



№3 (5) 2024

INTERNATIONAL
SCIENCE REVIEWS



Natural Sciences and
Technologies series





INTERNATIONAL SCIENCE REVIEWS

Natural Sciences and Technologies series

Has been published since 2020

№3 (5) 2024

Astana

INTERNATIONAL SCIENCE REVIEWS. NATURAL SCIENCES AND TECHNOLOGIES SERIES ЖУРНАЛЫНЫҢ РЕДАКЦИЯСЫ

БАС РЕДАКТОР

Қалимолдаев Мақсат Нұрадилович, техникалық ғылымдар докторы, ҚР ҰҒА академигі, профессор, ҚР ҒЖБМ ҒК «Ақпараттық және есептеу технологиялары институты» бас директорының кеңесшісі, бас ғылыми қызметкері (Қазақстан)

БАС РЕДАКТОРДЫҢ ОРЫНБАСАРЫ

Мырзағалиева Анар Базаровна, биология ғылымдарының докторы, профессор, бірінші вице-президент, Астана халықаралық университеті (Қазақстан);

РЕДАКТОРЛАР:

- **Сейткан Айнура Сейтканқызы**, техника ғылымдарының кандидаты, PhD, жаратылыстану ғылымдары жоғары мектебінің деканы, Астана халықаралық университеті (Қазақстан);

- **Муканова Асель Сериковна**, PhD, Ақпараттық технологиялар және инженерия жоғары мектебінің деканы, Астана халықаралық университеті (Қазақстан);

- **Абдилдаева Асель Асылбековна**, PhD, қауымдастырылған профессор, Әл-Фараби атындағы ҚазҰУ (Қазақстан);

- **Хлахула Иржи** PhD, профессор, Познаньдағы Адам Мицкевич атындағы университет (Польша);

- **Редферн Саймон А.Т.**, PhD, профессор, Наньян технологиялық университеті (Сингапур);

- **Сяoley Фенг**, PhD, Наньян технологиялық университеті (Сингапур);

- **Шуджаул Мулк Хан**, PhD, профессор, Каид-және-Азам университеті (Пакистан);

- **Базарнова Наталья Григорьевна**, химия ғылымдарының докторы, профессор, Химия және химиялық-фармацевтикалық технологиялар институты (Ресей);

- **Черёмушкина Вера Алексеевна**, биология ғылымдарының докторы, профессор, РҒА СБ Орталық Сібір ботаникалық бағы (Ресей);

- **Тасболатұлы Нұрболат**, PhD, Ақпараттық технологиялар және инженерия жоғары мектебі деканының орынбасары, Астана халықаралық университеті (Қазақстан);

- **Байшоланов Сакен Советович**, география ғылымдарының кандидаты, доцент, Астана халықаралық университеті (Қазақстан);

- **Нуркенов Серик Амангельдинович**, PhD, қауымдастырылған профессор, Астана халықаралық университеті (Қазақстан).

**РЕДАКЦИЯ ЖУРНАЛА INTERNATIONAL SCIENCE REVIEWS.
NATURAL SCIENCES AND TECHNOLOGIES SERIES**

ГЛАВНЫЙ РЕДАКТОР

Калимолдаев Максат Нурадилович, доктор технических наук, академик НАН РК, профессор, ГНС, советник генерального директора Института информационных и вычислительных технологий КН МНВО РК (*Казахстан*)

ЗАМЕСТИТЕЛЬ ГЛАВНОГО РЕДАКТОРА

Мырзагалиева Анар Базаровна, доктор биологических наук, профессор, первый вице-президент, Международный университет Астана (*Казахстан*)

РЕДАКТОРЫ:

- **Сейткан Айнур Сейтканкызы**, кандидат технических наук, PhD, декан высшей школы естественных наук, Международный университет Астана (*Казахстан*);

- **Муканова Асель Сериковна**, PhD, декан Высшей школы информационных технологии и инженерии, Международный университет Астана (*Казахстан*);

- **Абдилдаева Асель Асылбековна**, PhD, ассоциированный профессор, Казахский национальный университет имени Аль-Фараби (*Казахстан*);

- **Хлахула Иржи** PhD, профессор, Университет имени Адама Мицкевича в Познани (*Польша*);

- **Редферн Саймон А.Т.**, PhD, профессор, Наньянский технологический университет (*Сингапур*);

- **Фенг Сяoley**, PhD, Наньянский технологический университет (*Сингапур*);

- **Шуджаул Мулк Хан**, PhD, профессор, Университет Каид-и Азама (*Пакистан*);

- **Базарнова Наталья Григорьевна**, доктор химических наук, профессор, Институт химии и химико-фармацевтических технологий (*Россия*);

- **Черёмушкина Вера Алексеевна**, доктор биологических наук, профессор, Центральный Сибирский Ботанический сад СО РАН (*Россия*);

- **Тасболатұлы Нұрболат**, PhD, заместитель декана Высшей школы информационных технологий и инженерии, Международный университет Астана (*Казахстан*);

- **Байшоланов Сакен Советович**, кандидат географических наук, доцент, Международный университет Астана (*Казахстан*);

- **Нуркенов Серик Амангельдинович**, PhD, ассоциированный профессор, Международный университет Астана (*Казахстан*);

**EDITORIAL TEAM OF THE JOURNAL INTERNATIONAL SCIENCE REVIEWS.
NATURAL SCIENCES AND TECHNOLOGIES SERIES**

CHIEF EDITOR

Maksat Kalimoldayev, Doctor of Technical Sciences, Academician of NAS RK, Professor, SRF, CEO's counselor «The Institute of Information and Computational Technologies» CS MSHE RK (Kazakhstan)

DEPUTY CHIEF EDITOR

Anar Myrzagaliyeva, Doctor of Biological Sciences, Professor, First Vice-President, Astana International University (Kazakhstan)

EDITORS:

