



КОМПЬЮТЕРНЫЕ НАУКИ, ПРИБОРОСТРОЕНИЕ И АВТОМАТИЗАЦИЯ
КОМПЬЮТЕРЛІК ҒЫЛЫМДАР, АСПАП ЖАСАУ ЖӘНЕ АВТОМАТТАНДЫРУ
COMPUTER SCIENCE, INSTRUMENTATION AND AUTOMATION

ӘОЖ 004.8

ҒТАХА 28.23.00

https://doi.org/10.53364/24138614_2026_42_3_15

Д.С. Абраймова¹, С.А. Нургалиева^{2*}, Ж.О. Башеева², Ш.С. Туткышбаева²,
 А.Р.Сыздықпаева³

¹С.Сейфуллин атындағы Қазақ агротехникалық зерттеу университеті, Астана, Қазақстан

²Astana IT university, Астана, Қазақстан

³С.Аманжолов атындағы Шығыс Қазақстан университеті, Өскемен, Қазақстан

*E-mail: symbat.nurgaliyeva@astanait.edu.kz

ЖАСАНДЫ ИНТЕЛЛЕКТ НЕГІЗІНДЕГІ МАНСАПТЫҚ ҰСЫНЫМ
ЖҮЙЕЛЕРІНІҢ ӨНІМДІЛІГІ: ЖҮЙЕЛІ ШОЛУ

***Аңдатпа.** Еңбек нарығының жылдам өзгеруі және білім беру деректерінің қолжетімділігінің артуы жағдайында жасанды интеллект (ЖИ) негізіндегі мансаптық ұсыным жүйелерін әзірлеу өзекті бағыттардың біріне айналды. Алайда деректер түрлері мен алгоритмдерді таңдау ерекшеліктерінің модельдердің болжамдық дәлдігі мен есептеу тиімділігіне әсері жеткілікті деңгейде жүйеленбеген. Осы зерттеудің мақсаты – ЖИ негізіндегі мансаптық ұсыным жүйелеріне арналған заманауи зерттеулерді жүйелеу және деректер түрлері, алгоритмдер мен бағалау көрсеткіштері арасындағы байланысты анықтау. Зерттеу PRISMA 2020 әдіснамасы негізінде жүргізілген жүйелі шолуға сүйенеді. Талдау «деректер түрі-алгоритм-бағалау көрсеткіші» байланысы тұрғысынан орындалды. Нәтижелер қарастырылған зерттеулерде жіктеу міндеттерінің басым екенін (56%) және құрылымдалған кестелік деректердің ең жиі қолданылатынын (76%) көрсетті. Ең жоғары болжамдық дәлдік негізінен ансамбльдік және нейрондық желі модельдерінде байқалды. Сонымен қатар зерттеулердің көпшілігінде есептеу тиімділігі жүйелі түрде бааланбайтыны анықталды: жұмыстардың 76%-ында бұл көрсеткіштер қарастырылмаған. Бұдан басқа, алгоритмдерді таңдау көбіне пайдаланылатын деректер түрлерімен байланысты екені көрсетілді. Зерттеу нәтижелері мансаптық ұсыным жүйелеріндегі алгоритмдік тәсілдерді жүйелеуге мүмкіндік береді және болашақ зерттеулер үшін бағалау хаттамаларын стандарттау, есептеу тиімділігін жүйелі есепке алу және мультимодальды деректерді кеңінен пайдалану қажеттігін көрсетеді. Сонымен қатар ұлттық еңбек нарығы мен білім беру ерекшеліктеріне бейімделген ЖИ негізіндегі ұсыным жүйелерін дамыту өзекті бағыт болып табылады.*

***Түйін сөздер:** жасанды интеллект, мансаптық ұсыным жүйесі, деректер түрлері, машиналық оқыту алгоритмдері, болжамдық дәлдік, есептеу тиімділігі.*

Кіріспе.

Қазіргі таңда жасанды интеллект (ЖИ) негізіндегі жүйелер білім беру және еңбек нарығы салаларында кеңінен қолданылып, жеке тұлғаның ерекшеліктері мен әртүрлі деректер көздерін ескере отырып, оқыту және кәсіби бағдар беру бойынша дербестендірілген ұсынымдар жасауға мүмкіндік беруде. Дәстүрлі мансаптық бағдарлау құралдарымен салыстырғанда, ЖИ жүйелері анағұрлым дәл және бейімделген шешімдер ұсынғанымен, кәсіп таңдау сияқты маңызды өмірлік шешімдерде олардың болжамдық

дәлдігімен қатар алгоритмдердің ашықтығы, түсіндірілетіндігі және тиімділігін бағалау мәселелері де ерекше маңызға ие.

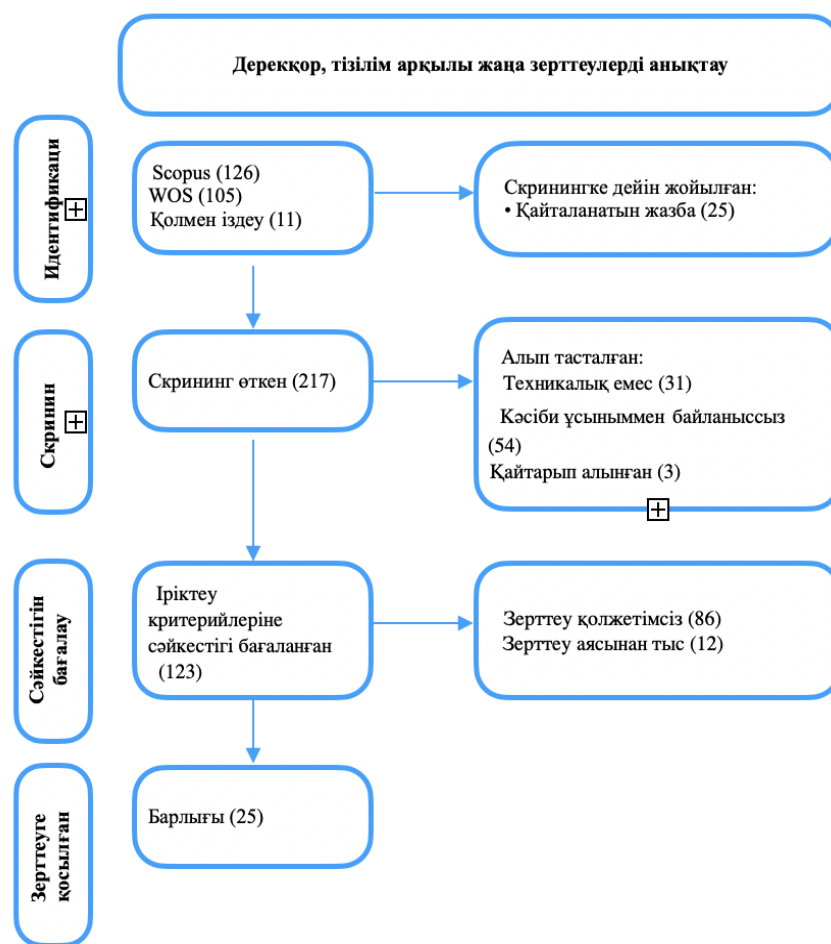
Қазіргі уақытта жасанды интеллект негізіндегі мансаптық бағдарлау жүйелеріне арналған зерттеулер алгоритмдік тәсілдердің кең ауқымын қамтиды. Олардың қатарында бақылаулы оқытудың дәстүрлі әдістері, ұсынымдық жүйелер және гибридтік модельдер бар. Мұндай әртүрлілік мансаптық ұсынымдар саласына арналған жүйелі шолу зерттеулерінде де көрсетілген [1]. Қолданбалы зерттеулерде білім беру траекториясын болжау мен мансаптық ұсынымдар қалыптастыру үшін классикалық машиналық оқыту модельдерімен қатар құрылымдалмаған мәтіндерді өңдеуге негізделген заманауи NLP тәсілдері де кеңінен қолданылады [2]. Дегенмен, зерттеулерде әртүрлі міндет қойылымдары мен бағалау көрсеткіштері пайдаланылатындықтан, алынған нәтижелерді бірыңғай дәлелдік қор ретінде қарастыру қиын [3]. Осыған байланысты алгоритмдердің тиімділігін тікелей салыстыру әдіснамалық тұрғыдан әрдайым дұрыс бола бермейді [4]. Сонымен қатар нәтижелерді ұсыну тәсілдерінің біркелкі болмауы, әсіресе ұсынымдық және гибридтік жүйелерде қолданылатын көрсеткіштердің айырмашылығы, сандық жинақтауды шектейді. Мұндай әдіснамалық әртектілік зерттеулерді салыстыруды күрделендіріп, сипаттамалық синтез жүргізудің қажеттілігін арттырады. Сондықтан бұл жұмыстың негізгі мақсаты қолданыстағы тәсілдерді сипаттау ғана емес, жоғары әдіснамалық жағдайында олардың нәтижелерін жүйелі түрде синтездеу болып табылады.

