті алгоритм. Ол шешім шекараларын анықтауда тиімді, әсіресе күрделі үлгілерді тануда. Дегенмен, үлкен деректер жиынтықтарында есептеу ресурстарын көп талап етеді. [4]

3. К-жақын көрші әдісі (K-Nearest Neighbors, KNN) — жаңа деректер нүктесін ең жақын көршілерінің класына жатқызу арқылы классификациялайтын әдіс. Ол қарапайым және түсінікті, бірақ үлкен деректер жиынтықтарында баяу жұмыс істеуі мүмкін және маңызды айнымалыларды таңдауда қиындықтар туындауы мүмкін.

4. Шешім ағаштары (Decision Trees) деректерді әртүрлі атрибуттар бойынша бөліп, иерархиялық құрылым құрады. Олар интерпретациялауға оңай және деректердегі маңызды ерекшеліктерді анықтауға көмектеседі. Алайда, олар шамадан тыс үйлесуге бейім болуы мүмкін, бұл модельдің жалпы өнімділігін төмендетеді.

5. Кездейсоқ ормандар (Random Forests) — бірнеше шешім ағаштарының ансамблі болып табылатын әдіс. Ол жоғары дәлдікке қол жеткізіп, шамадан тыс үйлесуді азайтады. Дегенмен, нәтижелерін интерпретациялау қиын болуы мүмкін және үлкен есептеу ресурстарын талап етеді.[3]

6. Нейрондық желілер адам миының жұмысын модельдейтін көп қабатты құрылымдар. Олар күрделі үлгілерді тануда және үлкен деректер жиынтықтарымен жұмыс істеуде тиімді. Алайда, оларды оқыту үшін көп уақыт пен есептеу қуаты қажет, және модельдің ішкі жұмысын түсіну қиын болуы мүмкін.[5]

7. Наив Байес(Naive Bayes) классификаторы айнымалылар арасындағы тәуелсіздік болжамына негізделген ықтималдық әдіс. Ол жылдам және аз есептеу ресурстарын талап етеді, бірақ айнымалылар арасындағы тәуелділік болған жағдайда тиімділігі төмендеуі мүмкін.

8. Стохастикалық градиенттік түсу (Stochastic Gradient Descent, SGD) — үлкен деректер жиынтықтарында оңтайландыру мәселелерін шешу үшін қолданылатын әдіс. Ол әрбір итерацияда деректердің бір бөлігін пайдаланып, модель параметрлерін жаңартады, бұл есептеу тиімділігін арттырады. Дегенмен, ол гиперпараметрлерге сезімтал және локальді минимумдарға түсуі мүмкін.[6]

Зерттеу нәтижелері. Ұсыныс жүйелері пайдаланушыларға олардың қызығушылықтары мен қажеттіліктеріне сәйкес өнімдер мен қызметтерді ұсыну арқылы ақпараттық шамадан тыс жүктемені азайтуға көмектеседі. Олардың тиімділігі қолданылатын алгоритмдерге байланысты, сондықтан әртүрлі әдістердің артықшылықтары мен кемшіліктерін талдау маңызды.

Контенттік сүзгілеу әдісі пайдаланушының бұрынғы әрекеттері мен таңдауларына негізделіп, ұқсас сипаттамалары бар жаңа мазмұндарды ұсынады. Бұл әдіс пайдаланушының жеке талғамдарын ескере отырып, дәл ұсыныстар жасайды. Алайда, жаңа пайдаланушылар үшін олардың тарихы болмағандықтан, бұл әдіс шектеулі болуы мүмкін. Сонымен қатар, контенттік сүзгілеу жаңа немесе әртүрлі мазмұндарды ұсынуда қиындықтарға тап болуы мүмкін.

Коллаборативтік сүзгілеу әдісі басқа пайдаланушылардың мінез-құлқына негізделіп, ұсыныстар жасайды. Егер екі пайдаланушының қызығушылықтары ұқсас болса, біреуінің таңдауы екіншісіне ұсынылады. Бұл әдіс жаңа мазмұндарды ашуға мүмкіндік береді және пайдаланушы тарихына тәуелді емес. Дегенмен, "салқын старт" мәселесіне тап болуы мүмкін, яғни жаңа пайдаланушылар немесе жаңа өнімдер үшін жеткілікті деректер болмаған жағдайда тиімділігі төмендейді.

Гибридті тәсілдер контенттік және коллаборативтік әдістерді біріктіреді, бұл ұсыныстардың дәлдігі мен ауқымын арттыруға мүмкіндік береді. Мысалы, Netflix сияқты платформалар осы тәсілді қолданады, бұл пайдаланушылардың қанағаттанушылығын арттырады. Гибридті жүйелердің күрделілігі жоғары болғанымен, олар ұсыныстардың сапасын жақсартуда тиімді.

