

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

«Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ» КЕАҚ

**Студенттер мен жас ғалымдардың
«ҒЫЛЫМ ЖАҢЕ БІЛІМ - 2025»
XIX Халықаралық ғылыми конференциясының
БАЯНДАМАЛАР ЖИНАҒЫ**

**СБОРНИК МАТЕРИАЛОВ
XX Международной научной конференции
студентов и молодых ученых
«ҒЫЛЫМ ЖАҢЕ БІЛІМ - 2025»**

**PROCEEDINGS
of the XX International Scientific Conference
for students and young scholars
«ҒЫЛЫМ ЖАҢЕ БІЛІМ - 2025»**

**2025
Астана**

УДК 001(06)
ББК 72я631
F96

**«GYLYM JÁNE BILIM – 2025» студенттер мен жас ғалымдардың
XX Халықаралық ғылыми конференциясы = XX Международная
научная конференция студентов и молодых ученых «GYLYM JÁNE
BILIM – 2025» = The XX International Scientific Conference for
students and young scholars «GYLYM JÁNE BILIM – 2025». – Астана:
– 3813 б. - қазақша, орысша, ағылшынша.**

ISBN 978-601-08-5373-7

**Жинаққа студенттердің, магистранттардың, докторанттардың және жас
ғалымдардың жаратылыстану-техникалық және гуманитарлық ғылымдардың өзекті
мәселелері бойынша баяндамалары енгізілген.**

**The proceedings are the papers of students, undergraduates, doctoral students and young
researchers on topical issues of natural and technical sciences and humanities. В сборник
вошли доклады студентов, магистрантов, докторантов и молодых ученых по
актуальным вопросам естественно-технических и гуманитарных наук.**

УДК 001(06)
ББК 72я431
F96

ISBN 978-601-08-5373-7

Л.Н. Гумилев атындағы Еуразия
ұлттық университеті, 2025

		сауаттылығын арттыру	
203.	Эрболат А.	Орта мектепте нанотехнология ұғымын оқытудың тиімді әдістері	808

СЕКЦИЯ 2

СОВРЕМЕННЫЕ ИНФОРМАЦИОННЫЕ И КОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

Подсекция 2.1			
Цифровая трансформация образования			
204.	Адалбек Н.	«Традиционные и интеллектуальные подходы в обучении»	812
205.	Бакенова А.А.	«Цифровизация тестирования: разработка нейросетевого приложения для формирования заданий по английской грамматике»	816
206.	Бекмурат А.Е.	«Инновационные методы обучения информатике в школе на основе искусственного интеллекта»	821
207.	Назарова А.Т.	«Развитие цифровых компетенций учителей в условиях персонализированного обучения»	826
208.	Нуриева Д.Р.	«Цифровая трансформация педагогики: роль информационных технологий в повышении квалификации преподавателей»	830
209.	Абдуашимова П.М.	«Білім беру процесінде жасанды интеллект технологияларын қолданудың тиімділігі»	833
210.	Ажибаева А.Д.	«Мектеп информатикасын оқытудағы кемшіліктерді жою жолдары»	837
211.	Асылбек М.А.	«Орта мектепте білім беру үдерісінде үлкен деректерді қолдану әдістемесі»	842
212.	Аталова А.Е.	«Әлеуметтік желілерді информатика пәні бойынша оқыту құралы ретінде пайдалану»	845
213.	Балтабаев Н.П.	«Мектептерде сабақ кестесін автоматтандыруға арналған интеллектуалды жүйе құру»	851
214.	Балтабаев Н.П., Дәрменов Ә.М., Мұратова М.М.	«Жасанды интеллект негізінде жаратылыстану пәндерін оқытуды жетілдіру: BilimALL AI платформасының мүмкіндіктері»	854
215.	Баумуратова Х.Б.	«АКТ оқыту барысында бастауыш сынып оқушыларының цифрлық сауаттылықтарын қалыптастырудың әдістемесі»	856
216.	Баумуратова Ш.Б.	«Жасанды интеллект негізінде инклюзивті білім беруді жетілдіру»	859
217.	Ғазиз Ж.Е.	«Бастауыш мектепте ақпараттық-коммуникациялық технологияларды оқыту әдістемесі»	863
218.	Дәрменов Ә.М.	«Информатиканы қолжетімді ететін мобильді "BilimAll" қосымшасы»	866
219.	Дүйсегалиева Н.А.	«HIGH-TOUCH HIGH-TECH моделі арқылы болашақ информатика мұғалімдерін машиналық оқыту негізінде даярлаудың	870