Бұл зерттеудің мақсаты – жасанды интеллект негізіндегі мансаптық ұсыным жүйелеріне арналған ғылыми еңбектерді жүйелі түрде талдау арқылы әртүрлі міндет қойылымдары аясында ең жоғары болжамдық дәлдік пен тиімділікті көрсететін алгоритмдерді анықтау. Зерттеудің негізгі үлестері: (1) мансаптық бағдарлау саласындағы ЖИ-зерттеулердің деректер базасын қалыптастыру және оларды міндеттері, алгоритмдері мен бағалау көрсеткіштері бойынша жіктеу; (2) алгоритмдерді салыстырмалы талдауға мүмкіндік беретін кластерлеу тәсілін ұсыну; (3) алынған нәтижелерді жүйелі синтездеу; (4) зерттеулердегі шектеулер мен олқылықтарды анықтап, болашақ зерттеулерге арналған ұсыныстарды тұжырымдау. Осы мақсатқа жету үшін келесі зерттеу сұрақтары қарастырылды:

1. ЖИ негізіндегі мансаптық ұсыным жүйелерінде машиналық және терең оқытудың қандай алгоритмдері ең жоғары болжамдық дәлдікті көрсетеді?
2. Машиналық және терең оқыту алгоритмдерінің қайсысы есептеу тиімділігі бойынша жоғары нәтижелерге қол жеткізеді?
3. Мансаптық ұсыным модельдерінде қандай деректер түрлері пайдаланылады және олардың алгоритмдерді таңдау мен өнімділік көрсеткіштеріне әсері қандай?

Материалдар мен зерттеу әдістері.

Жүйелі әдебиет шолуы PRISMA 2020 нұсқаулығына сәйкес жүргізілді [5]. Зерттеу әдістемесі зерттеу сұрақтарын қалыптастыру, іздеу стратегиясын әзірлеу, зерттеулерді іріктеу, сапасын бағалау және нәтижелерді синтездеу кезеңдерін қамтиды. Әдебиеттерді іздеу Scopus және Web of Science деректер қорларында жүзеге асырылды. Зерттеулерді іріктеу екі кезеңнен тұрды: бастапқы кезеңде жарияланымдардың атаулары мен аңдатпалары қарастырылса, кейінгі кезеңде толық мәтіндер талданды. Алып тастау себептері PRISMA талаптарына сәйкес құжатталды (1-сурет).



Сурет 1 – Зерттеулерді іріктеу үдерісінің сызбасы (2021–2025)

Іздеу стратегиясы және деректер көздері. Зерттеу барысында әрбір енгізілген жұмыс үшін «деректер түрі – алгоритм – бағалау көрсеткіші» арасындағы байланыс тіркелді. Мұндай тәсіл нәтижелерді нақты міндеттер контекстінде талдауға және зерттеулердің әдіснамалық әртекстілігі жағдайында деректерді жүйелі түрде синтездеуге мүмкіндік берді.

Ғылыми жарияланымдарды іздеу алдын ала әзірленген стратегия негізінде Scopus және Web of Science деректер қорларында жүргізілді. Іздеу кезеңі 2021 жылдан 2025 жылдың қарашасына дейінгі уақыт аралығын қамтыды. Ғылыми жарияланымдарды іздеу барысында жасанды интеллект технологияларын сипаттайтын «artificial intelligence», «AI», «machine learning», «deep learning», «neural network», «recommender system» және «decision support system» сияқты кілттік сөздер пайдаланылды. Мансаптық бағдарлау саласын қамту үшін «career guidance», «career recommendation», «career counseling», «career planning» және «career development» терминдері қолданылды. Іздеу сұраныстары аталған терминдердің әртүрлі комбинациялары негізінде қалыптастырылды. Іріктеу Computer Science және Engineering пәндік салаларындағы ағылшын тілінде жарияланған мақалалар мен конференция материалдарымен шектелді. Негізгі іздеуден бөлек, толық мәтіндік сараптаудан өткен мақалалардың әдебиеттер тізімдерін талдау арқылы қосымша қолмен іздеу жүргізілді. Бұл тәсіл бастапқы іздеу барысында анықталмаған, бірақ зерттеу тақырыбына сәйкес келетін жарияланымдарды қамтуға мүмкіндік берді.

Қосу және алып тастау өлшемшарттары. Шолуға мансаптық ұсыным жүйелерінде жасанды интеллект алгоритмдерін қолданатын және пайдаланылған әдістер, деректер жиынтықтары немесе бағалау көрсеткіштері туралы ақпарат беретін зерттеулер енгізілді. Ал ЖИ әдістерін қолданбайтын педагогикалық немесе психологиялық зерттеулер,

мансаптық ұсыным жүйелерімен байланысты емес жарияланымдар, сондай-ақ шолу мақалалары, кері қайтарылған жұмыстар, толық мәтіні қолжетімсіз материалдар және 2021 жылға дейін жарияланған зерттеулер талдаудан шығарылды.

Сүзгілеу және іріктеу процесі. Зерттеулерді іріктеу PRISMA-2020 ұсынымдарына сәйкес бірнеше кезеңде жүргізілді [5]. Алдымен қайталанатын жарияланымдар алынып тасталды. Кейін жұмыстардың атаулары мен аңдатпалары қосу және алып тастау өлшемшарттарына сәйкестігі тұрғысынан бағаланды. Соңғы кезеңде толық мәтіндерге талдау жүргізіліп, олардың зерттеу сұрақтарына және әдіснамалық сапа талаптарына сәйкестігі тексерілді.

Деректерді алу. Деректерді жинау зерттеу сұрақтарына бағытталған арнайы әзірленген хаттама негізінде жүргізілді. Бұл зерттеуде алынатын айнымалылар жиынтығы болжам дәлдігін, есептеу тиімділігін және қолданылған деректер түрлерін талдауға бағдарланды. Құрылымдық талдау жүргізу мақсатында әрбір енгізілген зерттеуге бірегей код (S01–S25) берілді. Деректерді алу жартылай құрылымдалған кесте арқылы жүзеге асырылып, барлық ақпарат кейінгі талдау үшін жинақтаушы деректер қорына енгізілді. Алынған деректердің қорытынды құрылымы 1-кестеде көрсетілген.

Кесте 1 – Мансаптық ұсыным жүйелері бойынша зерттеулердің алгоритмдік кластерлер бойынша жіктелуі

Зерттеу коды	Библиографиялық сипаттама	Кластер
S01	Islam R., Keya K., Zeng Z., Pan S., Foulds J. (2021). Debiasing Career Recommendations with Neural Fair Collaborative Filtering	Кластер 4
S02	Mejia M., Jimenez C., Martínez-Santos J. (2021). Career Recommendation System for Validation of Multiple Intelligence to High School Students	Кластер 1
S03	Guo P., Xiao K., Ye Z., Zhu H., Zhu W. (2022). Intelligent Career Planning via Stochastic Subsampling Reinforcement Learning	Кластер 5
S04	Huang L. (2022). The Establishment of College Student Employment Guidance System Integrating Artificial Intelligence and Civic Education	Кластер 3
S05	Wang C., Wang K., et al. (2022). Do Humans Prefer Debaised AI Algorithms? A Case Study in Career Recommendation	Кластер 4
S06	Guo J., Qi L. (2022). Visualization Research of College Students' Career Planning Paths Integrating Deep Learning and Big Data	Кластер 2
S07	Fatima N., Gul S., Ahmed J., Khand Z., Mujtaba G. (2022). A Rule-Based Machine Learning Model for Career Selection through MBTI Personality	Кластер 1
S08	Wang Y., Yang L., Wu J., Song Z., Shi L. (2022). Mining Campus Big Data: Prediction of Career Choice Using Interpretable Machine Learning Method	Кластер 1
S09	Majjate H., Bellarhmouch Y., Jeghal A., Yahyaouy A., Tairi H., Zidani K. (2024). AI-Powered Academic Guidance and Counseling System Based on Student Profile and Interests	Кластер 1
S10	Byeon H. (2024). A Feature Interaction Based Neural Network Approach: Predicting Job Turnover in Early Career Graduates in South Korea	Кластер 2
S11	Gülten H., Baraçlı H. (2024). A Machine Learning-Based Forecast Model for Career Planning in Human Resource Management: A Case Study of the Turkish Post Corporation	Кластер 1