- **Ainur Seitkan**, Candidate of Technical Sciences, PhD, Dean of the Higher School of Natural Sciences, Astana International University (Kazakhstan);
- **Assel Mukanova**, PhD, Dean of the Higher School of Information Technology and Engineering, Astana International University (Kazakhstan);
- **Assel Abdildayeva**, PhD, Associate Professor, of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University (Kazakhstan);
- **Jiri Chlachula**, PhD, Dr.Hab., Full Professor, Adam Mickiewicz University, Poznań (Poland);
- **Simon A.T. Redfern**, PhD, Professor, Nanyang Technological University (Singapore);
- **Xiaolei Feng**, PhD, Nanyang Technological University (Singapore);
- **Khan Shujaul Mulk**, PhD, Professor, Quaid-i-Azam University (Pakistan);
- **Natal'ya Bazarnova**, Doctor of Chemical Sciences, Professor, Institute of Chemistry and Chemical-Pharmaceutical Technologies (Russia);
- **Vera Cheryomushkina**, Doctor of Biological Sciences, Professor, Central Siberian Botanical Garden SB RAS (Russia);
- **Nurbolat Tasbolatuly**, PhD, Deputy Dean of the Higher School of Information Technology and Engineering, Astana International University (Kazakhstan);
- **Saken Baisholanov**, Candidate of Geographical Sciences, Associate Professor, Astana International University (Kazakhstan);
- **Serik Nurkenov**, PhD, Associate Professor, Astana International University (Kazakhstan).

Editorial address: 8, Kabanbay Batyr avenue, of.316, Nur-Sultan,
Kazakhstan, 010000

Tel.: (7172) 24-18-52 (ext. 316)

E-mail: natural-sciences@aiu.kz

International Science Reviews NST - 76153

International Science Reviews

Natural Sciences and Technologies series

Owner: Astana International University

Periodicity: quarterly

Circulation: 500 copies

CONTENT

1. С.Е.Базаров, А.М.Касымханов, С.Б.Ниғметжанов, А.А.Аманжолов БҰҚТЫРМА СУ ҚОЙМАСЫНДАҒЫ КӘСІПШІЛІК МАҢЫЗЫ БАР ЖЫРТҚЫШ БАЛЫҚ ТҮРЛЕРІНІҢ ЭКОЛОГИЯЛЫҚ-БИОЛОГИЯЛЫҚ СИПАТТАМАСЫ.....	7
2. А.М. Касымханов СОВРЕМЕННОЕ ГИДРОХИМИЧЕСКОЕ СОСТОЯНИЕ ОЗЕРА ЖАЙСАН.....	12
3. Е.К.Жаксылыкова, С.Б. Абеуова ВНЕДРЕНИЕ ЭКОЛОГИЧЕСКИХ ПРАЗДНИКОВ: ШАГ К ЗЕЛЕНОМУ БУДУЩЕМУ	19
4. Ж.Балтабай ОРТА МЕКТЕПТЕ БИОЛОГИЯНЫ ОҚЫТУДЫҢ ЗЕРТТЕУ ӘДІСТЕРІ	24
5. M.Karibayeva ABOUT THE CHOICE OF DRINKING WATER PURIFICATION METHODS	30
6. А.Б.Карабалаева, Т.Абуова, А.Шәлтік, М.Ермек БИОДЕГРАДАЦИЯ (УТИЛИЗАЦИЯ) БЫТОВОГО ПЛАСТИКА МИКРОМИЦЕТАМИ	37
7. С.Б. Марзен, А.Е.Сулейменова М ИЗУЧЕНИЕ ОСОБЕННОСТЕЙ ПРИМЕНЕНИЯ БИОЛОГИЧЕСКИ АКТИВНОГО ВЕЩЕСТВА ДЗ	44
8. Г.Қ. Ситахметова, А.М. Касымханов, А.А. Аманжолов СОСТАВ И РАСПРЕДЕЛЕНИЕ АКТИВНОЙ МОЛОДИ РЫБ ПО РЕКЕ ЕРТИС	48
9. А. Б. Шуакбаева АУЫЗ СУДЫҢ ФИЗИКА-ХИМИЯЛЫҚ ЖӘНЕ МИКРОБИОЛОГИЯЛЫҚ КӨРСЕТКІШТЕРІ	54
10. Т.А.Бейсен, О.К.Суйинханов, А.Н.Сұлтанғазиева БІЛІМ БЕРУДЕП ЖАСАНДЫ ИНТЕЛЛЕКТ: АРТЫҚШЫЛЫҚТАРЫ МЕН БОЛАШАҒЫ.....	61
11. Т.А.Шалбай, М.Ж.Қалдарова МӘТІН БОЙЫНША АДАМНЫҢ КӨҢІЛ- КҮЙІН ТАЛУҒА ӘРТҮРЛІ ТӘСІЛДЕРДІ ТАЛДАУ	67
12. Ж.Семейхан, М.Ж.Калдарова ФРЕЙМВОРКИ ДЛЯ FRONTEND- РАЗРАБОТКИ: СРАВНИТЕЛЬНЫЙ АНАЛИЗ	78
13. Тишбаева А.Д., Рыспаева Д.С. ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕДИЦИНЕ КАЗАХСТАНА	88
14. Жәніс Д.Е., Бурибаева А.К. ЖАСАНДЫ ИНТЕЛЛЕКТ ЖӘНЕ ОНЫҢ ҚОҒАМҒА ӘСЕРІ: ЖИ ЕНГІЗУДІҢ ЭТИКАЛЫҚ, ЭКОНОМИКАЛЫҚ ЖӘНЕ ӘЛЕУМЕТТІК АСПЕКТІЛЕРІН ЗЕРТТЕУ	93
15. А.Т. Баймуханбетова ПЕРВЫЙ ШАГ В PYTHON ДЛЯ АНАЛИЗА ДАННЫХ: ИСПОЛЬЗОВАНИЕ БИБЛИОТЕК PANDAS И MATPLOTLIB	101