	инновациялық тәсілдері туралы»	
220.	Еликбай А.Ж. «Ақпараттық дәуірде білім берудің жаңа кезеңі – Инфографика»	874
221.	Жаңабекқызы А. «EDCAFE AI көмегімен сабақты жоспарлау»	879
222.	Жумабекова У.Б., Сабырова М.Е., Сабыров Т.С. «Информатика пәнін жобалап оқыту технологиясы»	883
223.	Кендебай Н.А. «EDUVISION білім беру процесін қадағалайтын қосымша»	888
224.	Көшенова А. «Цифрлық сауаттылықтың мектеп курсы бойынша интеллектуалдық оқу басылымдарына арналған дидактикалық материалдар»	891
225.	Куанышева Д.Ж. «Инклюзивті білім беруде педагогтың ақпараттық-коммуникациялық технологияларды (АКТ) қолдану даярлығын жетілдіру»	893
226.	Мауленова М.А. «Үлкен деректерді өңдеуде машиналық оқытудың әдістері мен құралдары»	897
227.	Мылтыкбаева Ж.Т. «Жаратылыстану пәндерін STEM білім беру мен ROS операциялық жүйесі негізінде кешенді оқыту»	901
228.	Надирхан Г.Е. «Ауыл мектептерінде цифрлық оқытуды дамыту мүмкіндіктері»	903
229.	Орынбаев М.Ж. «Компьютерлік көру алгоритмдерін машиналық оқыту негіздері бойынша қолданудың оқу-әдістемелік негіздері»	907
230.	Сабитова А.Б., Ражапова А.Н. «Жасанды интеллект және білім: болашақ мұғалімдерге арналған жаңа мүмкіндіктер»	910
231.	Сағындықова А.С. «Болашақ информатика мұғалімдерін магистратураға даярлаудағы онлайн-курстардың рөлі»	915
232.	Сайлау Ж.Б. «Халықаралық зерттеуге оқушыларды АКТ арқылы дайындаудағы педагогтердің құзыреттілігін арттыру жолдары»	918
233.	Төрәлі Қ.Н. «Бастауыш сынып оқушыларының цифрлық сауаттылығын дамытудың ерекшеліктері»	923
234.	Турмаганбетова З.П., Алтыбаева А.Н. «Ерекше білімді қажет ететін оқушыларға мектеп информатика курсы оқытуды ұйымдастыру»	927
235.	Халхабай А. ««Алгоритмдеу және бағдарламалау» курсы бойынша мобильді қосымшаны оқу үдерісінде қолдану»	931
236.	Ысмайыл Н. «Мектеп информатика курсына жобалық оқыту әдісін енгізу»	936
237.	Ізбасарова М.Р. «Білім берудегі тестілеу жүйелері»	938

Подсекция 2.2

Интеллектуальные информационные системы

238.	Amantayeva Gulden Turarkyzy «Comparative analysis of models and methods in heart disease prediction problems»	944
------	---	-----

239.	Tanirbergenov Meirbek Sagyndykovich «Facial Recognition-Based Attendance Management»	947
240.	Toleubay Daniyar Manatuly «Cardiac disease prediction using machine learning algoritms»	952
241.	Yerezhepov Rakhat Aibulatovich «Detecting logical fallacies in web content with nlp-powered crawling»	957
242.	Ажикенов Арман Русланович, Абашев Арслан Азатабекович «Оптимизация дорожного трафика в Астане через симуляцию транспортных потоков»	962
243.	Аманжол Альфараби Маликович, Сабит Мадияр, Кушербаев Бекзат Алибекулы «Система визуализации и анализа данных о передвижении нефти на основе интерактивной карты»	968
244.	Аскапова Мадина Куанышбековна «Параллельді қазақ-түрік сөйлеу корпусы қалыптастырудың әдісі мен моделін құру»	972
245.	Бекқожин Дастан Ақанұлы «Терең оқыту негізінде қолжазба таңбаларын тану программалық құралын әзірлеу»	975
246.	Дакенов Алишер Мырзахметұлы «Анализ сигналов ЭЭГ нейросетевыми методами для ранней диагностики нейродегенеративных заболеваний»	978
247.	Доспол Нәзгүл Нурланқызы, Жеткенбай Лена «Балабақшадағы балалардың эмоциялық жағдайын бақылауға арналған эмоцияларды тану жүйесін әзірлеу»	987
248.	Ермекбай Айболат, Молдабек Елжан «Жасанды интеллект негізінде веб-қосымша әзірлеу»	992
249.	Жұмал Жания Ержанқызы, Абдурахман Жансая Берікжанқызы «Применение голосового ИИ-помощника в геймифицированной образовательной среде»	1001
250.	Каримов Руслан Жасинович «Эффективность существующих ИИ-решений в основных направлениях транспортной логистики»	1007
251.	Кубиева Сабина Талгатовна, Утепбергенова Зарина Арманкызы «Разработка iot системы по уходу за растениями на базе искусственного интеллекта»	1012
252.	Кудобаев Даниал Дулатович «Разработка информационной системы для автоматизации стоматологических услуг»	1017
253.	Мусина Данель Тлеухановна «Интеллектуальные инструменты автоматизированной диагностики надежности информационных систем»	1024
254.	Рогова Ксения Александровна, Қабдыбек Ризат Досмжанұлы, Джумадиева Тогжан Бекежановна «Мониторинг инженерных конструкций на основе искусственного интеллекта»	1030

255.	Сафонова Софья Александровна «Современные аспекты информационной безопасности в облачных вычислениях: модели, угрозы и методы защиты»	1034
256.	Смаилова Назгүл Батырбекқызы «Терең оқыту арқылы кітап ұсыныстарын әзірлеу: collaborative filtering, content-based және nlp әдістерінің комбинациясы»	1041
257.	Тажібай Аружан Айдосқызы, Кудубаева Сауле Альжановна «Көру қабілеті әлсіз адамдарға арналған ai дауыстық көмекші: нақты уақытта объектілерді анықтау және қашықтықты бағалау»	1046
258.	Тайжанов Азамат Жанкелдіұлы «Python тілінде фильмдердің интеллектуалды ұсыныс жүйесін әзірлеу»	1051
259.	Умирзахов Сундетали Кабылбекович «Сұраныстарды интеллектуалды талдау негізінде ұйымның сайты үшін чат-бот құру»	1055
260.	Шайхстан Марғұлан «ИОТ Сенсорлары негізінде ауа ластану деңгейін болжау»	1060