S12	Abdalkareem M., Min-Allah N. (2024). Explainable Models for Predicting Academic Pathways for High School Students in Saudi Arabia	Кластер 1
S13	Liu P. (2024). Research on the Path of Enhancing Employment and Entrepreneurship Ability of Deaf College Students Based on Knowledge Graph	Кластер 4
S14	Feng K. (2024). Research on College Students' Career Planning and Employment Guidance Strategies Assisted by Artificial Intelligence	Кластер 4
S15	Chen Y., Wei W., Wang L., Dong Y., Liang C. (2024). Where Do They Go Next? Causal Inference-Based Prediction and Visual Analysis of Graduates' First Destination	Кластер 3
S16	Mahboob K., Asif R., Haider N. (2024). Career Planning Matters: Intelligence-Based Career Path Predictions Using Data Mining Models – A Longitudinal Study	Кластер 2
S17	Seesukong S., Angskun T., Keandoungchun N., Thippongton A., Angskun J. (2024). CARIA: A Personalized Career Recommender Based on Individual Competency Similarity Measure	Кластер 5
S18	Zhang L., Gu X. (2025). A Novel Career Prediction Method Based on Fuzzy Model–Fuzzy Clustering Approach	Кластер 5
S19	Agrawal R., Nayak A., Hemnani P., et al. (2025). Predicting AI Job Market Dynamics: A Data Mining Approach to Machine Learning Career Trends on Glassdoor	Кластер 1
S20	Moreno-Novoa M., Puertas E., Martinez-Santos J. (2024). Enhancing Vocational Guidance with Machine Learning: Predicting STEM Career Viability for High School Students	Кластер 1
S21	Cheng Y. (2025). Development of a Career Planning Assistance Platform for College Students Combining Fuzzy Logic and Deep Learning	Кластер 3
S22	El-Keiey S., ElMenshaway D., Hassanein E. (2025). Career Recommendation Based on Feature Selection for Undergraduate Students Using Machine Learning Techniques	Кластер 1
S23	Hassan R., Hassan M., Sameem M., Rafique M. (2024). Personality-Aware Course Recommender System Using Deep Learning for Technical and Vocational Education and Training	Кластер 2
S24	Senger E., Campbell Y., van der Goot R., Plank B. (2025). Toward More Realistic Career Path Prediction: Evaluation and Methods	Кластер 2
S25	Faruque S.H., Sharun K., Akter S. (2025). Decision Support System to Reveal Future Career over Students' Survey Using Explainable AI	Кластер 2

Нәтижелер және оларды талқылау.

Енгізілген зерттеулердің әдіснамалық сапасын бағалау (Quality Assessment (QA)) үшін зерттеу мақсаттарына бейімделген төрт критерийден тұратын арнайы сапаны бағалау сызбасы қолданылды [6]. Бағалау нәтижелері 2-кестеде көрсетілген.

Кесте 2 – Енгізілген зерттеулердің сапасын бағалау нәтижелері

Зерттеу коды	QA1	QA2	QA3	QA4	Жалпы балл
S01	1	1	1	1	4
S02	1	0	1	1	3
S03	1	1	0	1	3

S04	1	0	1	1	3
S05	1	0	1	1	3
S06	1	1	1	1	4
S07	1	0	1	1	3
S08	1	1	1	1	4
S09	1	0	1	1	3
S10	1	1	1	1	4
S11	1	1	1	1	4
S12	1	1	1	1	4
S13	1	1	1	1	4
S14	1	1	1	1	4
S15	1	0	1	1	3
S16	1	0	1	1	3
S17	1	0	1	1	3
S18	1	1	1	1	4
S19	1	0	1	1	3
S20	1	0	1	1	3
S21	1	0	1	1	3
S22	1	0	1	1	3
S23	1	0	1	1	3
S24	1	1	1	1	4
S25	1	1	1	1	4

Ескерту:

QA1 – алгоритмнің сипатталуы;

QA2 – есептеу тиімділігі туралы ақпараттың болуы;

QA3 – болжам дәлдігі туралы ақпараттың болуы;

QA4 – деректер түрінің сипатталуы.

Жоғарыдағы кестеде көрсетілгендей, сапаны бағалау нәтижелері бойынша зерттеулердің 48%-ы (n=12) барлық төрт критерийге сәйкес келіп, жоғары сапа санатына жатқызылды. Тағы 48%-ы (n=12) орташа сапа деңгейін көрсетті. Тек бір зерттеу (4%, S03) төмен балл алды, бұл онда болжам дәлдігіне қатысты ақпараттың болмауымен байланысты. Алайда бұл нәтиже зерттеудің жалпы ғылыми сапасын емес, оның осы жұмыста қолданылған сапаны бағалау критерийлеріне сәйкестік деңгейін ғана сипаттайды.

Зерттеу нәтижелері «деректер түрі – алгоритм – бағалау көрсеткіші» байланысы негізінде ұсынылды, себебі модельдердің өнімділігі қолданылған деректердің сипаты мен зерттеу міндеттеріне тікелей тәуелді.

Зерттеулердің сипаттамалары. Жарияланымдардың географиялық таралуы жасанды интеллект негізіндегі мансаптық кеңес беру саласындағы зерттеулердің жекелеген елдерде шоғырланғанын көрсетті. Ең көп зерттеу Қытайда жүргізілген (n=11), бұл елдің аталған бағыттағы жетекші орнын айғақтайды. АҚШ, Германия, Колумбия, Пәкістан және Мароккодан екі зерттеуден анықталса, Бангладеш, Үндістан, Тайланд, Сауд Арабиясы, Түркия, Оңтүстік Корея, Италия, Ұлыбритания және Египеттен бір-бір жарияланым ғана табылды. Ал шолуға енгізілген зерттеулер арасында Қазақстан авторларының жұмыстары кездеспеді. Бұл жағдай елімізде ЖИ негізіндегі мансаптық ұсыным жүйелері бойынша зерттеулердің жеткіліксіз дамығанын көрсетеді және білім беру жүйесі мен еңбек нарығының цифрлық трансформациясы жағдайында маңызды зерттеу олқылығы ретінде қарастырылуы мүмкін.

Мазмұндық талдаудың тақырыптық нәтижелері. Бұл бөлімде зерттеулердің негізгі сипаттамалары мен зерттеу сұрақтарына сәйкес алынған нәтижелер ұсынылады. Іріктелген

жұмыстарда қолданылған алгоритмдер оқыту парадигмаларының ерекшеліктеріне қарай бес тұжырымдамалық кластерге жіктелді. Мұндай топтастыру зерттеулердің жалпы корпусын жүйелеуге және алгоритмдерді салыстырмалы түрде талдауға мүмкіндік берді (3-кесте).

Кесте 3 – Мансаптық ұсыным жүйелерінде қолданылатын алгоритмдердің оқыту парадигмалары бойынша жіктелуі

Кластер нөмері	Кластер атауы	Ең жоғары нәтижені көрсеткен алгоритмдер
Кластер 1	Classical & Ensemble Machine Learning	Logistic Regression, SVM, KNN, Huber Regressor Random Forest, Gradient Boosting, XGBoost
Кластер 2	Deep Neural Networks	MLP, CNN, LSTM, FINN, TVET-CRS model
Кластер 3	Hybrid & Multimodal Deep Learning	GRU+MLP+attention, CIRF-MLP, CNN+LSTM+Transformer
Кластер 4	ML-based Recommender Systems	Feature-based & Time-aware Hybrid Collaborative Filtering Neural CF NFCF Hybrid CF + KG + clustering + TF-IDF + time-decay
Кластер 5	Non-comparable AI Paradigms	Reinforcement Learning–based methods (SSSRL c Q-learning) Fuzzy Logic & Soft Computing Systems (Takagi–Sugeno type + FCM + LSE) Similarity-based / Distance-based Methods (MES)

Іріктелген зерттеулерді талдау барысында әрбір жұмыста авторлар ұсынған әдістеме шеңберінде ең жоғары нәтижеге қол жеткізген алгоритмдер анықталды. Алайда бұл нәтижелер нақты деректер жиынтықтары мен зерттеу жағдайларына тәуелді болғандықтан, белгілі бір алгоритмнің әмбебап артықшылығын білдірмейді.

Талдау зерттеу сұрақтарына сәйкес үш бағыт бойынша жүргізілді: (1) шешілетін міндеттер мен қолданылған бағалау көрсеткіштері; (2) алгоритмдердің есептеу тиімділігі; (3) кіріс деректерінің түрлері және олардың алгоритмдерді таңдау мен өнімділік нәтижелерімен байланысы. Нәтижелер міндеттер мен бағалау көрсеткіштері бойынша ұсынылса, кластерлер арасындағы ұқсастықтар мен айырмашылықтар талқылау бөлімінде қарастырылды.

