

Тоқтарова Айгерім Бастарбекқызының 8D06115 – «Ақпараттық жүйелер» білім беру бағдарламасы бойынша доктор (PhD) дәрежесін алуға ұсынылған «Табиғи тілдерді өңдеу және машиналық оқыту әдістерін қолдана отырып онлайн контенттегі ғадауат сөздерді анықтау» диссертациялық жұмысына

АНДАТПА

Зерттеу жұмысының өзектілігі. Бүгінгі күні онлайн әлеуметтік желілердегі ғадауат тілді сөздерді бар пікірлерді, яғни дініне, этникалық тегіне, жынысына немесе басқа факторларға негізделе отырып жеке адамға немесе топқа қарсы қорлау немесе кемсіту ниетінде ауызша немесе жазбаша сөз тіркестері немесе сөйлемдерді деректер базасы ретінде жинақтау үшін машиналық оқыту алгоритмдерін пайдалану арқылы оңай және автоматтандырылған қадамдарды жүзеге асыруға болады.

Бүгінде, интернет пайдаланушыларының саны артқан сайын, онлайн контенттегі жағымсыз пікір қалдырушылар саны қаптап кетуі мәселе туғызып отыр.

Ғадауат тілді пікірлер немесе адамның психологиялық жағдайына қауіп төндіретін сөздер және сөз тіркестер – интернет немесе әлеуметтік желі құрбандарына құқық бұзушылардың қауіп төндіруі күн сайын артып келеді.

Ғадауат тілді сөздер немесе қоқан-лоққы пиғылды пікірлерді, қауесеттерді тарату, кемсіту немесе қорлау мақсатында жасалған суреттерді, бейнелерді тарату немесе интернеттегі әлеуметтік желілерден адамдарды бұғаттау сияқты бірнеше жолмен көрінеді.

Ғадауат тілді пікірлерді жариялау, бөлісу салдары өте зиянды болуы мүмкін. Жәбірленушілерде алаңдаушылық, үмітсіздік, өзін-өзі бағалаудың төмендеуі және өзіне-өзі қол жұмсау туралы ойлардың белгілері көрінуі мүмкін. Қазақстандық жастардың 75 пайызы интернеттегі қауіптің алуан түрлерін басынан кешірген. Мұны Facebook компаниясы әлеуметтік желі қолданушылары арасында жүргізген онлайн сауалнаманың нәтижесінен байқауға болады. Сонымен қатар, оларда ұйқысыздық, тамаққа тәбеттің азаюы, өз –өзіне көңіл бөлудің жеткіліксіздігі және сабақ оқуда және қарым – қатынасс қалыптастыру өнімділіктері төмендеуі мүмкін. Кейбір жағдайларда ғадауат тілді пікірлердің көптігі физикалық зардаптарға әкелуі мүмкін, себебі жәбірленуші өз ортасынан оқшаулануды сезінуі немесе қауіптің жоғарлауын сезінуі мүмкін. 2024 жылы 165 жасөспірім өз-өзіне қол жұмсап, онлайн контенттен балағаттау немесе кемсіту мағынасына ие пікірлердің салдарынан 378 кәмелетке толмаған жасөспірім өз-өзіне қол жұмсауға әрекеттенген.

Жалпы алғанда, жасөспірімдердің 5% интернеттен ғадауат тілді пікірлерді оқу, көру арқылы келетін қауіпті бастан өткерген немесе айына 2-3 рет немесе одан да көп онлайн қауіп тууы мүмкін. Жасөспірімдердің 12% кем дегенде бір рет интернеттен келетін қауіпке ұшыраған. Егер олардың іс - әрекеті заңсыз деп

жіктелсе, әлеуметтік және кәсіби салдарлары, соның ішінде беделіне нұқсан келтіру, достық қарым-қатынас орнатудағы қиындықтар немесе жұмысқа орналасуды қамтамасыз етуде өзіндік жағымсыз салдары бар екенін көруге болады. Статистика көрсеткендей, интернеттен келетін қауіпке тап болғандардың үштен бірінен астамы жеке хабарламалар арқылы қорлау, кемсіту пиғылдары кездесетін хабарлама алған, 24 пайызы өздері туралы кемсіту мақсаты айқын жазбаларды көрген және үштен бір бөлігі (31%) фотосуреттерінің астында жағымсыз пікірлерге тап болған.

Адамдарға психологиялық тұрақтылықтың салдары туралы білім беру онлайн қауіпті азайтудың бір стратегиясы болып табылады. Білімділікті арттыруға арналған вебинарлар, нұсқаулықтар және тәрбиешілер мен ата-аналарға арналған материалдар өз кезегінде төнетін онлайн қауіптің жетегінде кетпеуге көмектесе алады.