Подсекция 2.3

Современные тенденции в программной инженерии и управлении в условиях цифровой индустрии

261.	Bekenova A.B. «Development of a registration panel for users and doctors with integration into the database»	1077
262.	Bolat A.Zh. «Data analysis methods and decision making using big data and machine learning tools»	1081
263.	Алтайұлы А. «Visual studio интегралды ортасында «қойма қызметкерлеріне арналған» мәліметтер қорын жобалау»	1086
264.	Арап А.Қ. «Ақылды сурет салушы роботты әзірлеу»	1088
265.	Артыкбекқызы А. «Ақылды үйлердегі заттар интернеті(iot) мен робототехниканың өзара әрекеттесуі»	1091
266.	Ахметова А.Д. «Тоңазытқыштағы өнімдерді бақылауға және тағам әзірлеу ұсынысын беруге арналған программалық қосымша»	1096
267.	Дәрібай Д.Д. «Робототехниканы қолдану арқылы қойма логистикасындағы қолданыстағы басқару жүйелерін талдау»	1100
268.	Жамбулов С.Ж. «Білім алушыларды информатика және программалау олимпиадаларына дайындауда жасанды интеллекттің қолданысы»	1102
269.	Каиржан Р.С. «Development of system for recognition of emotional states of employees based on computer vision methods on Raspberry Pi»	1108
270.	Кайрекенова Н.Р. «Өнеркәсіптік роботты көру үшін машиналық оқытудың заманауи тәсілдері: әдістер, деректер жиынтығы және оптимизациялау»	1111

271.	Калижан А.К. «Разработка системы биометрической аутентификации с предотвращением deepfake атак»	1113
272.	Касылкасова К.Н. «Программное обеспечение smartmed для обработки медицинских данных и диагностики»	1118
273.	Қабдешев Ә.Е. «Жөтелді талдау негізінде денсаулықты диагностикалаудың интеллектуалды программасын әзірлеу»	1120
274.	Махаев Е.Е. «Разработка облачного приложения для автоматизации деятельности сети аптек»	1123
275.	Муратов М.М. «Эффективность единой информационной системы агропромышленного комплекса»	1126
276.	Нуржанова А.Б. «Современные методы классификации эмоций: анализ подходов и перспективы развития»	1130
277.	Нурпеисова З.Р. «Обзор и исследование методов искусственного интеллекта для анализа рынка недвижимости»	1134
278.	Рақымбек А.С. «Кітапқұмарларға арналған платформа: кітаптарды оқу және бөлісу үшін әлеуметтік желіні жобалау және іске асыру»	1138
279.	Сагидуллина Д.С. «Visual studio интегралды ортасында «қаржылық транзакцияларды қадағалау және талдауға арналған» мәліметтер қорын жобалау»	1144
280.	Төлеубай Д.М. «Yolov10 қолдану арқылы рентген суреттерінде сүйек сынуын анықтауды кешенді зерттеу»	1147
281.	Утегенова Д.Б. «Visual studio интегралды ортасында «фитнес орталық қызметкері үшін» мәліметтер қорын жобалау»	1152
282.	Шаймуратов А.Ж. «Проектирование аппаратно-программного комплекса для автоматизированного учета железнодорожного подвижного состава»	1154
Подсекция 2.4		
Информационная безопасность		
283.	Akniyet N. «Smart home automation and security system using arduino uno r4 and esp32 microcontrollers with telegram integration»	1158
284.	Askhatov A. «Analysis of social engineering methods and development of a defense strategy for corporate structures»	1165
285.	Bekturganov A.B. «Development of an early detection model for ddos attacks based on network traffic analysis»	1170
286.	Gabdullin A. «Analysis of modern wireless network security protocols and prospects for their development»	1174

287.	Garifullin A. «Modern information security management systems: construction and implementation in the digital era»	1179
288.	Igumenshev D.V. «Methods of embedding malicious code into pdf files»	1182
289.	Issabay T.B. «Utilizing sandboxes for cybersecurity training: a hands-on approach»	1187
290.	Kalybayev S. «Overview of modern authentication methods in telecommunication systems: from passwords to biometrics»	1191
291.	Kerim A. «Owasp top 10 and alternative methods of its compilation»	1194
292.	Yergazin A. «Analysis of a protection of hybrid intrusion detection and prevention system (idps) for low-latency 5g networks with adaptive learning using edge computing»	1199
293.	Yerzhanova Y.Y. «Key attacks in web forensics: xss, sql injection and rce»	1204
294.	Zhakay A. «Fundamentals of modern cryptography: from encryption to digital signatures»	1209
295.	Айдарова А.А. «Visualvm көмегімен cast-128 және kuznyechik блоктық шифрларының кілт генерациясын салыстыру және стандарттарға шолу»	1214
296.	Акимбекова Д.М., Каиржанова Д.Ж. «Жергілікті желінің қауіпсіздігін қамтамасыз ететін негізгі параметрлер»	1220
297.	Аскарлов А.Д. «Разработка и исследование эффективности метода и инструмента для выявления фейковых новостей в социальных сетях»	1224
298.	Ауесхан Н. «Аномалияларды анықтау әдістерін талдау»	1229
299.	Ерболатов А. «Анализ вредоносных программ с помощью ии и криптографическая защита»	1332
300.	Ерболатова А.Ж. «Neuvector және kubernetes: контейнерлік ортадағы қауіпсіздікті қамтамасыз ету тәсілдері»	1336
301.	Жанатаев М.К. «Стеганография на основе lsb: реализация сокрытия данных в медиафайлах»	1338
302.	Жарасхан Н.Ж., Қайупов Е.К. «Crystals-kyber алгоритмін ресурсы шектеулі құрылғыларға оңтайландыру»	1343
303.	Жолдасбаев М.Ә. «Заманауи операциялық жүйелердегі жады дампы кескінін алу құралдарын талдау және салыстыру»	1348
304.	Жолмұратұлы Б., Маратов Ә.Б., Ховдабай Н.А. «Екі факторлы	1353