МӘТІН БОЙЫНША АДАМНЫҢ КӨҢІЛ-КҮЙІН ТАЛУҒА ӘРТҮРЛІ ТӘСІЛДЕРДІ ТАЛДАУ

Т.А.Шалбай, М.Ж.Қалдарова

Астана халықаралық университеті, Астана, Қазақстан
e-mail: takhminashh@gmail.com

Аңдатпа. Мақалада маркетингтік талдаулардан бастап әлеуметтік медиа мониторингіне дейінгі әртүрлі салаларда туындайтын мәтін бойынша көңіл-күйді талдаудың өзекті мәселесі қарастырылады. Жұмыс бұл мәселені шешудің әртүрлі тәсілдерін, соның ішінде дәстүрлі статистикалық әдістерді, машиналық оқытуды (SVM) және терең оқытуды талдайды. Олардың артықшылықтары мен кемшіліктері контекстті, тілдің түрлі ерекшеліктерін және іске асырудың күрделілігін ескеру қабілетіне назар аудара отырып салыстырылады. Сонымен қатар, дәлдік пен тұрақтылықты арттыру үшін әртүрлі тәсілдерді біріктіретін гибриді модельдер зерттелуде. Әртүрлі зерттеулерді талдау әрбір тәсілдің күшті және әлсіз жақтары бар екенін және оңтайлы шешімді таңдау нақты тапсырма мен деректер сипаттамаларына байланысты екенін көрсетеді.

Кілтті сөздер: Көңіл-күйді талдау, мәтінді талдау, табиғи тілді өңдеу (NLP), машиналық оқыту, терең оқыту, статикалық әдістер, гибриді әдістер.

КІРІСПЕ

Әлеуметтік желілерді қолданушылардың санының артуы маңызды ақпаратты алу үшін түрлі электрондық ақпараттарды зерттеу қажеттіліктерін тудырды [1]. Мұндай кең деректер жиынтығын дәстүрлі қолмен зерттеу көп уақытты қажет етеді. Оның ауқымын ескергенде, бұл тіпті іс жүзінде мүмкін емес. Осыған байланысты мәтіннің сезімдік, эмоционалды фонның автоматты талдау өте пайдалы құрал болып табылады [2]. Ол өнімдер, қызметтер, оқиғалар және олардың сипаттамалары туралы Интернеттегі қоғамдық пікірге сүйене отырып, мәтіндердегі эмоциялар мен қатынастарды өлшеу үшін алгоритмдік есептеулерді қолданады. Бұл міндет табиғи тілді өңдеуде өте маңызды [3].

Қазіргі әлемдегі ақпаратпен өзара әрекеттесу көбінесе мәтіннің эмоционалды бояуын бағалауды қамтиды. Мәтін авторының суреттелетін объектіге қандай көзқарас білдіретінін анықтау қажет - оң, теріс немесе бейтарап. Цветков пен Рыбанов дуэті [4] фильмге шолулардың эмоционалды реңктерін талдауға арналған бағдарламалық қамтамасыз ету жүйесін әзірлеп, дәл осы тапсырманы орындады. Олардың жүйесі шолулардағы қатынастың үш түрін сәтті анықтайды: теріс, бейтарап және оң. Бұл мысал эмоционалды ақпаратты талдау негізінде неғұрлым негізделген шешімдер қабылдауға көмектесетін нақты өмірлік жағдайларда көңіл-күйді автоматты талдаудың практикалық мәнін көрсетеді.

Бұл жұмыстың мақсаты – көңіл-күйді автоматты түрде талдаудың қолданыстағы әдістеріне шолу және талдау жүргізу, олардың тиімділігін бағалау және одан әрі зерттеу үшін перспективалық бағыттарды анықтау.

Жұмыста көңіл-күйді анықтаудың әдістері қарастырылады, олардың тиімділігі бағаланады және одан әрі дамудың перспективалы бағыттары іздестіріледі.

Семантикалық әдістерді, машиналық оқытуды, терең оқытуды және гибридік әдістерді қоса, олардың күшті және әлсіз жақтарын талдай отырып, әртүрлі тәсілдер қарастырылады. Гибридік әдістерге ерекше назар аударылады, өйткені олар көңіл-күйді талдаудың дәлірек және жан-жақты үлгісін уәде етеді.

Көңіл-күйді талдау адамдардың оқиғаларға, өнімдерге, қызметтерге және саяси шешімдерге қатынасы туралы құнды ақпарат беретін маркетингтен денсаулық сақтау саласына дейінгі әртүрлі салаларда қолдануды табуда.

Мәтін арқылы көңіл-күйді тану – сөздерді, сөз тіркестерін және стилистикалық ерекшеліктерді талдау арқылы мәтін авторының эмоционалды күйін автоматты түрде анықтауға мүмкіндік беретін технология. Бұл технология адамдардың оқиғаларға, өнімдерге, қызметтерге және саяси шешімдерге деген көзқарасы туралы құнды ақпарат беретіндіктен, оған сұраныс әртүрлі салаларда өсуде [5].

Маркетингте көңіл-күйді талдау компанияларға өз тұтынушыларын жақсы түсінуге көмектеседі. Өнімдер мен қызметтерге шолуларды талдау арқылы маркетингтік проблемаларды анықтай алады, өнім сапасын жақсартады және тұтынушылардың наазрын аударады. Заманауи ақпараттық технологияларды пайдалану үлкен көлемдегі мәтіндерді көңіл-күй үшін автоматты түрде талдауға мүмкіндік береді. Мұндай зерттеулердің мысалы ретінде [6] әлеуметтік желілердегі тауарларға немесе қызметтерге шолулардың көңіл-күйін талдау болып табылады.

Әлеуметтік желілерде қоғамдық пікірді бақылау үшін де осы технологиялық талдау қолданылады [7]. Бұл адамдардың оқиғаларға реакциясын қадағалауға, проблемаларды талқылау үрдістерін анықтауға, тіпті болу ықтималдығы жоғары қақтығыстарды болжауға мүмкіндік береді. Бұған қоса, көңіл-күйді талдау өшпенділіктің, интернеттік қақтығыстардың және өзара әрекеттесудің басқа жағымсыз түрлерінің таралуын болдырмауға көмектеседі, қауіпсіз және оңтайлы әлеуметтік қонысты жасайды.