Бірінші зерттеу сұрағына сәйкес, зерттеулерде ұсынылған алгоритмдердің болжамдық дәлдік көрсеткіштері салыстырылып, әр жұмыста ең жоғары нәтиже көрсеткен модельдер анықталды. Алынған нәтижелер алгоритмдік кластерлер бойынша жүйеленді:

а) Жіктеу. Жіктеу міндеттерінде ең жоғары нәтижелерді ансамбльдік машиналық оқыту әдістері мен терең оқыту модельдері көрсетті. Кластер 1 зерттеулерінде Random Forest моделі Accuracy = 99% және AUC = 99.97% (S12), ал Gradient Boosting Accuracy = 0.978 және AUC = 0.983 (S11) мәндеріне қол жеткізді. Терең оқыту модельдері арасында CNN Mean Accuracy = 93.75% және AUC-ROC = 99.02% (S25), ал MLP моделі 95.73% дәлдік көрсетті (S16) (Кластер 2). Сонымен қатар гибридік тәсілдер де жоғары нәтижелер берді: CIRF-MLP моделі Accuracy = 0.87 және AUC = 0.95 (S15) (Кластер 3) көрсетсе, Feature-based & Time-aware Hybrid Collaborative Filtering моделі AUC көрсеткіші бойынша 95%-дан жоғары нәтижеге жетті (S14) (Кластер 4). Жалпы алғанда, жіктеу міндеттерінде ансамбльдік әдістер мен нейрондық желілер ең жоғары болжамдық дәлдікті көрсетті.

б) Ұсынымдарды дәрежелену. Ұсынымдарды дәрежелену міндеттерінде зерттеулер әртүрлі бағалау көрсеткіштерін қолданғандықтан нәтижелерді тікелей салыстыру қиын болды. Ең жиі кездескен көрсеткіштердің бірі NDCG болып табылады. Осы көрсеткіш бойынша TVET-CRS моделі $NDCG = 0.96$ нәтижесін көрсетті (S23). NCF модельі үшін NDCG шамамен 0.48 (S01), ал NCF+LR моделі үшін $NDCG = 0.0050$ (S05) болды. Басқа зерттеулерде Precision@K (S17), HR және NDCG (S01), сондай-ақ MRR және Recall@K (S24) көрсеткіштері пайдаланылды. Сондықтан бұл санаттағы нәтижелер көрсеткіштердің әртүрлілігіне байланысты сипаттамалық түрде жинақталды.

в) Регрессия міндеттері. Регрессиялық міндеттерде модельдердің тиімділігі қателік көрсеткіштері мен детерминация коэффициенті арқылы бағаланды. Ең төмен қателік көрсеткіштерінің бірі Huber Regressor моделінде байқалды ($MSE = 0.0017$, $RMSE = 0.0422$) (S09). Random Forest моделі $RMSE = 0.0646$ және $NRMSE = 0.1966$ мәндерін көрсетті (S19). Ал гибридік ұсынымдық тәсіл үшін $MAE = 0.591$ нәтижесі тіркелді (S13) (Кластер 4). Бұл нәтижелер регрессиялық міндеттерде ансамбльдік модельдердің жоғары тиімділігін көрсетеді.

г) Кластерлеу. Кластерлеу міндетіне арналған жалғыз зерттеуде LSTM-Canopy алгоритмі k-means әдісімен салыстырылды (S06). Нәтижелер бойынша DS3 деректер жиынтығында LSTM-Canopy Accuracy = 84.30% мәніне жетіп, Emin көрсеткіші бойынша дәстүрлі k-means алгоритмінен жақсы нәтиже көрсетті.

Бірінші зерттеу сұрағы нәтижелері бойынша ең жоғары болжамдық дәлдік көрсеткіштері негізінен ансамбльдік машиналық оқыту әдістерінде (Random Forest, Gradient Boosting) және терең оқыту модельдерінде (CNN, MLP) байқалды. Алайда алгоритмдердің артықшылығы зерттеу міндетіне, қолданылған деректерге және бағалау көрсеткіштеріне тәуелді болды, сондықтан барлық жағдай үшін әмбебап ең үздік модель анықталған жоқ.

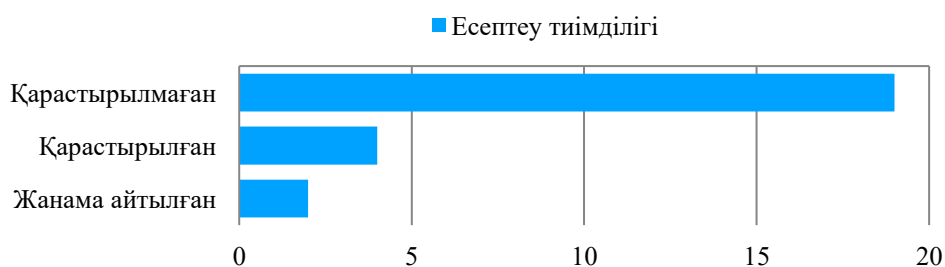
Екінші зерттеу сұрағы бойынша машиналық және терең оқыту алгоритмдерінің есептеу тиімділігі талданды. Есептеу тиімділігі туралы ақпарат енгізілген 25 зерттеудің аз бөлігінде ғана қарастырылған. Зерттеулерде бұл мәселе архитектуралық шешімдер, алгоритмдік күрделілік, гиперпараметрлерді оңтайландыру, оқыту параметрлері және болжау кідірісі тұрғысынан бағаланды. Алайда жұмыстардың көпшілігінде есептеу тиімділігін сипаттайтын сандық көрсеткіштер берілмеген.

Бірқатар зерттеулерде есептеу шығындарын азайтуға бағытталған тәсілдер ұсынылған. Мысалы, (S03) зерттеуінде нейрондық желілерді негізгі есептеу циклінен шығару арқылы жүйенің масштабталуы мен тиімділігі арттырылған. (S08) жұмысында алгоритмнің есептеу күрделілігі $O((N \log N) - (\log^2 k))$ аралығында бағаланып, TPE негізіндегі гиперпараметрлерді оңтайландыру оқыту тиімділігін жақсартатыны көрсетілген. (S11) зерттеуінде Gradient Boosting және Random Forest модельдерінің тиімділігі оңтайлы гиперпараметрлерді таңдау арқылы қамтамасыз етілетіні айтылған.

Есептеу тиімділігі бойынша нақты сандық нәтижелер шектеулі болды. Мысалы, (S14) зерттеуінде 300 белгі пайдаланылған жағдайда модельдің жауап беру уақыты 5.62 мс болғаны көрсетілген, бұл оның нақты уақыт режиміне жақын қолданбаларға жарамдылығын дәлелдейді. Ал (S24) жұмысында бір жасырын қабаттан (512 нейрон), шағын пакет өлшемінен (16) және 10 эпохтан аспайтын оқыту циклінен тұратын модель салыстырылған тәсілдер арасындағы ең жеңіл және жылдам архитектура ретінде сипатталған.

Жалпы алғанда, есептеу тиімділігі енгізілген зерттеулердің көпшілігінде жүйелі түрде қарастырылмаған. Тек 4 зерттеуде бұл мәселе тікелей талданса, 2 зерттеуде жанама түрде қарастырылған, ал қалған 19 жұмыста есептеу тиімділігі туралы ақпарат берілмеген (2-сурет). Сондықтан қазіргі әдебиеттерде алгоритмдердің болжамдық дәлдігіне көбірек көңіл бөлініп, есептеу шығындары мен ресурстық тиімділікті бағалау жеткілікті деңгейде

жүргізілмейтіні байқалады.