	аутентификацияның қауіпсіздігі және оның қолданылуы»	
305.	Кадринов Д.М. «Автоматизация внедрения альтернативной soag платформы на основе средств со свободной лицензией»	1357
306.	Казбаганбетова М.А. «Wireshark бағдарламасын пайдаланып желілік трафикті талдау және ақпараттық қауіпсіздікті қамтамасыз ету»	1361
307.	Кәкімбек Ә.Қ., Серікбай А.Е., Наурызбаев Д.Е. «MITM шабуылы туралы»	1366
308.	Кеттеш Б.Н. «ELF талдауындағы capstone: сызықтық және рекурсивті дизассемблерлеу»	1370
309.	Көшкінбаева Ф.Қ. «Linux қорғаудың заманауи әдістеріне талдау.openvas және nmap көмегімен осалдықтарды анықтау»	1374
310.	Қадыр Н.Е. «Заманауи фишинг түрлері мен олардың ұйымдық ақпараттық жүйелерге ықпалы»	1379
311.	Қажкен Е.Е., Темиржан С.А. «Қауіпсіздік инциденттеріне қалай жауап беруге болады?»	1384
312.	Қартбай Е.Ғ., Тынарбай Н.И. «MITM шабуылы (адамның ортадағы шабуылы)»	1388
313.	Маратов Б.Ж. «Әлеуметтік инженерия қауіпсіздікке қатер ретінде: қызметкерлерді қорғау және оқыту әдістері»	1393
314.	Мағзумов А.М. «Websocket протоколындағы осалдықтарды талдау»	1397
315.	Майданов А.С. «Автоматизация процесса анализа оперативной памяти с использованием python»	1401
316.	Мақсат Ә., Нурсейтов С. «Блокчейн қажеттілік пе, әлде сән бе?»	1406
317.	Қ. Мырзағалиұлы. «Инциденттерді анықтауда желілік логтарды талдаудың маңызды рөлдері»	1409
318.	Нурбатуров С.К. «Интеграция honeypot в ит-инфраструктуру компании»	1412
319.	Нуриева Д.Р., Исайнова А.Н. «Анализ рисками безопасности данных в медицинских учреждениях»	1415
320.	Нұрлан А.Т. «Кескіндердегі статистикалық стегоанализ әдістері»	1420
321.	Оралбеков Е.А. «Ddos-шабуылдардың жаңа буыны»	1424

322.	Рамазанова Ж, Нұрлан А, Жайсанбаева А. «Бұлтты технологияларды пайдалану кезіндегі тәуекелдер мен қауіпсіздік шараларын зерттеу»	1430
323.	Сахатбекқызы Т., Бахтиярқызы Т.А. «IoT құрылғыларының желідегі қауіпсіздігін қалай қамтамасыз етуге болады: стратегиялар және packet tracer көмегімен модельдеу»	1434
324.	Серғазы М. «Повышение производительности разработчиков с помощью интегрированных искусственных интеллектов и соображения кибербезопасности»	1440
325.	Султанов А.М. «Стеганография в кибербезопасности казахстана»	1443
326.	Танатаров Е., Іргебай С., Султанов А. «WI-FI желісінде шақырылмаған қонақтарды автоматты түрде анықтау жүйесі»	1447
327.	Таубай М.Е. Рамагуллаев Ә.А. «Фишинг: желідегі beef әдісі арқылы алдау және одан сақтану»	1452

СЕКЦИЯ 3 ЕСТЕСТВЕННЫЕ НАУКИ

		ПОДСЕКЦИЯ 3.1 АКТУАЛЬНЫЕ ПРОБЛЕМЫ СОВРЕМЕННОЙ БИОЛОГИИ	
328.	Акимкара А.Б.	Гербарийдің ботаникалық зерттеулерде қолданылуы және гербарий қорындағы кеппе шөптің қалыптасу ерекшеліктері	1457
329.	Ақылбек А.	Астана қаласындағы ботаникалық бағының ландшафттағы <i>geranium sanguineum</i> биологиялық ерекшеліктеріне сипаттама беру	1459
330.	Әділхан Ж.	Мобильді байланыс пен қолданбалардың адамның мінез-құлқына әсерін анықтау	1463
331.	Базарбаева Қ.	Жасөспірімдерде девиантты мінез-құлықтың даму қаупі	1467
332.	Байдосова А.Б.	Методика использования игровых технологий на уроках биологии	1471
333.	Байдосова А.Б.	Актуальные проблемы современной биологии с использованием игровых технологий в образовании	1474
334.	Ғазизова Ә.	Сәулеленген егеуқұйрықтардың бүйректеріндегі морфофункционалдық өзгерістерді салыстырмалы бағалау	1477
335.	Еркін З.Б.	Биология сабақтарында оқушылардың сыни ойлау қабілетін жетілдіруде блум таксономиясын пайдалану	1482
336.	Жанабергенова	Кенеттен жүрек өлімі: генетикалық аспектілері	1486