Денсаулық сақтау саласында, көңіл-күйді талдау психикалық ауруларды ерте диагностикалауға көмектеседі. Осы тәсілді [8] қолданғандардың бірі – Пронин А. Г., ол пациенттердің мәтіндік хабарламаларын талдай отырып қолданды, мәселені шешудің ең қолайлы тәсілін зерттеді. Бұл еңбектің нәтижесінде дәрігерлер депрессия, мазасыздық және басқа психикалық бұзылулардың белгілерін анықтап, уақытында қажетті көмек көрсете алу жолын таба алады. Ол сондай-ақ кейбір

әдістер сынақ үлгілері бойынша жақсы нәтиже беретінін және психология саласында жасанды интеллектті пайдаланудың болашағы зор екенін атап өтті.

Саясатта, көңіл-күйді талдау саяси мәселелер бойынша қоғамдық пікірді бақылауға, сайлау нәтижелерін болжауға және саясаткерлердің сөйлеген сөздерін талдауға мүмкіндік береді, бұл саяси партиялар мен кандидаттарға тиімді шешімдер қабылдауға көмектеседі. Мысалы, Черничкин және Кривенко [9] Астрахан облысының әлеуметтік медиа деректерін талдау үшін PolyAnalyst платформасын қолданды. Пікірлердің эмоционалды фонын анықтауға басты назар аударылды, бұл оларға өзекті аймақтық мәселелерді анықтауға және азаматтардың аймақтық саяси қайраткерлерге деген көзқарасын талдауға мүмкіндік берді.

Білім берудегі көңіл-күйді талдау мұғалімдерге студенттердің оқу процесіне қалай жауап беретінін түсінуге және оны жеке қажеттіліктеріне бейімдеуге көмектеседі. Сонымен қатар, бұл технологияны зерттеушілер көркемдік әдебиет мазмұнын талқылауға да қолдануда. Мысалы, филология ғылымдарының кандидаты Пронин Е. В. «Біздің заманның қаһарманы» деген Лермонтовтың еңбегінің негізінде мәтіннің эмоционалды реңкінің талдау әдісін көркем шығарманы зерттеудің құралы ретінде қарастырған. [10]. Зерттеу барысында NLP табиғи тілді өңдеу әдісін қолданған. Көңіл-күй талдауында тілдерге шек жоқ. Тағы бір еңбек – өзбек тіліндегі мәтіндердің эмоционалды фонын мен поэтиялығын талдау. Ол үшін Барахнин және Менглиев алгоритмиялық, яғни статистикалық әдісті қолданған. [11]

Жалпы, мәтін арқылы көңіл-күйді тану әлеуметтік қарым-қатынасты жақсартуға, шешім қабылдауға, проблемаларды анықтауға және әртүрлі қызмет салаларында өмір сүру сапасын жақсартуға кең мүмкіндіктер ашады.

Бұл тақырыптың өзектілігі графикалық веб-интерфейстер сияқты оқытылатын модельдермен жұмыс істеуге арналған құралдардың дамуымен расталады. Мысалы, TensorFlow модулінің негізінде деректерді жіктеу моделі және keras қондырмасының көмегімен нейрондық желіні құру процесі зерттелді [12].

МӘТІННІҢ КӨҢІЛ-КҮЙІН ТАЛДАУ ӘДІСТЕРІ

Мәтіннің көңіл-күйін талдаудың әртүрлі тәсілдері бар, олардың әрқайсысының өзіндік артықшылықтары мен кемшіліктері бар:

Статистикалық әдістер сөздердің, өрнектердің және басқа лингвистикалық бірліктердің жиілігін анықтауға, сондай-ақ заңдылықтарды зерттеу үшін мәтіндердің корпустарын статистикалық талдау жасауға мүмкіндік береді [13].

Көңіл-күйді танудағы статистикалық әдістер сөздердің эмоционалды түске ие болу принципіне негізделген. Олар полярлық мәндері (оң, теріс, бейтарап) берілген сөздерді қамтитын көңіл-күй сөздіктерін (Sentiment Lexicons)

пайдаланады. Оның жұмыс принципі өте қарапайым. Кілт сөздіктерін қолдануға негізделген тәсіл мәтіндегі жеке сөздердің эмоционалды бояуын және олардың мәтіннің жалпы көңіл-күйіне қосқан үлесін талдауды қамтиды. Бұл тұрғыда көңіл-күй сөздіктері қолданылады, мұнда әр сөзге мәтіннің эмоционалды бояуы үшін маңыздылығын көрсететін мағына беріледі. Содан кейін мәтіндегі әр сөздің пайда болу жиілігін ескеретін мамандандырылған функцияны қолдана отырып, мәтіннің жалпы эмоционалды бояуын есептеу жүзеге асырылады.

Көңіл-күйді тануға көмектесу үшін әртүрлі көңіл-күй сөздіктері қолданылады. Олардың бірі- SentiWordNet, онда WordNet-тің әр сөзіне үш параметр бойынша баға беріледі: оң, теріс және бейтарап. NRC Emotion Lexicon сөздерді сегіз негізгі эмоциямен байланыстырады: қуаныш, қайғы, қорқыныш, ашу, сенім, таңдану, жиіркеніш және қызығушылық. Afinn сөздігі де бар, онда сөздерге -5-тен (өте теріс) +5-ке дейін (өте оң) баға беріледі.

Мәтіннің ырғағын талдау үшін қолданылатын статистикалық әдістер құнды құрал болғанымен, мазмұнды, сарказмды және табиғи тілдің басқа ерекшеліктерін ескеруде шектеулерге ие. Мысалы, мәтінде ирония немесе сарказм болса, олар көңіл-күйді дұрыс анықтамауы мүмкін.

Мәтіннің көңіл-күйін танудың заманауи тәсілдері модельдердің дәлдігі мен сенімділігін арттыру үшін машиналық оқыту әдістерін белсенді қолданады.

Машиналық оқыту – бұл компьютерлерге мәліметтер негізінде білім алуға және болжам жасауға мүмкіндік беретін жасанды интеллекттің негізгі құралдарының бірі. Қолмен жасалған дәстүрлі талдау әдістерінен айырмашылығы – машиналық оқыту бұл процесті күрделі алгоритмдер арқылы автоматтандырады. Ең танымал әдістердің бірі-SVM (Support Vector Machine, тірек векторлар әдісі).