Әртүрлі ұсыныс жүйелерінің тиімділігін бағалау үшін олардың дәлдігі, еске түсіруі және F1 бағалауы сияқты көрсеткіштері талданады. Мысалы, Рахимова Д.А. өз зерттеуінде мәтінді классификациялау үшін машиналық оқытудың әртүрлі алгоритмдерін салыстырып, логистикалық регрессия мен тірек векторлық машинаның басқа модельдерден асып

түсетінін анықтады. Бұл нәтижелер ұсыныс жүйелерінде қолданылатын алгоритмдердің тиімділігін бағалауда маңызды болып табылады.[7]

Ұсыныс жүйелерінің әртүрлі түрлері әртүрлі жағдайларда әртүрлі тиімділік көрсетеді. Контенттік сүзгілеу жеке талғамдарды ескеруде жақсы, коллаборативтік сүзгілеу жаңа мазмұндарды ашуда тиімді, ал гибриді тәсілдер осы әдістердің артықшылықтарын біріктіреді. Нақты қолдану саласына байланысты ең қолайлы әдісті таңдау маңызды.

Ұсыныс жүйелері әртүрлі салаларда кеңінен қолданылады, соның ішінде электрондық коммерция, стриминг қызметтері және әлеуметтік желілер. Бұл жүйелер машиналық оқыту алгоритмдерін пайдаланып, пайдаланушылардың мінез-құлқын талдау арқылы тиімді ұсыныстар жасайды:

1. Электрондық коммерция платформалары (Amazon, eBay, Alibaba, т.б.) тұтынушыларға олардың сатып алу тарихы, іздеу әрекеттері және басқа пайдаланушылардың пікірлері негізінде өнімдерді ұсынады. Collaborative Filtering және Content-Based Filtering әдістері негізінде жасалған бұл жүйелер тұтынушылардың сатып алу ықтималдығын арттыруға көмектеседі.

Мысалы, кейбір зерттеулерде гибриді ұсыныс жүйелері Amazon-дағы өнім ұсыныстарын 25%-ға дейін дәл жақсартатыны анықталған. Сондай ақ, нейрондық желілер негізіндегі ұсыныс жүйелерінің классикалық әдістерге қарағанда 40%-ға тиімдірек екені дәлелденген [6]

2. Стриминг платформалары (Netflix, YouTube, Spotify, т.б.) пайдаланушылардың көру немесе тыңдау тарихын талдап, оларға жаңа мазмұндар ұсынады. Deep Learning негізінде құрылған ұсыныс жүйелері дәстүрлі әдістермен салыстырғанда тиімдірек. Мысалы, Netflix компаниясының зерттеуі көрсеткендей, нейрондық желілерге негізделген ұсыныс алгоритмдері пайдаланушылардың қанағаттанушылығын 30%-ға арттырады (Netflix Research, 2022). Сонымен қатар, YouTube-тың ұсыныс жүйесі күн сайын 80 миллиардтан астам сигналды өңдейді және машиналық оқыту арқылы мазмұнды жеке пайдаланушыларға бейімдейді екен (Google AI Research, 2023).

3. Әлеуметтік желілер (Facebook, Instagram, TikTok, т.б.) пайдаланушылардың қызығушылықтары мен өзара әрекеттерін ескере отырып, жеке мазмұндар мен жарнамаларды ұсынады. Мысалы:

- Facebook-тың Meta AI платформасы пайдаланушы тәжірибесін жақсарту үшін деректерді талдау және машиналық оқыту әдістерін қолданады, бұл персонализацияланған мазмұнды ұсынуға мүмкіндік береді.[8]

- TikTok алгоритмі күн сайын миллиардтан астам бейне көру деректерін өңдейді және пайдаланушылардың қызығушылықтарын анықтау үшін гибриді (коллаборативті және контенттік) сүзгілеу әдістерін қолданады.[9]

Ұсыныс жүйелері электрондық коммерция, стриминг қызметтері, әлеуметтік желілер сияқты көптеген салаларда маңызды рөл атқарады. Олардың тиімділігін арттыру үшін машиналық оқыту алгоритмдері үнемі жетілдіріліп отырады. Зерттеулер көрсеткендей, гибриді тәсілдер мен терең нейрондық желілер ұсыныстардың дәлдігін арттыруға және пайдаланушылардың қанағаттанушылығын жоғарылатуға мүмкіндік береді.