Біз барлығы бір-біріне құрметпен және жанашырлықпен қарайтын ортаны қалыптастыру үшін қолдан келгеннің бәрін жасауымыз керек. Біз өзара сыйластық пен жайлы атмосфераны қалыптастыру арқылы әлемді барлық адамдар үшін жақсы ортаға айналдыра аламыз.

Мысалы, 2024 жылы жүргізілген әлем мемлекеттерінің интернетті қолдану деңгейі статистикасы бойынша Қазақстан мемлекеті 20 075277 миллион жалпы халықтың 18,9 миллион адамы, яғни 92,3% интернетті тұрақты түрде қолданатынын көрсетіп отыр. Соның ішінде әлеуметтік желі қолданушылар саны жалпы халықтың 14 миллионын, яғни, 71,5% құрап отыр.

Онлайн желідегі ақпараттар, жаңалықтар аптасына 7 күн, 24 сағат бойы қолжетімділікке ие, интернеттен қасақана немесе абайсызда күннің кез-келген уақытында келетін «шабуылдан» қорғана алмайсың, себебі электрондық пошта, ұялы телефондардағы мессенджерлер немесе жеке әлеуметтік жерлілерге, ақпараттық – танымдық сайттарға кез- келген уақытта келіп түсу мүмкіндігі әрдайым жоғары. Бұл желіні, интернетті онлайн пайдаланушыларға, оның ішінде жастар мен жеткіншектерге психологиялық түрде кері әсерін бірден тигізуі мүмкін. Сонымен қатар, интернетте, әлеуметтік желілерде өз аты – жөнін жасырып, лақап атпен желі қолданушысы бола білу мүмкіндігі бар, яғни интернеттегі «шабуылдың» кімнен келгенін білу көрсеткіші төмендеуі мүмкін.

Эмоционалдық «шабуылдың» зорлық – зомбылықтың физикалық түріне қарағанда адамды психологиялық ауытқуларға әкеліп соғу қаупі жоғары екенін психологтар да мойындап отыр.

Адамдардың психологиялық жағдайына кері әсерін беретін зиянды мәліметтерді анықтау қазіргі таңда ақпараттық техникалық құралдардың негізі міндеті деп алсақ, бұл міндетті іске асырып, өңдеуде машиналық оқыту әдістерін пайдалану арқылы автоматтандыру әлдеқайда тиімді деп саналып отыр.

Ғылыми зерттеу жұмысының мақсаты - мәтіндік деректерде ғадауат тілді сөздерді автоматты түрде анықтау үшін терең нейрондық желі моделін құру.

Ғылыми зерттеу жұмысының міндеттері: Зерттеу мақсатына жету үшін келесі міндеттер шешілді:

1. Ғадауат тілді пікірлерді анықтауға арналған екілік және көп класты жіктеуге арналған машиналық оқыту алгоритмдерін талдау.

2. Ғылыми зерттеу жұмысы үшін ML және DL алгоритмдерін оқыту үшін қазақ тілінде деректерді жинау және алдын ала өңдеу.

3. Терең оқыту архитектурасын талдау. DL алгоритмдерінің әртүрлі түрлерін оқыту, мысалы:

а) Конволюциялық нейрондық желілер

ә) Терең нейрондық желілер

б) Қайталанатын нейрондық желілер

г) Қысқа мерзімді ұзақ мерзімді жады бар желілер

ғ) Көпқабатты перцептрондар

д) жасанды интеллект механизмін пайдаланатын терең нейрондық желілер

ж) Модель нәтижелерін жақсарту үшін эксперименталды зерттеулер жүргізу, салыстыру, үлгі таңдау, гиперпараметрлерді баптау.

Зерттеу нысаны: әлеуметтік желілер (Вконтакте, Инстаграм, Ютуб, твиттер), жаңалықтар порталдары (nur.kz, tengri news).

Зерттеу тақырыбы: Мәтіндік деректерде ғадауат тілді пікірлерді анықтау үшін машиналық оқыту және терең оқыту алгоритмдері.

Зерттеудің теориялық маңыздылығы: мәтіндік деректердегі ғадауат тілді пікірлерді анықтау бойынша бар жұмыстарды зерттеу, табиғи тілдің өңдеу құралдарын талдау.

Зерттеудің практикалық маңыздылығы: терең нейрондық желілерді әзірлеу және оқыту, табиғи тілді өңдеу, әлеуметтік медиа кеңістігінде ғадауат тілді пікірлерді анықтау бойынша эксперименттер жүргізу.

Қорғауға ұсынылатын негізгі ережелер:

1. Қазақ тілі үшін деректер жинағы жасақталды, алдын ала өңделді және одан әрі машиналық және терең оқыту тапсырмалары үшін қолмен сыныптарға бөлінді.