Көңіл-күйді талдау шеңберінде SVM деректерді сыныптарға бөлу арқылы мәтіннің кілтін анықтау үшін қолданылады. Негізгі идея-деректер кластары арасындағы алшақтықты барынша арттыратын, жіктеуді ең айқын және қатеге төзімді ететін гипержазықтықты анықтау. Автор [8] еңбекте SVM және логистикалық регрессия модельдері көңіл-күйді тануда ең жақсы нәтиже көрсететінін атап өтті. Дегенмен, олардың күрделілігі көбінесе оқу деректерінің көлемі мен сапасына байланысты. Psychological Data from the Open Science Framework деректерінде жүргізілген зерттеулер SVM-дің Аңғал Байес жіктеуіштері, Random Forest және градиентті күшейту алгоритмі сияқты басқа дәстүрлі машиналық оқыту әдістерінен артықшылығын көрсетеді.

Дегенмен, біз қандай мәселені шешіп жатқанымызды және қандай деректер бар екенін ескерген жөн. Open Science Framework жағдайында, зерттеуде үлкен деректер болғандықтан, оған Random Forest немесе градиентті күшейту қолданған тиімдірек. Себебі бұл екі әдіс басқа машиналық оқыту әдістеріне қарағанда

деректер шуына және қайта оқытуға төзімдірек. Егер жылдамдық маңыздырақ болса, онда логистикалық регрессия немесе Аңғал Байес жіктеуіштері сияқты қарапайым модельдерді қолданған дұрыс.

Машиналық оқыту мен мәтінді генерациялаудың үлкен мүмкіндігіне қарамастан, олар сөздер мен сөз тіркестерінің мағынасын дұрыс түсінбеу сияқты қиындықтарға тап болуы мүмкін, бұл дұрыс емес немесе байланыссыз мәтіндердің пайда болуына әкелуі мүмкін. Сонымен қатар, мәтінді генерациялауды қолдану өзіндік ерекшелігі мен шығармашылығының жоқтығы үшін сынға ұшырауы мүмкін [13]. Бұл орайда мәтіндердің эмоционалды фонын анықтау қиынға соғады.

Мәтіндік көңіл-күйді тану үшін машиналық оқыту үлгілерін әзірлеу және қолдану алгоритмдерді таңдауда, деректерді дайындауда және нәтижелерді бағалауда мұқият қарауды қажет ететінін атап өту маңызды.

Терең оқуға негізделген көңіл-күйді талдау – табиғи тілді түсінуге және мәтіннен ақпаратты алудың қуатты құралы болып табылады. Оны көңіл-күйді тану, тақырыптарды анықтау және нақты құрылымы жоқ деректерден мағыналы ақпаратты алу үшін пайдалануға болады. Көңіл-күйді талдаудың ең танымал терең оқыту әдістері LSTM (ұзақ қысқа мерзімді жад) және BERT трансформаторларының екі жақты кодтауы болып табылады. Бұл сөйлемдегі сөздер арасындағы ұзақ мерзімді тәуелділікті үйренуге қабілетті қайталанатын нейрондық RNN желінің бір түрі. Ол белгіленген деректерге үлгіні үйрету арқылы мәтіндегі көңіл-күйді жіктеу үшін пайдаланылуы мүмкін. Ол көңіл-күйді талдау тапсырмалары үшін өте қолайлы, бірақ ұзақ мерзімді тәуелділіктерді түсіру мүмкіндігімен шектелген [14].

Аңғал Байес әдісі көңіл-күйді талдауда және спамды анықтауда жақсы. Өлшемдері көп мәтіндік деректер үшін SVM өте қолайлы. Күрделі үлгілерді тану нейрондық желілердің, әсіресе CNNs конволюционды нейрондық желілердің және RNN қайталанатын нейрондық желілердің артықшылығы болып табылады. Нейрондық желілер мен LSTM тілдерді модельдеу және машиналық аударма үшін тиімді жұмыс істейді, өйткені олар дәйекті енгізуді өңдей алады. Мәтіннің үлкен көлемдеріндегі контексті түсіну және күрделі лингвистикалық нюанстарды түсіру арқылы BERT және GPT сияқты трансформаторға негізделген модельдер табиғи тілді өңдеуде (NLP) революция жасады [15].

Тапсырма мәтіннің эмоционалдық реңкін анықтау болып табылатын мәтіндік сезімді талдауда гибридті әдістер бір ғана әдісті қолданудан гөрі дәлірек және сенімді тәсілді қамтамасыз етеді. Олар модельдерді жоғары дәлдікпен және беріктікпен қамтамасыз ету үшін лексикалық және машиналық оқыту әдістерінің артықшылықтарын біріктіреді.

Лексикалық тәсіл белгілі полярлығы бар сөздердің жиілігін талдау арқылы мәтіннің сезімін бағалау үшін сезім сөздігі немесе сезімдік тезаурус арқылы сөздер

мен сөз тіркестерін талдауға негізделген. Машиналық оқыту әдістері, керісінше, контекст, сарказм, ирония және басқа да тілдік өзгешіліктерді ескере отырып, белгілі сезімдері бар мәтіннің үлкен мәтін корпусындағы үлгілерді жаттықтырады.

Мәтінді өңдеудің қуатты құралдары бақыланатын оқыту үлгілерін (SVM, нейрондық желілер және Naïve Bayes сияқты), сондай-ақ бақыланбайтын әдістерді (тақырыпты модельдеу және кластерлеу сияқты) қамтиды. Трансформатор, атап айтқанда, күрделі тіл үлгілерін түсінуді жақсартатын тереңдетілген оқыту моделі болып табылады. Бұл үлгілердің сенімділігі дәлдік, нақтылық, еске түсіру, F1 ұпайы және AUC-ROC сияқты көрсеткіштерді пайдалану арқылы бағалау арқылы қамтамасыз етіледі.