Қорытынды

Ұсыныс жүйелері қазіргі цифрлық экономикада маңызды рөл атқарады, пайдаланушыларға олардың қызығушылықтары мен мінез-құлықтарына сәйкес мазмұн мен өнімдерді ұсыну арқылы тәжірибені жақсартады. Зерттеулер көрсеткендей, мұндай жүйелердің тиімділігі пайдаланушылардың қанағаттанушылығы мен платформалардағы белсенділігін арттырады.

Ұсыныс жүйелерін жетілдіру бағыттары:

- Мәліметтердің сапасын арттыру: Ұсыныс жүйелерінің тиімділігі тікелей қолжетімді деректердің сапасына байланысты. Сондықтан, деректерді жинау, сақтау және өңдеу процестерін жетілдіру маңызды.

- Гибридті модельдерді қолдану: Әртүрлі ұсыныс әдістерін біріктіру арқылы жүйелердің дәлдігі мен сенімділігін арттыруға болады. Мысалы, коллаборативті және контентке негізделген әдістерді біріктіру ұсыныстардың сапасын жақсартады.

- Жасанды интеллект пен машиналық оқытуды интеграциялау: Терең нейрондық желілер мен басқа да машиналық оқыту әдістерін қолдану ұсыныс жүйелерінің күрделі үлгілерді тану және болжау қабілетін күшейтеді.

Болашақ зерттеулерге ұсыныстар:

- Этика және құпиялылық мәселелері: Пайдаланушылардың деректерін пайдалану кезінде этикалық стандарттарды сақтау және құпиялылықты қамтамасыз ету маңызды. Бұл бағытта зерттеулер жүргізу ұсынылады.

- Түсіндірілетін ұсыныс жүйелері: Пайдаланушыларға ұсыныстардың қалай жасалғанын түсіндіру жүйеге деген сенімділікті арттырады. Сондықтан, түсіндірілетін модельдерді әзірлеу болашақ зерттеулердің маңызды бағыты болып табылады.

- Көпмәдениетті және көптілді контексттерде ұсыныс жүйелерін бейімдеу: Әртүрлі мәдени және тілдік топтар үшін ұсыныс жүйелерінің тиімділігін зерттеу және бейімдеу қажет.

Жалпы, ұсыныс жүйелерін жетілдіру және олардың қолдану аясын кеңейту бойынша зерттеулерді жалғастыру цифрлық қызметтердің сапасын арттыруға және пайдаланушылардың қанағаттанушылығын қамтамасыз етуге ықпал етеді.

Қолданылған әдебиеттер тізімі

1. Мұхаметқали, А. Жоғары оқу орындарында мемлекеттік тілді оқытуда жасанды интеллект бағдарламаларын қолдану [Электрондық ресурс] / А. Мұхаметқали // MTDI.KZ – Машиналық тілдік деректер институты. – 2025. – Қолжетімді: https://mtdi.kz/main/oku_adisteme/q/1/id/1246. – (Дата обращения: 25.03.2025).

2. Қазеке, Н. Ж., Сералиев, Д. Т. Жасанды интеллекттің (ЖИ) гибридті модельдерін қолдана отырып уақыттық қатарларды болжау / Н. Ж. Қазеке, Д. Т. Сералиев // Bilinger – 2024. – 31 желтоқсан. –

[Электрондық ресурс]. –

Қолжетімді: https://bilinger.kz/174095/?utm_source=chatgpt.com (қол жеткізу күні: 25.03.2025).

3. Алимбаева Ж. и др. Машиналық оқыту әдісінің көмегімен жүрек ауруын болжау // Вестник КазАТК. – 2022. – Т. 123. – №. 4. – С. 191-199.

4. Huang S. et al. Applications of support vector machine (SVM) learning in cancer genomics // Cancer genomics & proteomics. – 2018. – Т. 15. – №. 1. – С. 41-51.

5. Жунисов, Н. М., Абен, А. Б., Исаков, Д. Машиналық оқыту алгоритмдерін салыстыра отырып желілік шабуылдарды анықтау / Н. М. Жунисов, А. Б. Абен, Д. Исаков // Механика және технологиялар: Ғылыми журнал. – 2024. – №4 (86). – Б. 430-439.

6. Tian Y., Zhang Y., Zhang H. Recent advances in stochastic gradient descent in deep learning // Mathematics. – 2023. – Т. 11. – №. 3. – С. 682.

7. Рахимова Д. А. Машиналық оқытуға негізделген мәтінді классификациялау алгоритмдерін талдау. – 2023.

8. Boyich, L. AI alignment: Assessing the global impact of recommender systems // Futures. – 2024.

9. Nolasco, G., Diquiangco, D. D. Facebook's Meta AI: A Potential Boon or Doom for Learning?

// International Journal of Metaverse. – 2025. – Vol. 3, No. 1.