**ТЕРЕҢ ОҚЫТУ АРҚЫЛЫ КІТАП ҰСЫНЫСТАРЫН ӘЗІРЛЕУ:
COLLABORATIVE FILTERING, CONTENT-BASED ЖӘНЕ NLP
ӘДІСТЕРІНІҢ КОМБИНАЦИЯСЫ**

Смаилова Назгүл Батырбекқызы

smailovanazgul58@gmail.com

«Жасанды интеллект технологиялары» кафедрасының студенті,
Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан
Ғылыми жетекшісі – У.К. Турусбекова

Аңдатпа. Электрондық коммерция, ағындық платформалар және электрондық кітапханалар сияқты салаларда дербестендірілген мазмұнды қамтамасыз етуге мүмкіндік беретін ұсыныс жүйелері цифрлық дәуірде маңызды бөлігіне айналды. Кітап ұсыныстары контекстінде бұл жүйелер пайдаланушыларға олардың қалауы мен мінез-құлық үлгілерін пайдалана отырып, тиісті атауларды табуға көмектеседі. Цифрлық контенттің үнемі өсіп келе жатқан көлеміне байланысты кітап таңдау процесін автоматтандыратын, пайдаланушы тәжірибесін жақсартатын және ұсыныстардың сапасын арттыратын интеллектуалды алгоритмдерді құру мәселесі өзекті болып табылады.

Ұсынылған жұмыстың мақсаты - пайдаланушылардың жеке қалауларын ескере отырып, ең өзекті жұмыстарды ұсына алатын терең оқытуға негізделген дербестендірілген кітап ұсыныстарының үлгісін жасау.

Түйін сөздер: ұсыныс жүйелері, терең оқыту, кітап ұсынысы

Кіріспе. Соңғы жылдары ұсыныс жүйелері ақпараттық технологиялардың негізгі элементіне айналды, әсіресе кітаптар мен медиа-мазмұнды таңдау саласында. Бұл жұмыстың мақсаты бірлескен сүзу (CF) және мазмұнды сүзу (CBF) сияқты дәстүрлі әдістердің бар шектеулерін жеңе алатын тиімді гибриді ұсыныс үлгісін әзірлеу болып табылады. Зерттеудің міндеті-пайдаланушылардың элементтермен өзара әрекеттесуін ескеретін алгоритм құру. Сонымен қатар, алгоритм дәлірек және жекелендірілген ұсыныстарды қамтамасыз ету үшін метадеректерді талдауы керек.

Жұмыстың өзектілігі деректердің сирек кездесуі мен суық бастау мәселелерін шеше алатын тиімді ұсыныс жүйелеріне деген қажеттіліктің артуына байланысты. Көп жағдайда пайдаланушылар қол жетімді кітаптардың аз ғана бөлігін бағалайды, бұл CF әдістерін қолдану үшін ақпараттың жетіспеушілігіне әкеледі. Сонымен қатар, жаңа пайдаланушылар мен жаңадан қосылған кітаптар тарихи деректердің болмауына байланысты тиісті ұсыныстар алуға қиындықтарға тап болады. CF және CBF біріктіретін гибриді әдістер бұл мәселелерді шешуді ұсынады, бірақ метадеректерді өңдеуде әлі де шектеулер бар.

Бұл жұмыста құрылымдалмаған метадеректерді өңдеу үшін BERT көмегімен бірлескен сүзуді, мазмұнға негізделген сүзуді және мәтінді кодтауды қамтитын терең оқытудың гибриді моделін ұсынамыз. Біздің көзқарасымыз әртүрлі деректер көздерін біріктіру арқылы ұсыныстардың дәлдігін арттыруға, суық бастау жағдайында өнімділікті жақсартуға және модельдегі мүмкіндіктердің маңыздылығын талдау арқылы түсіндірілетін ұсыныстарды беруге бағытталған.

Жұмыстың мақсатына жету үшін бірнеше негізгі міндеттерді шешу қажет:

- мазмұнды, бірлескен және гибриді сүзуді қоса алғанда, ұсыныстар жүйесін құрудың әртүрлі тәсілдерін талдау;

- ұсыныстар жүйесінде терең оқыту әдістерін қолдануды, атап айтқанда, пайдаланушылар мен кітаптардың жоғары сапалы көріністерін жасау үшін көп қабатты перцептрондарды (MLP) және embedding пайдалануды зерттеу;

- жүйені оқыту үшін стохастикалық градиент түсіру (SGD) және адаптивті алгоритмдер (Adam, RMSProp) сияқты оңтайландыру әдістерін қолдану.

Пайдаланылған деректердің сипаттамасы. Зерттеу екі негізгі деректер жиынтығын пайдаланады: кітаптар (books), рейтингтер (ratings). Бұл мәліметтер кітаптардың анықтамалық жүйесін құру үшін қажет.

Кітаптар туралы мәліметтер 1-суретте көрсетілген - деректер жиынтығы кітаптар туралы ақпаратты қамтиды.