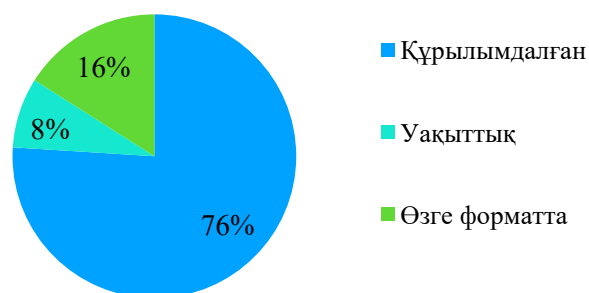


Сурет 2 – Есептеу тиімділігі мәселесін қарастыру белгісі бойынша зерттеулерді санаттау нәтижелері

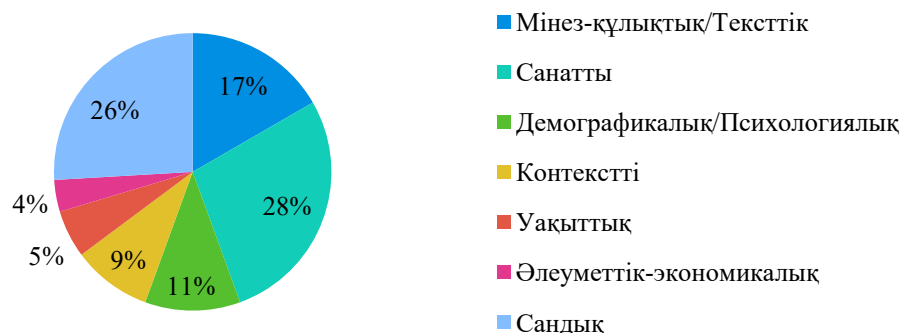
Екінші зерттеу сұрағы бойынша қорытындылай келе, есептеу тиімділігі тұрғысынан нақты көшбасшы алгоритмді анықтау мүмкін болмады, себебі зерттеулердің басым бөлігінде уақыт, жад шығыны немесе ресурстық тұтыну көрсеткіштері жүйелі түрде есептелмеген. Соған қарамастан, жеңілдетілген нейрондық желі архитектуралары, оңтайландырылған ансамбльдік әдістер және күрделілігі төмен модельдер есептеу тиімділігі жоғары тәсілдер ретінде ерекшеленді.

Үшінші зерттеу сұрағы бойынша мансаптық ұсыным модельдерінде қолданылатын деректер түрлері және олардың алгоритмдерді таңдау мен өнімділік көрсеткіштерімен байланысы талданды. Зерттеу барысында деректер форматтарының таралуы, әртүрлі деректер түрлеріне сәйкес келетін алгоритмдер және олардың өнімділігін бағалау үшін қолданылған көрсеткіштер қарастырылды.

а) Қолданылатын деректер түрлері. Талдау нәтижелері мансаптық ұсыным жүйелерінде құрылымдалған кестелік деректердің басым қолданылатынын көрсетті. Енгізілген 25 зерттеудің 19-ында негізгі деректер көзі ретінде құрылымдалған кестелік белгілер пайдаланылған (3-сурет). Деректердің семантикалық құрамы бойынша санаттық және реттік белгілер (15 зерттеу), сондай-ақ сандық белгілер (14 зерттеу) жиі кездеседі. Бірқатар жұмыстарда бірнеше деректер түрі бір мезгілде қолданылғандықтан, деректер санаттарының жиынтық саны зерттеулер санынан асып түседі (4-сурет).



Сурет 3 – ЖИ -мансаптық ұсыным жүйесі зерттеулеріндегі деректер форматтарының таралуы

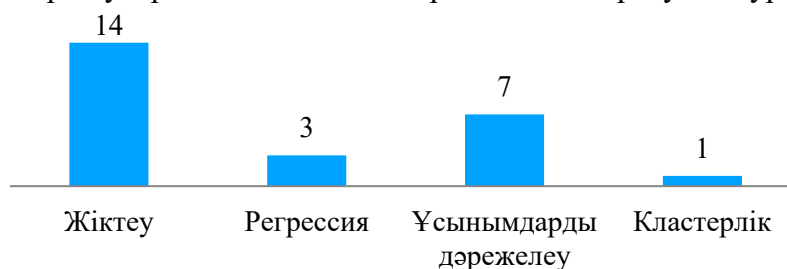


Сурет 4 – Шолудағы деректер түрлерінің пайдаланылуы (N=25).

б) Деректер түрлерінің алгоритмдерді таңдаумен байланысы. Құрылымдалған кестелік деректер пайдаланылған зерттеулерде ең жоғары нәтижелерді негізінен классикалық машиналық оқыту және ансамбльдік әдістер көрсетті. Әсіресе Random Forest [S12, S19, S22] және Gradient Boosting (S11) модельдері жиі кездеседі. Ал мәтіндік немесе аралас форматтағы деректер қолданылған жағдайда SVM және Logistic Regression (S07), CNN және LSTM (S21), CNN (S25), сондай-ақ CIRF-MLP (S15) сияқты нейрондық желі және гибридік модельдер жоғары өнімділік көрсетті. Мінез-құлықтық және реттік деректер үшін үздік нәтижелер ұсынымдық және реттік міндеттерге бейімделген модельдерде байқалды. Олардың қатарында MLP (S24), Hybrid CF + KG + time-decay (S13) және feature-based & time-aware hybrid collaborative filtering (S14) сияқты гибридік ұсынымдық тәсілдер бар. Психометриялық деректер жеке деректер көзі ретінде де, аралас құрылымдалған деректер жиынтықтарының құрамында да пайдаланылды. Бұл санаттағы зерттеулерде алгоритмдердің әртүрлілігі байқалды: KNN (S20), FINN (S10), MES (S17), сондай-ақ FCM және LSE компоненттері бар Takagi–Sugeno моделі (S18) жоғары нәтижелер көрсетті. Ал психометриялық белгілер басқа құрылымдалған деректермен біріктірілген жағдайларда XGBoost (S02) және SVM/LR (S07) сияқты дәстүрлі машиналық оқыту алгоритмдері тиімді болды.

Жалпы алғанда, нәтижелер деректер түрінің алгоритмдерді таңдауға тікелей әсер ететінін көрсетті. Құрылымдалған кестелік деректер үшін ансамбльдік модельдер басым болса, мінез-құлықтық және реттік деректер үшін гибридік ұсынымдық жүйелер, ал психометриялық деректер үшін әртүрлі машиналық оқыту және бұлдыр логикаға негізделген тәсілдер қолданылған.

в) Деректер түрлерінің өнімділік көрсеткіштерімен байланысы. Деректер түрлері мен модельдердің өнімділігі арасындағы байланысты талдау мақсатында зерттеулерде қолданылған бағалау көрсеткіштері төрт негізгі санатқа жіктелді: жіктеу, регрессия, ұсынымдарды дәрежелу және өзге мамандандырылған көрсеткіштер. Мұндай жіктеу әртүрлі деректер түрлеріне сәйкес қолданылатын бағалау тәсілдерін салыстыруға мүмкіндік берді. Зерттеулердің аталған санаттар бойынша таралуы 5-суретте көрсетілген.



Сурет 5 – Кіріс деректері түрлері негізінде қолданылған негізгі бағалау көрсеткіштерінің жіктелуі.

Ескерту: Мәндер тиісті көрсеткіштер тобы пайдаланылған зерттеулер санын көрсетеді.

Құрылымдалған кестелік деректер қолданылған зерттеулерде өнімділік көбінесе жіктеу және регрессия көрсеткіштері арқылы бағаланды. Бұл санаттағы жұмыстарда Accuracy, Precision, Recall, F1-score, AUC, MSE, RMSE және R-squared көрсеткіштері кеңінен қолданылды. Мысалы, (S12) зерттеуінде Accuracy және AUC көрсеткіштері 99%-дан жоғары нәтижелерге жетсе, (S09) жұмысында регрессиялық модельдер үшін төмен қателік мәндері мен жоғары детерминация коэффициенті тіркелген.

Реттік және мінез-құлықтық деректер пайдаланылған зерттеулерде бағалау көбіне ұсынымдық жүйелерге тән көрсеткіштер арқылы жүргізілді. Атап айтқанда, MRR, Recall@K, HR және NDCG сияқты дәрежелік көрсеткіштері (S01, S24), сондай-ақ ұсынымдық міндеттер үшін MAE көрсеткіші (S13) қолданылды.

Мәтіндік және мультимодальды деректер пайдаланылған зерттеулерде өнімділік негізінен жіктеу көрсеткіштері арқылы бағаланды. Бұл жұмыстарда Accuracy, Precision, Recall, F1-score және AUC-ROC көрсеткіштері жоғары мәндер көрсетті (S07, S21, S25).

Жалпы алғанда, нәтижелер деректер түрі мен бағалау көрсеткіштері арасында белгілі бір байланыс бар екенін көрсетті: құрылымдалған кестелік деректер көбіне жіктеу және регрессия көрсеткіштерімен, ал реттік және мінез-құлықтық деректер ұсынымдарды дәрежелік көрсеткіштерімен бағаланған. Мәтіндік және мультимодальды деректер жағдайында жіктеу көрсеткіштері басым қолданылды.

Талқылау. Бірінші зерттеу сұрағы нәтижелері ең жоғары болжамдық дәлдік көрсеткіштері бақылаулы оқытуға негізделген жіктеу міндеттерінде байқалатынын көрсетті. Бұл санатта ансамбльдік әдістер мен нейрондық желі модельдері Accuracy, AUC және F1 көрсеткіштері бойынша жоғары нәтижелерге қол жеткізді. Алайда алгоритмдерді салыстыру тек бірдей міндет қойылымдары мен бірдей бағалау көрсеткіштері шеңберінде ғана орынды. Себебі әртүрлі бағалау көрсеткіштері модель өнімділігінің әртүрлі қырларын сипаттайды. Мысалы, AUC жіктеу сапасын бағаласа, NDCG@K ұсынымдарды дәрежелік сапасын көрсетеді. Сондықтан әртүрлі міндеттерге арналған зерттеулер арасындағы тікелей салыстыру әдіснамалық тұрғыдан дұрыс емес.