2. Ғадауат тілді сөздерді анықтау тапсырмасында екілік класты жіктеу тапсырмалары үшін назар аудару механизмі бар терең нейрондық желі әзірленді және оқытылды.

3. Ғадауат тілді анықтау тапсырмасында машиналық оқыту мен терең оқыту алгоритмдері арасында салыстырмалы талдау жүргізілді.

Зерттеу жұмысының ғылыми жаңалығы:

1. Қазақ тілінде ғадауат тілді сөздерден құралған деректер жинағын әзірлеу:

Қазақ тілінде ғадауат тілді сөздерден тұратын толық деректер жинағы жиналды, аннотацияланды және алдын ала өңделді, машиналық оқыту мен терең оқыту алгоритмдерін оқыту. Деректерді ұсынудың дәйектілігі мен дәлдігін қамтамасыз ету үшін қолмен санаттау және лингвистикалық нормалау жүзеге асырылды.

2. Гибридті терең нейрондық желі моделін құру:

BERT-ті зейін механизмімен біріктіретін бірегей гибриді терең оқыту архитектурасы әзірленді және енгізілді. Модель желідегі мәтіндегі ғадауат тілді сөздерді екілік классқа бөлу үшін әзірленген және ғадауат тілді сөзді контекстік және семантикалық ерекшеліктерін түсінуде жоғары нәтижеге жете білді.

3. Машиналық оқыту мен терең оқыту алгоритмдерін салыстырмалы бағалау:

Ұсынылған гибриді модельмен салыстырғанда CNN, RNN, LSTM, MLP және BiLSTM сияқты бірнеше алгоритмдерге жан-жақты зерттеу және салыстырмалы талдау жүргізілді. Зерттеу көптілді параметрлерде күрделі тіл үлгілерін өңдеудегі трансформаторға негізделген модельдердің жақсартылған контекстік түсінігі мен икемділігін көрсетті.

4. Шулы онлайн деректер үшін алдын ала өңдеудің тиімді әдістері: Зерттеу бейресми тілмен, кодты ауыстырумен, эмоджилермен және әлеуметтік медиа байланысында жиі кездесетін арнайы таңбалармен жұмыс істеуге жарамды сенімді алдын ала өңдеу құрылымдарын әзірледі. Бұл практикалық қолдану контекстінде модельдің сенімділігін айтарлықтай арттырды.

5. Ұсынылған модельдің тиімділігі Қазақстандық Contur.kz және Serke.org ақпараттық сайттарының арқылы бағаланды. Бұл платформалар көбінесе саяси қоғамдық пікірді және эмоцияға толы пікірталастарды көрсететін әлеуметтік-саяси түсініктемелер мен редакциялық мазмұнға ерекше назар аударуы үшін таңдалды. Бағалау үлгінің ресми, эмоционалды жаңалықтардағы ғадауат тілді сөздерді тесттілік режимде дәл анықтауда тиімді екенін көрсетті.

Докторанттың жеке қосқан үлесі.

Диссертациялық ғылыми зерттеу тақырыбы бойынша әдеби материалдарды және патенттік талдау мен жалпылау, талдау мен зерттеу әдістерін таңдау, теориялық және эксперименттік ғылыми зерттеулер, зерттеулер мен тәжірибелік жұмыстарын жүргізді.

Зерттеу нәтижелерін апробациялау. «Компьютерлік инженерия» кафедрасының ғылыми семинарлары мен мәжілістерінде, сондай-ақ Хожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университетінде және халықаралық конференцияларда баяндалды. Атап айтқанда:

1. «Компьютерлік инженерия» кафедрасының кеңейтілген мәжілісі, №6, 2025 жылғы 19 ақпан;
2. 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST), Астана қ.;
3. The 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021 ж., Үндістан.

Диссертациялық жұмыстың негізгі нәтижелері келесі ғылыми басылымдарда және халықаралық конференцияларда баяндалды:

Web of Science және Scopus базасында индекстелген ғылыми журналдарда жарияланған мақалалар:

1. Toktarova, A., Abushakhma, A., Adylbekova, E., Manapova, A., Kaldarova, B., Atayev, Y., ... & Aidarkhanova, A. (2023). Offensive language identification in low resource languages using bidirectional long-short-term memory network. International Journal of Advanced Computer Science and Applications, 14(6). DOI: <http://dx.doi.org/10.14569/IJACSA.2023.0140687>

2. Toktarova, A., Syrlybay, D., Myrzakhmetova, B., Anuarbekova, G., Rakhimbayeva, G., Zhylanbaeva, B., ... & Kerimbekov, M. (2023). Hate speech detection in social networks using machine learning and deep learning methods. International Journal of Advanced Computer Science and Applications, 14(5).