ID	Названия	Цена	Рейтинг	Сложность книги	Обложка	Автор	Издательство	ISBN	Кол-во страниц	Год издания	Язык
1	Керімбай С., Құспан А., Бейсен	2 455 T	4.35	410		Керімбай С., Құспан А.	"Обшасы"	978-6-01-...	204	2021	қазақш
2	Санжар Керімбай, Әділбек Нәби	2 995 T	4.88	327		Санжар Керімбай	"Обшасы"	978-601-7...	254	2023	қазақш
3	Уағым ROOTEST Жыл Қоғамы	4 990 T	4.52	185		ROOTEST	Rootest.kz	978-601-8...	216	2024	қазақш
4	Құрманғали Тұрмағалиұлы Рау...	4 024 T	4.96	431		Құрманғали Құрма...	AYAT SYRL	978-601-0...	300	2023	қазақш
5	Берік Жұбаналиев Қоғамстан т...	1 960 T	4.89	32		Берік Жұбаналиев	Азамат	976-519-2...	212	2024	қазақш
6	Гәлімбек Тобылан: Құран аян...	1 700 T	4.89	171		Гәлімбек Тобылан	Самға	978-601-7...	80	2010	қазақш
7	Сейітбаев Л., Аманжолбаев Р., С...	4 688 T	4.52	321		Ларра Сейітбаев, Р...	Арыс	9786-016...	283	2021	қазақш
8	Мырзабай С., Балқаштан дана...	2 400 T	4.84	138		Мырзабай С.	"Обшасы"	978-601-7...	224	2021, 2022	қазақш
9	Четви Ф., Гайбег Бақса болса...	2 300 T	4.8	254		Четви Ф.	Танғы	978-601-7...	128	2022	қазақш
10	Жұманбаев М., Павловский	3 400 T	4.94	90		Мажан Жұманбаев	Майн Ривк	978-601-3...	108	2022	қазақш
11	Сәкен Смағұл: Әр күннің ас...	1 960 T	4.58	254		Сәкен Смағұл	Танғы	978-601-7...	376	2023	қазақш
12	Қарғы и атасы Ермұхамед С...	1 100 T	4.99	32		Ермұхамед С. В.	888	46702489...	56 + 32	2023	қазақш
13	Ишан Рахым Жүрек сөздері	5 750 T	4.88	288		Ишан Рахым	Самға	978-601-7...	488	2021	қазақш
14	Аманат САМЖӨ: Мей. Сөздер...	4 717 T	4.99	94		Аманат САМЖӨ	QALYNA	97860108...	198	2024	қазақш
15	Жандарбек Бүткел: Сөз жет...	5 950 T	4.93	168		Жандарбек Бүткел	Самға	97860176...	218	2021	қазақш
16	Сәкен Д., Шыңғыс мейр	3 988 T	4.88	74		Сәкен Д.	Малғұм	978-601-8...	351	2023	қазақш
17	Пәлбике Айбек Сөзшілбек Ду...	1 950 T	4.87	34		Айбек Сөзшілбек	Азамат	97860126...	194	2024	қазақш
18	Асан Оразабай: Сөз жеткі...	4 313 T	4.97	201		Асан Оразабай	Самға	978-601-8...	190	2024	қазақш

1-сурет. Books деректер жиынтығы

Рейтинг деректері - бұл деректер жиынтығы пайдаланушылардың кітаптарды қандай бағаға бағалағаны туралы ақпаратты қамтиды (2-сурет).

Название книги	Имя пользователя	Оценка
Керімбай С., Құспан А., Бейсен және болмыс	Жұлдыз	5
Керімбай С., Құспан А., Бейсен және болмыс	Гауһаргас	5
Керімбай С., Құспан А., Бейсен және болмыс	Айнұр	5
Керімбай С., Құспан А., Бейсен және болмыс	Архат	5
Керімбай С., Құспан А., Бейсен және болмыс	Улжалғас	5
Керімбай С., Құспан А., Бейсен және болмыс	Дастан	5
Керімбай С., Құспан А., Бейсен және болмыс	Гульмира	5
Керімбай С., Құспан А., Бейсен және болмыс	Шыңғыс	5
Керімбай С., Құспан А., Бейсен және болмыс	Айым	4
Санжар Керімбай, Әділбек Нәби: Қуыскеуде	Мәдина	5
Санжар Керімбай, Әділбек Нәби: Қуыскеуде	Ахерке	5
Санжар Керімбай, Әділбек Нәби: Қуыскеуде	Ләйла	5
Санжар Керімбай, Әділбек Нәби: Қуыскеуде	Дінмұхаммед	5

2-сурет. Ratings деректер жиынтығы

Деректерді түрлендіру: BERT көмегімен кітаптардың метадеректерін токенизациялау, TF-IDF матрицасын құру. Деректерді түрлендіру екі негізгі кезеңді қамтиды: BERT көмегімен кітап метадеректерін токенизациялау және TF-IDF матрицасын құру. Кітаптардың атауы, авторы, жанры және сипаттамасы сияқты мәтіндік сипаттамаларымен жұмыс істеу үшін алдын ала оқытылған BERT тілдік моделі қолданылды. BERT мәтіндік деректерді эмбединг түрінде тиімді ұсынуға, содан кейін олар ұсыныс жүйесінде қолдануға мүмкіндік береді [1]. Токенизация процесі Transformers кітапханасының BertTokenizer көмегімен жүзеге асырылды, онда кітаптың әрбір мәтіндік сипаттамасы токенизация арқылы өтіп, токен индекстеріне айналды және одан әрі тығыз, тұрақты өлшемді векторларға айналды. Кейін бұл векторлар BERT моделіне берілді, ол кітаптың семантикалық мағынасын білдіретін соңғы жасырын қабаттың эмбедингтерін шығарды [2].