Ұсынымдық міндеттерде бағалау көрсеткіштерінің (Precision@K, HR, NDCG және т.б.) және K параметрінің әртүрлілігі ерекше байқалады. Бұл жағдай мақалааралық сандық салыстыру жүргізуді қиындатып, нәтижелерді негізінен сипаттамалық синтез деңгейінде қарастыруға мәжбүр етеді. Жүргізілген талдау негізінде екі негізгі заңдылық анықталды. Біріншіден, ансамбльдік және нейрондық желі модельдері жіктеу міндеттерінде тұрақты түрде жоғары дәлдік көрсетеді. Екіншіден, гибридік және ұсынымдық жүйелерде қолданылатын бағалау көрсеткіштерінің әркелкілігі бірыңғай дәлелдік қор қалыптастыруды қиындатады. Осыған байланысты болашақ зерттеулерде бағалау хаттамаларын стандарттау маңызды болып табылады. Атап айтқанда, жіктеу міндеттері үшін бірыңғай көрсеткіштер жиынтығын, Top-N ұсынымдары үшін бірдей K мәніндегі ортақ дәрежелік көрсеткіштерін, ал регрессия үшін RMSE және MAE көрсеткіштерін қатар ұсыну ұсынылады. Бұл зерттеулер арасындағы нәтижелердің салыстырмалылығын арттырып, мансаптық ұсыным жүйелерінің тиімділігі туралы қорытындылардың сенімділігін күшейтеді.

Кластерлер арасындағы айырмашылықтарды статистикалық талдау. Жіктеу дәлдігі туралы сандық деректер берілген зерттеулер негізінде алгоритмдік кластерлердің сипаттамалық статистикасы есептелді. Орташа мәндер 1-кластер ($M=0.950$), 2-кластер ($M=0.947$) және 3-кластер ($M=0.875$) арасында белгілі бір айырмашылықтардың бар екенін көрсетті. Алайда шағын іріктеме көлеміне байланысты Краскел–Уоллис критерийі бойынша айырмашылықтар статистикалық мәнділік деңгейіне жеткен жоқ ($N=4.00$, $p=0.135$). Әсер мөлшерін бағалау нәтижелері 1-кластер мен 3-кластер, сондай-ақ 2-кластер мен 3-кластер арасында практикалық тұрғыдан елеулі айырмашылықтардың болуы мүмкін екенін көрсетті. Дегенмен, зерттеулер санының шектеулілігі және тәжірибелік жағдайлардың жоғары гетерогенділігі бұл нәтижелерді алдын ала сипаттағы қорытынды

ретінде қарастыруды талап етеді. Сондықтан алынған нәтижелерді жалпылау кезінде сақтық қажет.

Екінші зерттеу сұрағы бойынша нәтижелер есептеу тиімділігі ЖИ-мансаптық ұсыным жүйесі зерттеулерінде жүйелі бағаланатын өлшемшарт ретінде қарастырылмайтынын көрсетті. Күрделі модельдерге деген қызығушылықтың артып келе жатқанына қарамастан, тек аз ғана жұмыстар есептеу шығындарының нақты немесе сандық өлшенетін көрсеткіштерін келтіреді, мысалы кідіріс [7], күрделілік немесе оқыту параметрлері. Бұл орайда келесі заңдылық байқалды. Тоқталатын болсақ, есептеу тиімділігі жоғары шешімдер, әдетте, архитектуралық қарапайымдылығымен ерекшеленеді. Жеңіл нейрондық архитектуралар, классикалық машиналық оқыту алгоритмдері және аз параметрлі модельдер оқыту мен болжау жылдамдығы тұрғысынан жақсы нәтиже көрсетеді. Ал, ансамбльдер мен әсіресе LLM сияқты күрделі тәсілдер айтарлықтай есептеу ресурстарын қажет етеді және жиі дәлдік пен іс жүзінде қолданудың ымырасын болжайды. Көп жағдайда авторлар сапалық тұжырымдармен шектеледі [8] немесе тиімділікті тек жанама түрде қарастырады [9]. Бұл берілген өлшемшарт бойынша алгоритмдердің үстемдігі туралы жалпылалған қорытындыларды едәуір қиындатады. Зерттеушілік тұрғыдан бұл жерде айқын олқылық байқалып отыр. Есептеу тиімділігі тәуелсіз бағалау өлшемшарты ретінде сирек қолданылады. Бұл дәлдік, түсіндіру мүмкіндігі мен есептеу шығындары арасындағы баланс іздеуге бағытталған болашақ жұмыстарға перспективалар ашады. Сонымен қатар, осы шолу аясында алгоритмдерді есептеу тиімділігі бойынша статистикалық салыстыру жүргізу мүмкін болмады. Қамтылған 25 зерттеудің тек 6-сы ғана осы қырынан қарастырады, олардың ішінде тек 2 зерттеу сандық көрсеткіштерді (жауап беру уақыты, есептеу күрделілігі) ұсынады. Қалған жұмыстар сапалық сипаттамалармен ғана шектелген. Салыстыруға болатын сандық деректердің болмауы қандай да бір статистикалық тестті қолдану мүмкіндігін жоққа шығарады.

Деректер түрлерін талдаудағы зерттеулерде құрылымдалған кестелік деректер айқын басымдық танытатыны анықталды. Мәтіндік, мінез-құлықтық, реттік және мультимодальды деректер көздері [10-11] айтарлықтай сирек пайдаланылады. Белгілер арасында санаттық және сандық белгілер ең кең тараған болғаны анықталды. Деректер түрі мен ең тиімді деп танылған алгоритмдер класы арасында тұрақты байланыс байқалады. Кестелік деректер үшін классикалық және ансамбльдік машиналық оқыту әдістері, мәтіндік деректер үшін құрылымдалмаған ақпаратты өңдеу алгоритмдері, мінез-құлықтық және реттік деректер үшін ұсынымдық және уақытша міндет қойылымдарына бағытталған модельдер жиі таңдалады екен. Психометриялық деректер ең жоғары алгоритмдік вариативтілігімен ерекшеленеді және ешбір басымды тәсілге жуымайды.

Анықталған заңдылықтар алгоритм таңдауы бірінші кезекте кіріс деректерінің сипаттамаларымен анықталатынын көрсетеді. Тіркелген белгі векторлары бар сценарийлерде кестелік құрылымды жақсы өңдей алатын алгоритмдерге артықшылық беріледі. Мәтіндік компоненттер қосылған кезде зерттеушілер құрылымдалмаған мазмұннан белгілерді шығара алатын модельдерге жүгінеді. Деректер динамиканы немесе өзара әрекеттесулердің реттілігін бейнелейтін жерлерде оқиғалардың ретін және уақытша тәуелділігін ескеретін тәсілдер қолданылады. Психометриялық деректер үшін басымды алгоритмнің болмауы психологиялық конструктілерді ұсыну тәсілдерінің алуандылығымен және мансаптық ұсыным жүйелерінде оларды модельдеудің қалыптасқан тәжірибесінің жоқтығымен түсіндіріледі.

Үшінші зерттеу сұрағының мақсаты сандық көрсеткіштерді салыстыру емес, деректер түрлерінің таралуын сипаттау және олардың алгоритмдерді таңдаумен байланысын анықтау болып табылады. Сол себепті статистикалық тест қолданылған жоқ. Қарастырылған зерттеулерде құрылымдалған кестелік деректер негізінен жіктеу көрсеткіші арқылы бағаланатындығына көз жеткізілді. Регрессиялық көрсеткіштер айтарлықтай сирек қолданылады және тек жекелеген жұмыстарда кездеседі [12-13]. MRR, HR, NDCG және

Precision@k сияқты ұсынымдық және ранг көрсеткіштері негізінен мінез-құлықтық және реттік деректерге ғана тән [14]. Барлық тиісті зерттеулерде мәтіндік және мультимодальды деректер Accuracy, AUC, Precision, Recall және F1 қоса алғанда тек жіктеу көрсеткіштері арқылы бағаланады.