Еңбек [13] авторлары көңіл-күйді дәлірек талдау үшін әртүрлі әдістерді біріктіретін әртүрлі гибридті әдістерді сипаттайды. Бір әдіс - кілт сөздерді немесе сөз тіркестерін шығару үшін статистикалық әдістерді қолдану. Мысалы, мәтіндегі түйінді сөздерді анықтау үшін TF-IDF (Term Frequency-Inverse Document Frequency) қолдануға болады, содан кейін олардың полярлығын бағалау үшін лексикалық талдауды қолдануға болады. Гибридті модельдер сонымен қатар мәтіннің лексикалық ерекшеліктеріне үйретуге болатын SVM (қолдау векторлық машинасы) немесе логистикалық регрессия сияқты классикалық машиналық оқыту әдістерін пайдаланады. Сонымен қатар, гибридті модель мәтіннің грамматикалық және синтаксистік ерекшеліктерін алу үшін NLP табиғи тілді өңдеу технологиясын пайдаланады, бұл эмоционалды талдаудың дәлдігін жақсартуға көмектеседі. Бұл әдістерді біріктіру тіл мен мазмұн аспектілерін ескеретін дәлірек және сенімді сезімді талдау үлгілеріне мүмкіндік береді.

Гибридті әдістің мысалы ретінде [16] ұсынған сарапшылардың шолуларын талдауға арналған GDM (Group Decision Making) әдісі болып табылады. GDM мәтіндегі оң және теріс эмоцияларға назар аудара отырып, шолудан құнды ақпаратты алу үшін көңіл-күй талдауын пайдаланады. Бұл мәлімдеменің мазмұнын ғана емес, эмоционалдық реңкін де бағалауға мүмкіндік береді. Мысалда мәтіннің грамматикалық және синтаксистік ерекшеліктерін шығару және кілт сөздердің эмоционалдық реңкін анықтау арқылы лексикалық талдау үшін NLP біріктіреді, сарапшылардың пікірлерінің көңіл-күйін дәлірек және сенімді бағалауды қамтамасыз етеді.

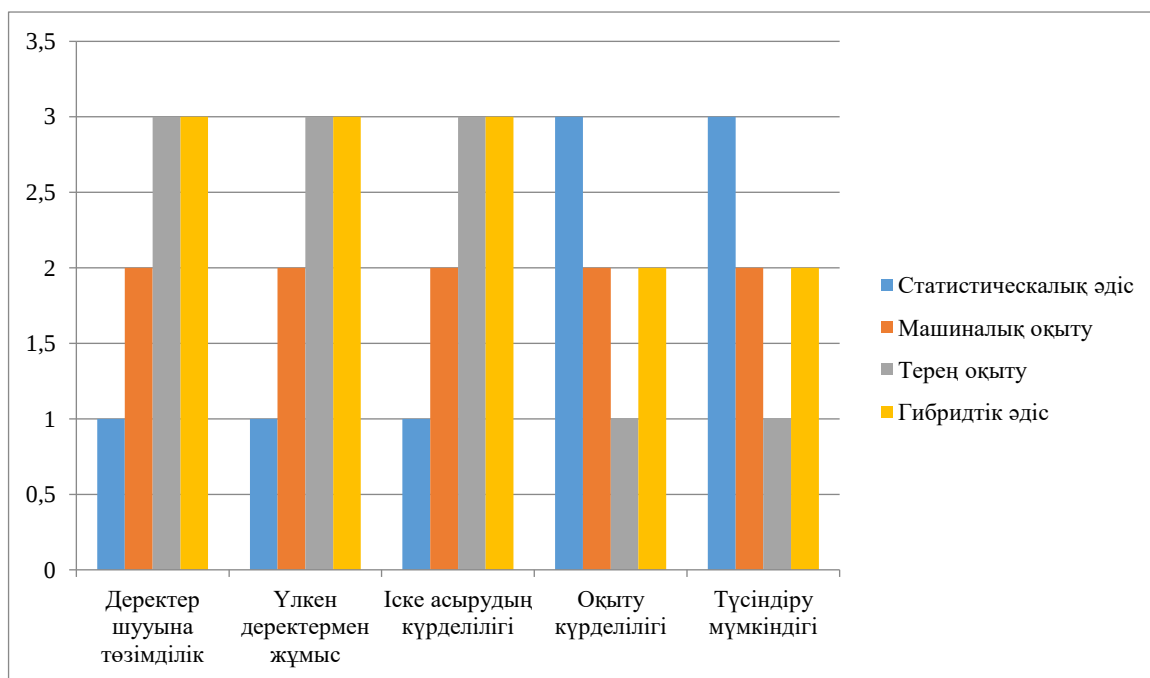
Гибридті әдістер тілдің әртүрлі ерекшеліктерін ескере отырып, мәтіндік сезімді дәлірек және сенімді бағалауға мүмкіндік береді.

Алдыңғы зерттеулер біршама ілгерілеушілікке қол жеткізгенімен, мәтін реңктерін талдаудың қолданыстағы үлгілері әлі де әртүрлі семантикалық ақпараттың жеткіліксіз интеграциялануы және кескіннің жеткіліксіз интерпретациялануы мәселесімен бетпе-бет келеді. Дегенмен, ғылыми прогресстің дамуы түрлі әдістерді біріктіріп, гибридті модельдерді жасауға мүмкінділік

туғызуда. Нәтижесінде, [17] мәтінді талдау үшін кванттық механиканың іргелі принциптерін қамтитын терең оқыту моделін жасалды, бұл технологияның жаңа салалардағы әлеуетін көрсетті. Бұл модель ұзақ қысқа мерзімді жады (LSTM) және өзін-өзі тану механизмі (SAM) бар желілерге негізделген мүмкіндіктерді шығару деңгейлеріне сәтті қол жеткізді. Жұмыс барысында, сонымен қатар, мәтіннің тығыздық матрицасын есептеу үшін кванттық комплекстік сандар принципін қолданды және функцияның конденсациясы мен өлшемділігін азайту үшін екі өлшемді конволюционды нейрондық желіні (CNN) қолданып, оның сезімді талдаудың басқа әдістерінен артықшылығын дәлелдеді.

ЗЕРТТЕУ НӘТИЖЕЛЕРІ

Зерттеуде мәтіннің эмоциясын анықтайтын төрт негізгі әдісін қарастырылды: статистикалық, машиналық оқыту, терең оқыту және гибриді. Олардың сипаттамаларын салыстырмалы бағалау үшін гистограмма салынды (Сурет-1).