BERT эмбедингтерін алғаннан кейін формула бойынша есептелген TF-IDF (term Frequency-Inverse Document Frequency) көмегімен метадеректерді векторлау жүргізілді:

$$TF - IDF(t, d, D) = TF(t, d) \times IDF(t, D) = \frac{f_{t,d}}{\sum_{t \in d} f_{t,d}} \times \log \left(\frac{N}{1 + |\{d \in D: t \in d\}|} \right),$$

мұндағы $f_{t,d}$ - d құжатындағы t терминінің жиілігі, N - D корпусындағы құжаттардың жалпы саны, ал IDF бөлгіші t термині бар құжаттардың санын білдіреді [3].

TF-IDF (Term Frequency-Inverse Document Frequency) жанрларды, авторларды және кітап атауларын талдау үшін қолданылды, өйткені ол деректер корпусындағы ең маңызды терминдерді бөліп көрсетуге мүмкіндік береді, бұл жиі кездесетін, бірақ маңызды емес сөздердің әсерін азайтады [3]. Векторлау `sklearn.feature_extraction.text` - `Termin TfIdfVectorizer` көмегімен жүзеге асырылды, онда мәтінді алдын-ала тазарту, тоқтату сөздерін жою және лемматизация жүргізілді. Алынған TF-IDF векторлары BERT эмбедингтерімен біріктіріліп, кейіннен ұсыныс жүйесін оқыту үшін пайдаланылатын бірыңғай кіріс матрицасын құрады.

Гибридті модельді әзірлеу және оқыту. Гибридті модель кітаптарды ұсыну үшін бірлескен және мазмұнды сүзуді біріктіреді. Бұл пайдаланушылардың қалауын (бірлескен тәсіл) және кітаптардың семантикалық ерекшеліктерін (мазмұндық тәсіл) ескеруге мүмкіндік береді. Модель бірнеше негізгі компоненттерді қамтиды: пайдаланушылар мен кітаптардың эмбедингі, мәтіндік деректерді өңдеуге арналған LSTM және қайта оқытудың алдын алу үшін реттеу механизмдері (Dropout, Batch Normalization) және Early Stopping. Бұл тәсіл ұсынымдық жүйелерді құрудағы заманауи тенденцияларға сәйкес келеді, мұнда гибридті әдістер жеке сүзу әдістерімен салыстырғанда жоғары дәлдікті көрсетеді. Модельді оқыту Adam оңтайландырғышын және MSE жоғалту функциясын қолдану арқылы жүзеге асырылады, бұл ұсыныс жүйелеріндегі регрессия тапсырмалары үшін стандартты. Валидация үшін Early Stopping қолданылады, ол валидация үлгісіндегі шығындарды бақылайды және 5 дәуір бойы жақсартулар болмаған кезде оқуды тоқтатады. Бұл әдіс нейрондық желілерді қайта оқытудың алдын алуда және оқытуды жеделдетуде тиімді екенін дәлелдеді [4].

Гибридті модель архитектурасы: бірлескен және мазмұнды тәсілдерді біріктіру. Кодта енгізілген модель архитектурасы пайдаланушы деректерін, кітап идентификаторларын және мәтіндік метадеректерді өңдеуді нақты бөледі. Пайдаланушылар мен кітаптар модельмен бірге оқытылатын эмбединг түрінде ұсынылған. Бұл нейрондық бірлескен сүзгіге сәйкес келеді, мұнда пайдаланушылар мен заттардың эмбедингі деректердегі жасырын қатынастарды анықтауға мүмкіндік береді [5]. Кітаптардың мәтіндік сипаттамалары бөлек өңделеді: алдымен олар BERT көмегімен таңбаланады, содан кейін эмбедингке айналады және LSTM қабатына беріледі. Бұл тәсіл сипаттамалардағы сөздердің контексті мен реттілігін ескеруге мүмкіндік береді, бұл әсіресе жанрлар мен тақырыптарды талдау үшін маңызды. Соңғы кезеңде белгілер біріктіріледі (біріктіру) және модель Dropout және Batch Normalization толық байланысқан қабаттар сериясынан өтеді.

Пайдаланушылардың эмбедингтері мен кітаптарын пайдалану. Кодта пайдаланушы мен кітап эмбедингтері дискретті идентификаторларды тығыз, бекітілген өлшемді векторларға түрлендіретін эмбединг қабаты арқылы жасалады. Бұл тәсіл нейрондық желінің бірлескен сүзгі идеяларына негізделген, мұнда әрбір пайдаланушы және әрбір элемент ортақ векторлық кеңістікте ұсынылады [5]. Мұндай эмбедингтерді оқыту модельге жасырын қатынастарды автоматты түрде анықтауға мүмкіндік береді, мысалы, ұқсас қалаулары бар пайдаланушыларды немесе бір жанрдағы кітаптарды топтастыру.

Модельдің нәтижелері мен талдауы

Модельдің сапасын метрика бойынша бағалау. Модель болжамдардың жоғары дәлдігін көрсетеді, бұл негізгі көрсеткіштермен расталды:

RMSE (Root Mean Square Error): 0.3644 - Төмен `rmse` мәні болжамды бағалаулардың нақты бағалардан минималды ауытқуын көрсетеді. Бұл әсіресе кішігірім қателер пайдаланушы тәжірибесіне әсер етуі мүмкін ұсыныс жүйелері үшін өте маңызды (3-сурет).