Жүйелі шолудың шектеулері. Маңызды шектеудің бірі - толық мәтіндік сүзгілеу кезеңіндегі 123 жарияланымның 86-ы (69,9%) ақылы болу салдарынан қолжетімсіз болуы. Бұл іріктеу қатесінің тәуекелін арттырды. Демек, алынған нәтижелерді абайлап талдаған жөн. Қолжетімсіз зерттеулер өзге, соның ішінде қайшылықты деректерді қамтуы мүмкін. Енгізілген жұмыстардың жоғары гетерогенділігі де өз шектеулерін туғызды. Деректер жиынтықтарындағы, бағалау көрсеткіштеріндегі және міндет қойылымдарындағы айырмашылықтар ресми мета-талдау жүргізуге мүмкіндік бермеді. Нәтижесінде нәтижелерді синтездеу сапалық сипаттамалық талдаумен шектелді.

Берілген жұмыс әдебиеттерге жүйелі шолу болып табылады және авторлар тарапынан алынған түп нұсқалық тәжірибелік деректерді қамтымайды. Бұл — PRISMA 2020 нұсқаулығына сәйкес жүйелі шолу жанрының талаптарына сай қабылданған саналы әдіснамалық шешім. Зерттеудің осы түрінің негізгі мақсаты жаңа тәжірибелік нәтижелер алу емес, қолда бар ғылыми дәлелдерді жүйелеу, синтездеу және сыни тұрғыдан бағалау болып табылады.

Жүргізілген шолу бар әдебиеттегі бірнеше маңызды олқылықтарды анықтауға мүмкіндік береді. Біріншіден, реттік және мультимодальды деректер шектеулі түрде ұсынылған. Бұл динамикалық мансаптық траекторияларға бағытталған жүйелерге арналған қорытындылардың жалпыламалылығын төмендетеді. Екіншіден, психометриялық және әлеуметтік-экономикалық деректер ішінара пайдаланылады және олар үшін алгоритмдер таңдауы мен бағалау хаттамалары бойынша ортақ ұстаным жоқ. Үшіншіден, қолданылатын көрсеткіштерінің гетерогенділігі байқалады. Бұл процесс зерттеулер арасындағы өнімділікті салыстыруды қиындатады және жинақталған талдау мүмкіндігін шектейді.

Зерттеулерге арналған ұсынымдар. Болашақ зерттеулерде қолданылатын деректер көздерін кеңейту қажет. Әсіресе реттік, мәтіндік және мультимодальды деректерге, сондай-ақ психометриялық және әлеуметтік-экономикалық белгілерді жүйелі енгізуге көбірек назар аударған жөн. Сонымен қатар бағалау көрсеткіштерін міндет түрлеріне сәйкес стандарттау зерттеулер арасындағы нәтижелердің салыстырмалылығын арттыруға мүмкіндік береді. Алынған нәтижелер мансаптық ұсыным жүйелерінде алгоритм таңдауы ең алдымен қолжетімді деректердің сипатына негізделуі тиіс екенін көрсетті. Деректер ерекшеліктерін ескермей әмбебап модельдерді қолдану жүйенің тиімділігі мен талдау мүмкіндіктерін төмендетуі мүмкін.

Алгоритмді таңдау бойынша практикалық ұсынымдар. Қамтылған зерттеулердің нәтижелерін синтездеу негізінде мансаптық ұсыным жүйелерін әзірлеушілер үшін келесі ұсынымдар тұжырымдалды:

- Құрылымдалған кестелік деректер (академиялық үлгерім, тест нәтижелері, әлеуметтік-демографиялық сипаттамалар) үшін Random Forest және Gradient Boosting сияқты ансамбльдік әдістерді қолдану ұсынылады, себебі олар жоғары дәлдік пен жақсы түсіндіру мүмкіндігін қамтамасыз етеді.

- Мәтіндік және мультимодальды деректерді өңдеу үшін CNN және LSTM негізіндегі модельдер тиімді, алайда олардың есептеу ресурстарына қойылатын талаптары жоғары.

- Мінез-құлықтық және реттік деректер жағдайында уақыттық тәуелділіктерді ескеретін гибридік ұсынымдық тәсілдер қолайлы болып табылады.

- Психометриялық деректер үшін әмбебап тиімді алгоритм анықталмағандықтан, бірнеше модельді нақты деректер жиынтығында салыстырмалы түрде сынақтан өткізу ұсынылады.

- Бағалау көрсеткіштері міндет түріне сәйкес таңдалуы тиіс: жіктеу үшін Assurasy және F1-score, регрессия үшін RMSE немесе MAE, дәрежелеу міндеттері үшін NDCG және Precision@K.
- Есептеу тиімділігін стандарттау мақсатында болашақ зерттеулерде модельді оқыту уақыты, болжау уақыты, параметрлер саны және қолданылған аппараттық орта туралы ақпаратты міндетті түрде көрсету ұсынылады.
- Алгоритмдердің асимптотикалық күрделілігін (O-нотация) көрсету нәтижелердің қайта жаңғыртылуын жақсартып, модельдерді практикалық қолдану тұрғысынан салыстыруға мүмкіндік береді.

Қорытынды.

Зерттеу барысында жасанды интеллект негізіндегі мансаптық ұсыным жүйелеріне арналған әдебиеттерге жүйелі шолу жүргізілді. Нәтижелер алгоритмдердің өнімділігі міндет түріне, бағалау көрсеткіштеріне және пайдаланылатын деректердің сипатына тәуелді екенін көрсетті.

Бірінші зерттеу сұрағы бойынша ең жоғары дәлдік көрсеткіштері негізінен жіктеу міндеттерінде ансамбльдік және нейрондық желі модельдері арқылы алынатыны анықталды. Екінші зерттеу сұрағы есептеу тиімділігінің қазіргі зерттеулерде жеткілікті деңгейде бағаланбайтынын көрсетті. Үшінші зерттеу сұрағы нәтижелері алгоритмдерді таңдау мен олардың өнімділігі деректер түрлерімен тығыз байланысты екенін дәлелдеді.

Жалпы алғанда, болашақ зерттеулер үшін бағалау хаттамаларын стандарттау, есептеу тиімділігін жүйелі түрде есепке алу және мультимодальды деректерді кеңінен пайдалану қажеттілігі анықталды. Жұмыстың ғылыми жаңалығы алгоритмдердің бес кластерден тұратын жіктелуін ұсынуымен, деректер түрлері мен алгоритмдер арасындағы байланысты анықтауымен және осы саладағы негізгі зерттеу олқылықтарын жүйелеуімен сипатталады.

Әдебиеттер тізімі

1. Trujillo, F., Pozo, M., & Suntaxi, G. (2025). Artificial intelligence in education: A systematic literature review of machine learning approaches in student career prediction. *Journal of Technology and Science Education*, 15(1), 162-185. <https://doi.org/10.3926/jotse.3124>
2. Zayed, Y., Salman, Y., & Hasasneh, A. (2022). A Recommendation System for Selecting the Appropriate Undergraduate Program at Higher Education Institutions Using Graduate Student Data. *Applied Sciences*, 12(24), 12525. <https://doi.org/10.3390/app122412525>
3. Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ. Computer science*, 7, e623. <https://doi.org/10.7717/peerj-cs.623>
4. Gunawardana, A., & Shani, G. (2009). A Survey of Accuracy Evaluation Metrics of Recommendation Tasks. *Journal of Machine Learning Research*, 10(100), 2935–2962. <https://jmlr.csail.mit.edu/papers/volume10/gunawardana09a/gunawardana09a.pdf>
5. Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
6. Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering (EBSE Technical Report EBSE-2007-01). Keele University and Durham University Joint Report.
7. Mahboob K, Asif R, Haider N. Career planning matters: Intelligence-based career path predictions using data mining models - A longitudinal study. *MEHRAN UNIVERSITY RESEARCH JOURNAL OF ENGINEERING AND TECHNOLOGY*. 2024;43(4):192-213. doi:10.22581/muet1982.3343.