Сурет-1. Мәтіннің көңіл-күйін талдау әдістерінің сипаттамаларының гистограммасы

Гистограмма бойынша бірнеше қорытынды жасауға болады. Терең оқыту және гибридік әдістер шуға жоғары төзімді және үлкен деректермен жұмыс істеуде тиімді. Статистикалық әдістерді енгізу және түсіндіру оңайырақ, бірақ шуға төзімділігі төмен және деректердің үлкен көлемімен жұмыс істеуде шектеулі. Терең оқыту ең баяу оқу жылдамдығына ие, бірақ жоғары дәлдікті қамтамасыз етеді.

Осы қорытындыларға сәйкес, мәтіндегі эмоционалды фонды анықтаудағы әдістердің әрқайсысының өзіндік артықшылықтары мен кемшіліктері бар (Кесте-1).

Кесте-1. Мәтіннің көңіл-күйін талдау әдістерін салыстыру

Әдіс	Артықшылықтар	Кемшіліктер
Статистическалық әдістер	Іске асыруға қарапайым, қол жетімді, түсіндіру оңай	Мәтіннің мазмұнын, ирония мен сарказмге назар аудармайды. Көп деректерді қажет етеді
Машиналық оқыту әдістері	Статистикалыққа қарағанда дәлірек, мәтін мазмұнын ескере алады	Оқыту үшін көбірек ресурстар қажет, қайта оқытуды қажет етуі мүмкін (overtraining)
Терең оқыту әдістері	Жоғары дәлдік, тілдің ерекшеліктерін ескеру мүмкіндігі	Оқыту үшін үлкен деректер жиынтығын және жоғары есептеу күрделілігін қажет етеді
Гибридтік әдістер	Жоғары дәлдік және тұрақты модельдер	Іске асыру қиын

Мәтін бойынша көңіл-күйді талдаудың әртүрлі тәсілдерін талдау бірде-бір әмбебап әдіс жоқ екенін көрсетеді. Статистикалық әдістерді қолдану оңай әрі жылдам, бірақ әрқашан тілдің мазмұндық және стильдік ерекшеліктерін ескермейді. Машиналық оқыту (мысалы, SVM) дәлірек модельдер жасауға мүмкіндік береді, бірақ сөздер мен сөз тіркестерінің мағынасын дұрыс түсінбеуі мүмкін. Терең оқыту (LSTM, BERT) күрделі үлгілерді өңдеуде өте тиімді, бірақ үлкен деректер жиынтығын қажет етеді және түсіндіру қиын. Әр түрлі тәсілдердің артықшылықтарын біріктіретін гибридті әдістер болашақ көзқарастарды көрсетеді, дегенмен оны жүзеге асыру қиын.

ТАЛҚЫЛАУ

Гистограммада және салыстыру кестелері түрінде ұсынылған нәтижелерде гибридті және терең оқыту әдістері дәлдік, шуға төзімділік және үлкен көлемдегі деректерді өңдеу тиімділігі бойынша бірдей тиімді екенін көрсетеді.

Бұл нәтиже жұмысымен расталады, онда ғалымдар LSTM және SAM көмегімен көңіл-күйді талдаудың сәтті үлгісін жасау үшін кванттық механика мен терең оқытуды біріктіретін гибридті тәсілді сәтті қолданды.

Гибридті әдістерді жүзеге асыру күрделірек екенін атап өту маңызды, бұл кейбір жағдайларда іске асыруға кедергі болуы мүмкін.

Мәтіндегі көңіл-күйді талдаудан басқа, бұл әдістерді табиғи тілді өңдеуде, әлеуметтік медианы талдауда, қызмет сапасын бағалауда және қаржы саласындағы тәуекелдерді анықтауда қолдануға болады.

Болашақта анағұрлым қуатты гибридті модельдерді әзірлеу, оқу деректерінің көлемін ұлғайту және сарказм мен мәдени айырмашылықтарды жою үшін көбірек зерттеулер қажет.

Тұтастай алғанда, гибридті және терең оқыту әдістері адам тілінің ерекшеліктерін қабылдауға қабілетті неғұрлым дәл және жалпы үлгілерді ұсынуға уәде беретін сезімді талдау жүйелерін дамытудың перспективалық бағыты болып табылады.

Машиналық оқытудың дамуымен мәтінді талдаудың неғұрлым жетілдірілген және дәл әдістері мүмкін болады, бұл әртүрлі салаларда маңызды артықшылықтар мен тереңірек түсінуді қамтамасыз етеді. Дегенмен, прогреске қарамастан, мәтіндік сезімді талдауда әлі де шешуді қажет ететін көптеген мәселелер бар.

ҚОРЫТЫНДЫ

Мәтіннің эмоционалдық түсін талдаудың қолданыстағы әдістерін (статистикалық, машиналық оқыту, терең оқыту және гибридті) қарастыру әрқайсысының өзіндік артықшылықтары мен кемшіліктері бар екенін көрсетеді. Статистикалық әдістер қарапайым болғанымен, дәлдігі шектеулі, ал машиналық оқыту әдістері жоғары дәлдікті көрсетсе де, үлкен оқу ресурстарын қажет етеді және шамадан тыс орнатудан зардап шегуі мүмкін. Өз кезегінде тереңдетіп оқыту ең жоғары дәлдікке ие, бірақ жоғары есептеу күрделілігімен және баяу оқу жылдамдығымен сипатталады. Гибридті әдістер әртүрлі әдістердің ең жақсы қасиеттерін біріктіреді және жоғары дәлдік пен беріктікті көрсетеді, бірақ олардың орындалуы күрделі.

Зерттеу гибридтік әдістер мен терең оқыту әдістері мәтіннің эмоционалдық бояуын дәлірек және жан-жақты түсінуді қамтамасыз ете алатын көңіл-күй талдау жүйесін дамытудың ең перспективалы бағыттары болып табылатынын көрсетті.