MAE (Mean Absolute Error): 0.1509 - 0.2-ден төмен MAE модельдің жоғары тұрақтылығын көрсетеді. Салыстыру үшін, оқытудың басында (Epoch 1) MAE 4.7098 болды, бұл оңтайландырудың тиімділігін көрсетеді (3-сурет).

R2 (анықтау коэффициенті): 0.5928 - R2 мәні 0.6-ға жақын, бұл деректердің дисперсиясының шамамен 60% түсіндіруді білдіреді. Бұл гетерогенді деректермен жұмыс істейтін гибридті модельдер үшін жақсы нәтиже болып табылады.

Accuracy (қателік ± 0.5): 93.60% - Модель 93.6% - бағалауды дұрыс болжайды(3-сурет).

Accuracy (с погрешностью ± 0.5): 0.9360 (или 93.60%)

Оценка качества модели:

RMSE: 0.3644

MAE: 0.1509

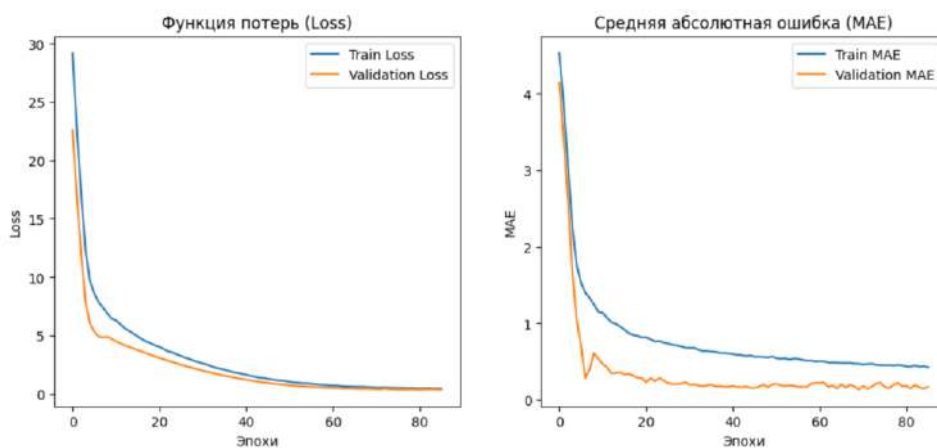
R²: 0.5928

3-сурет. Модельдің сапасын метрика бойынша бағалау

Оқу динамикасын талдау. Жоғалту және қате графиктері (4-сурет) тұрақты конвергенцияны көрсетеді:

Train Loss 31.2732 — ден (Epoch 1) 0.4200-ге (Epoch 86), ал validation Loss 22.5647-ден 0.3623-ке дейін төмендеді. Алғашқы 10 дәуірдегі шығындардың күрт төмендеуі (мысалы, val_loss 22.56-дан 4.88-ге дейін) модельдің деректерге тез бейімделуін көрсетеді. Epoch 20-дан кейін құлдырау баяулады, бұл дәл баптау кезеңіне тән.

Val MAE Epoch 70-те 0.1362 минимумға жетті, бұл Epoch 86-да оқуды тоқтатқан Early Stopping тиімділігін растайды (ең жақсы салмақ Epoch 81-мен қалпына келтірілді).



4-сурет. Loss және MAE графиктері

Қорытынды: Бірлескен және мазмұнды сүзуді біріктіретін ұсыныс жүйесінің гибриді моделі жоғары тиімділікті көрсетті. Мәтіндік метадеректерді токенизациялау үшін BERT және реттілікті талдау үшін LSTM пайдалану кітаптардың семантикалық ерекшеліктерін де, пайдаланушылардың қалауын да ескеруге мүмкіндік берді. Реттеуді (Dropout, Batch Normalization) және Early Stopping көмегімен модельді оқыту тұрақты конвергенцияны қамтамасыз етті және қайта оқытудың алдын алды.

Негізгі көрсеткіштер (MSE = 0.3644, MAE = 0.1509, Accuracy = 93.6%) модельдің дәлдігін растайды, бұл оны практикалық іске асыруға жарамды етеді. Ең жақсы нәтижелерге пайдаланушы эмбеддингтері мен кітаптардың үйлесімі және мәтіндік деректерді оңтайландырылған өңдеу арқылы қол жеткізілді.

Қолданылған әдебиеттер тізімі

1. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Minneapolis, Minnesota, 2019. P.4171-4186.
2. А.М. Мансур. Алгоритм на основе трансформеров для классификации длинных текстов // Известия ЮФУ. Технические науки. 2024. №3. С.186-196.
3. Paul Sheridan, Mikael Onsjö. The hypergeometric test performs comparably to TF-IDF on standard text analysis tasks // Multimedia Tools and Applications. 2023. №83(10). P.1-16. DOI:10.1007/s11042-023-16615-z
4. Eric Ziebell, Ratmir Miftachov, Bernhard Stankewitz, Laura Hucker. EarlyStopping: Implicit Regularization for Iterative Learning Procedures in Python // arXiv:2503.16753[stat.ML]. 2025. <https://doi.org/10.48550/arXiv.2503.16753>
5. Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, Tat-Seng Chua. Neural Collaborative Filtering // WWW '17: Proceedings of the 26th International Conference on World Wide Web: Perth Australia April 3 – 7, 2017. P.173-182.