8. Gülten, H., & Baraçlı, H. (2024). A Machine Learning-Based Forecast Model for Career Planning in Human Resource Management: A Case Study of the Turkish Post Corporation. *Applied Sciences*, 14(15), 6679. <https://doi.org/10.3390/app14156679>
9. Fatima N, Gul S, Ahmed J, Khand Z, Mujtaba G. A rule-based machine learning model for career selection through MBTI personality. *Mehran university research journal of engineering and technology*. 2022;41(2):185-196. doi:10.22581/muet1982.2202.18.
10. Cheng Y. Development of a Career Planning Assistance Platform for College Students Combining Fuzzy Logic and Deep Learning. *Journal of Advanced Computational Intelligence and Intelligent Informatics*. 2025;29(3):445-455. doi:10.20965/jaciii.2025.p0445.
11. José-García, A., Sneyd, A., Melro, A., Ollagnier, A., Tarling, G., Zhang, H., Stevenson, M., Everson, R., & Arthur, R. (2023). C3-IoC: A career guidance system for assessing student skills using machine learning and network visualisation. *International Journal of Artificial Intelligence in Education*, 33, 1094–1121. <https://doi.org/10.1007/s40593-022-00317-y>
12. Talib, A., Housni, M., & Radid, M. (2023). Utilizing M-technologies for AI-driven career guidance in Morocco: An innovative mobile approach. *International Journal of Interactive Mobile Technologies*, 17, 173–188. <https://doi.org/10.3991/ijim.v17i24.44263>
13. Zylowski, T., Sautchuk-Patricio, N., Hettmann, W., Anderer, K., Fischer, K., Wölfel, M., & Henning, P. (2024). User study on the trustworthiness, usability and explainability of intent-based and large language model-based career planning conversational agents. In *Proceedings of the 16th International Conference on Education Technology and Computers* (pp. 46–53). ACM. <https://doi.org/10.1145/3702163.3702409>
14. Alonso, R., Dessì, D., Meloni, A., Murgia, M., Recupero, D. R., & Scarpi, G. (2024). A seamless ChatGPT knowledge plug-in for the labour market. *IEEE Access*, 12, 155821–155837. <https://doi.org/10.1109/ACCESS.2024.3485111>

ПРОИЗВОДИТЕЛЬНОСТЬ СИСТЕМ КАРЬЕРНЫХ РЕКОМЕНДАЦИЙ НА ОСНОВЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: СИСТЕМАТИЧЕСКИЙ ОБЗОР

Аннотация. В условиях стремительных изменений на рынке труда и роста доступности образовательных данных разработка систем карьерного консультирования на основе искусственного интеллекта (ИИ) становится одним из актуальных направлений исследований. Однако влияние типов данных и особенностей выбора алгоритмов на прогностическую точность и вычислительную эффективность моделей до сих пор недостаточно систематизировано. Целью данного исследования является систематизация современных исследований в области ИИ-ориентированных систем карьерных рекомендаций и выявление взаимосвязей между типами данных, алгоритмами и показателями оценки. Исследование основано на систематическом обзоре, проведённом в соответствии с методологией PRISMA 2020. Анализ выполнен в рамках взаимосвязи «тип данных – алгоритм – показатель оценки». Результаты показали, что среди рассмотренных исследований преобладают задачи классификации (56 %), а наиболее часто используются структурированные табличные данные (76 %). Наивысшие показатели прогностической точности были достигнуты преимущественно ансамблевыми методами и моделями нейронных сетей. Вместе с тем выявлено, что вычислительная эффективность в большинстве исследований не оценивается систематически: в 76 % работ соответствующие показатели отсутствуют. Кроме того, установлено, что выбор алгоритмов в значительной степени зависит от используемых типов данных. Полученные результаты позволяют систематизировать алгоритмические подходы в системах карьерных рекомендаций и указывают на необходимость стандартизации протоколов оценки, систематического учёта вычислительной эффективности и более широкого использования мультимодальных данных в будущих исследованиях. Кроме того, разработка ИИ-ориентированных систем

рекомендаций, адаптированных к особенностям национального рынка труда и образовательной среды, представляет собой актуальное направление исследований.

Ключевые слова: искусственный интеллект, карьерные рекомендации, рекомендательные системы, типы данных, алгоритмы машинного обучения, прогностическая точность, вычислительная эффективность.

PERFORMANCE OF ARTIFICIAL INTELLIGENCE-BASED CAREER RECOMMENDATION SYSTEMS: A SYSTEMATIC REVIEW

Abstract. In the context of rapid changes in the labor market and the increasing availability of educational data, the development of artificial intelligence (AI)-based career recommendation systems has become an important research direction. However, the influence of data types and algorithm selection on the predictive accuracy and computational efficiency of such systems remains insufficiently systematized. The aim of this study is to systematize current research on AI-driven career recommendation systems and to identify relationships among data types, algorithms, and evaluation metrics. The study is based on a systematic literature review conducted in accordance with the PRISMA 2020 guidelines. The analysis was performed within the “data type–algorithm–evaluation metric” framework. The results show that classification tasks dominate the reviewed studies (56%), while structured tabular data represent the most frequently used data type (76%). The highest predictive performance was achieved primarily by ensemble machine learning methods and neural network models. At the same time, a significant gap was identified regarding the assessment of computational efficiency, as 76% of the analyzed studies did not report relevant computational performance measures. Furthermore, the findings indicate that algorithm selection is strongly associated with the characteristics of the underlying data. The study contributes to the systematization of algorithmic approaches used in career recommendation systems and highlights the need for standardized evaluation protocols, systematic reporting of computational efficiency, and broader adoption of multimodal data in future research. In addition, the development of AI-driven recommendation systems tailored to the specific characteristics of national labor markets and educational environments represents an important direction for future studies.

Keywords: artificial intelligence, career recommendations, recommender systems, data types, machine learning algorithms, predictive accuracy, computational efficiency.

Авторлар туралы мәлімет

Абраймова Динара Сержановна	Техника ғылымдарының магистрі, С.Сейфуллин атындағы Қазақ агротехникалық зерттеу университетінің, «Бизнес және цифрлық технологиялар» институтының докторанты, Астана қ., Қазақстан, E-mail: dabraymova@gmail.com
Нургалиева Сымбат Алтыбаевна	PhD, Astana IT university, Бағдарламалық инженерия кафедрасының ассистент профессоры, Астана қ., Қазақстан E-mail: symbat.nurgaliyeva@astanait.edu.kz
Башеева Жулдыз Орынбасаровна	PhD, Astana IT university, Бағдарламалық инженерия департаментінің ассистент профессоры, Астана қ., Қазақстан E-mail: zhuldyz.basheeva@astanait.edu.kz
Туткышбаева Шырын Сарыбайевна	PhD, Astana IT university, Бағдарламалық инженерия департаментінің ассистент профессоры, Астана қ., Қазақстан E-mail: Sh.Tutkyshbayeva@astanait.edu.kz
Сыздықпасва Айгуль Рамазановна	Сәрсен Аманжолов атындағы Шығыс Қазақстан университеті, жасанды интеллект және ақпараттық технологиялар кафедрасының қауымдастырылған профессоры, т.ғ.к., Өскемен қ., Қазақстан, E-mail: aigul uk@list.ru

Сведения об авторах

Абраймова Динара Сержановна	Магистр технических наук, докторант института «Бизнеса и цифровых технологий» Казахского агротехнического исследовательского университета имени С.Сейфуллина, г. Астана, Казахстан, E-mail: dabraymova@gmail.com
Нургалиева Сымбат Алтыбаевна	Ассистент профессор кафедры программной инженерии, PhD, Астана IT университет, Астана, Казахстан, E-mail: symbat.nurgaliyeva@astanait.edu.kz

Башеева Жулдыз Орынбасаровна	Ассистент профессор департамента программной инженерии, PhD, Астана IT университет, Астана, Казахстан E-mail: zhuldyz.basheeva@astanait.edu.kz
Туткышбаева Шырын Сарыбайевна	Ассистент профессор департамента программной инженерии, PhD, Астана IT университет, Астана, Казахстан E-mail: Sh.Tutkyshbayeva@astanait.edu.kz
Сыздыкпаева Айгуль Рамазановна	Ассоциированный профессор кафедры искусственного интеллекта и информационных технологий, PhD, Восточно-Казахстанский университет имени Сарсена Аманжолова, Усть-Каменогорск, Казахстан, E-mail: aigul_uk@list.ru

Information about authors

Abraimova Dinara Serzhanovna	Master of Technical Sciences, Doctoral student at the Institute of Business and Digital Technologies of the Kazakh Agrotechnical Research University named after S. Seifullin, Astana, Kazakhstan, E-mail: dabraymova@gmail.com
Nurgaliyeva Symbat Altybaevna	Assistant professor of department of software engineering, PhD, Astana IT University, Astana, Kazakhstan E-mail: symbat.nurgaliyeva@astanait.edu.kz
Basheeva Zhuldyz Orynassarovna	Assistant professor of department of software engineering, PhD, Astana IT University, Astana, Kazakhstan, E-mail: zhuldyz.basheeva@astanait.edu.kz
Shyryn Tutkyshbayeva	Assistant professor of department of software engineering, PhD, Astana IT University, Astana, Kazakhstan, E-mail: Sh.Tutkyshbayeva@astanait.edu.kz
Syzdykpayeva Aigul Ramazanovna	Associate Professor of the Department of Department of Artificial Intelligence and Information Technology, c.t.sc., Sarsen Amanzholov East Kazakhstan university, Ust-Kamenogorsk, Kazakhstan E-mail: aigul_uk@list.ru