Көңіл-күйді талдау - дамудың үлкен әлеуеті бар қуатты құрал, бірақ оның мүмкіндіктерін толық іске асыру үшін белгілі бір мәселелерді шешу қажет. Көңіл-күйді талдаудың қолданыстағы әдістерінің ғылыми еңбектерін талдау барлық жағдайларда ең жақсы рөл атқара алатын идеалды әдіс жоқ екенін көрсетеді.

Терең оқыту және гибридті әдістер деректер шууының жоғары дәлдігі мен беріктігін көрсетеді, бірақ олардың тиімділігі деректердің нақты сипаттамаларына және тапсырмаға байланысты.

Көңіл-күйді талдауды одан әрі дамыту үшін бірнеше негізгі бағыттарға назар аудару қажет. Біріншіден, әртүрлі әдістердің ең жақсы мүмкіндіктерін біріктіретін жақсырақ гибридті модельдерді әзірлеу маңызды. Екіншіден, мәдени-эмоционалды үлгілерді құруға ерекше назар аудару керек, бұл әсіресе халықаралық маркетингтік науқандарда көңіл-күйді талдау үшін маңызды.

Сондай-ақ модельдің дәлдігі мен тұрақтылығын жақсарту үшін оқу деректерінің көлемін арттыру қажет. Әртүрлі салаларда қолдануға болатын икемді және жалпы алгоритмдерді әзірлеу маңызды: әлеуметтік желілерді талдаудан

тұтынушыларға қызмет көрсету сапасын бағалауға дейін. Бұл тұтынушылармен тиімді әрекеттесуге мүмкіндік береді, қызмет көрсету сапасын жақсартады және корпоративтік бәсекеге қабілеттілігін арттырады. Әрине, талдау нәтижелеріне ирония мен мәдени айырмашылықтардың әсерін жою әдістерін әзірлеу қажет.

Нәтижелер көңіл-күйді талдау саласындағы әрі қарай зерттеудің маңыздылығын көрсетеді. Жетілдірілген аналитикалық әдістердің дамуы мәтіннің эмоционалдық бояуын тереңірек түсінуге мүмкіндік беріп қана қоймайды, сонымен қатар маркетингтен денсаулық сақтау саласына дейінгі салалардағы қолданбалар үшін жаңа мүмкіндіктер ашады, қазіргі әлемдегі ақпаратпен тиімді әрекеттесуге ықпал етеді.

ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ

1. M. Wankhade, A. C. S. Rao and C. Kulkarni, "A survey on sentiment analysis methods applications and challenges". *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5731-5780, oct. 2022.
2. L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, et al., "Review of deep learning: Concepts CNN architectures challenges applications future directions", *J. Big Data*, vol. 8, no. 1, pp. 1-74, Mar. 2021.
3. Hartmann, J., Heitmann, M., Siebert, C., & Schamp, C (2023). More than a Feeling: Accuracy and Application of Sentiment Analysis. *International Journal of Research in Marketing*, 40(1), 75-87. Article 40(1). <https://doi.org/10.1016/j.ijresmar.2022.05.005>
4. Цветков А. В., Рыбанов А. А. Разработка программной системы анализ эмоциональной окраски рецензий к фильмам. Развитие науки и образования: новые подходы и актуальные исследования. – 2021. – С. 128-133
5. Y. Tim, L. Cui and Z. Sheng, "Digital resilience: How rural communities leapfrogged into sustainable development", *Inf. Syst. J.*, vol. 31, no. 2, pp. 323-345, Mar. 2021.
6. Полюшина Д.В., Андронов А.Н. Анализ тональности комментариев, оставленных в социальной сети. Огарёвские чтения. - 2022. - С. 745-752
7. Марьенков А.Н., Кривенко В.И. Сбор и обработка текстовых данных в контексте оценки социальных настроений: Методологические аспекты. *Инженерный вестник дона*. 2022. №7 ivdon.ru/ru/magazine/archive/n7y2022/7797
8. Пронин А.Г. Исследование методов машинного обучения для анализа настроений в записях пациентов. *Международный научный журнал «Вестник Науки»*. - 2024 г. - Т. 1. - № 3 (72) - С. 448-453.
9. Черничкин Д.А., Кривенко А.И. Анализ общественных настроений в социальных медиа астраханской области. *Журнал исследований средств*

массовой информации. 2023. № 3. С. 145-169. <https://doi.org/10.46539/gmd.v5i3.372>

10. Пронина Е.В., Пронин Д.Д. Метод анализа эмоциональной тональности текста как цифровой инструмент изучения текста произведения литературы (на примере романа М.Ю. Лермонтова «Герой нашего времени»). Современный ученый. – 2022 - №6. – С. 8-14
11. Барахнин В. Б., Менглиев Д. Б. Алгоритмический подход к анализу эмоциональной окраски и поэтичности узбекских текстов. (2024). «Современные технологии компьютерной лингвистики», 2(22.04), 354-358.
12. Косых Н.Е., Хомоненко А.Д., Куранова О.Н. Разработка web-приложения для анализа настроений текста с помощью фреймворка Flask и языка Python. Наукоемкие технологии в космических исследованиях земли. 2022. Т. 14. № 1. С. 45-52. doi: 10.36724/2409&5419&2022&14&1&45&52
13. Джумабаева М.Ш., Бурнашев Р.Ф. Информационные технологии в обработке лингвистической информации. Science and education. - 2023. - Т. 4. - №. 4. - С. 643-653.
14. Тапе Х.Ж.М., Погуда А.А. Сравнение методов анализа настроений глубокого обучения, включая LSTM и машинное обучение. Проблемы информатизации экономики и управления. – 2023. – Т. 27. - № 4. – С. 60-71
15. Кузьмина Е.С., Горюнов Д.А. Методы анализа текстовых данных с использованием машинного обучения. Экономика и социум. - 2024. - №6(121)-1 – С. -1174-1178
16. JR Trillo, E Herrera-Viedma, JA Morente-Molinera, FJ Cabrerizo. A Large Scale Group Decision Making System Based on Sentiment Analysis Cluster. Information Fusion.– 2022. - №91, 633-643
17. Li B., Zhang D., Zhao Zh., Gao J., Yuan Y. Quantum-inspired Interpretable Deep Learning Architecture for Text Sentiment Analysis. arxiv preprint arxiv:2408.07891. – 2024.