DOI: <http://dx.doi.org/10.14569/IJACSA.2023.0140542>

ҚР ҒЖБ БҒСҚК ұсынған ғылыми журналдарда:

1. B. Toktarova, B.S.Omarov, G.N. Kazbekova, S.A.Mamikov, F.E. Temirbekova Collecting hate speech database on social network in Kazakh language by using machine learning/ Bulletin of the National Academy of Sciences of the Republic of Kazakhstan. Series of Physics and Mathematics, 1991-346X, No. 1, pp. 191-203, 2023, DOI:<https://doi.org/10.32014/2023.2518-1726.177>

2. A.B. Toktarova, Zh.Zh. Azhibekova, D.R. Sultan, M.A. Kerimbekov “Collecting hate speech in Kazakh language in online content using machine learning”/ Bulletin of Abai KazNPU, series “Physics and Mathematics Sciences”, vol. 81, No. 1, 2023

3. A.B. Toktarova, B.S. Omarov, Zh.Zh. Azhibekova, S. A. Mamikov “The importance of artificial intelligence in identifying offensive words in online content” Bulletin of Toraigyrov University. ISSN 2710-3420 Energy series. No. 1, p. 311-322, 2023

4. A.B. Toktarova, B.S. Omarov, Zh.Zh. Azhibekova, Automated classification of offensive words using the “emotional” opinions of network users / Bulletin of the National Academy of Sciences of the Republic of Kazakhstan, No. 2 (88), DOI:<https://doi.org/10.47533/2023.1606-146X.9>

5. A.B. Toktarova, B.S. Omarov, B.A. Kaldarova, Using bilstm in identifying offensive words from low-resource languages Bulletin of the National Academy of Sciences of the Republic of Kazakhstan. Physics and Mathematics Series, 1991-346X, No. 3, pp. 174-189, 2024, DOI: 10.32014/2024.2518-1726.299

6. A.B. Toktarova, B.S.Omarov, G.I. Beissenova, R.B. Abdrakhmanov Analysis of hate speech words in online content by using data mining Bulletin of the National Academy of Sciences of the Republic of Kazakhstan. Series: Physics and Mathematics, No. 2 (346), pp. 237-251, 2023, <https://doi.org/10.32014/2023.2518-1726.196>

7. A.B. Toktarova, B.S. Omarov, F.S. Temirbekova "Classification of Kazakh language obscene words and adaptation of machine learning methods to their detection" / Bulletin of KazNPU named after Abai, series "Physical and Mathematical Sciences", Vol. 82, No. 2 (2023)

8. A.Toktarova, Zh.Azhibekova, A.Aliyeva, N.Sarsenbieva "Bidirectional long short-term memory in hate speech detection problem on networks" / Bulletin of KazNPU

named after Abai, series "Physical and Mathematical Sciences" Vol. 87, No. 3 (2024), DOI: 10.51889/29595894.2024.87.3.010

Халықаралық конференциялар еңбектерінде:

1. Toktarova, A., Sultan, D., & Azhibekova, Z. (2024, May). Review of Machine Learning Models in Cyberbullying Detection Problem. In 2024 IEEE 4th International Conference on Smart Information Systems and Technologies (SIST) (pp. 233-238). IEEE.

2. Sultan, D., Suliman, A., Toktarova, A., Omarov, B., Mamikov, S., & Beissenova, G. (2021, January). Cyberbullying detection and prevention: Data mining in social media. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 338-342). IEEE.

Авторлық құқықпен қорғалатын құқықтардың мемлекеттік тізілімге мәліметтерді енгізу туралы куәлік. «Табиғи тілдерді өңдеу және машиналық оқыту әдістерін қолдана отырып онлайн контенттегі ғадауат сөздерді анықтау» тақырыбындағы мәліметтер қоры авторлық құқық объектісі ретінде мемлекеттік тіркеуден өтті. ҚР Әділет министрлігінің Республикалық зияткерлік меншік институтының тізімінде №54674 тіркеу нөмірімен 2025 жылғы 14 ақпанда тіркелді.

Ақпараттық сайт процессіне ендіру актісі. Ұсынылған модель 2025 жылғы 31 қаңтарда contur.kz және serke.org ақпараттық сайттарында тестілік режимде сынақтан өткізілді.

Диссертацияның көлемі мен құрылымы.

Диссертациялық жұмыс кіріспеден, төрт тараудан және қорытындыдан тұрады. Жұмыс машиналық 121 мәтіндік бетінен, компьютерлік мүмкіндіктерді пайдалана отырып, иллюстрациялар, сызбалар мен кестелер арқылы дайындалған. Пайдаланылған әдебиеттер тізімі 139 атаудан тұрады.

Зерттеу жұмысы барысында көрсеткен баға жетпес көмегі мен кеңестері үшін автор ғылыми жетекшісі – философия ғылымдарының докторы (PhD), Халықаралық ақпараттық технологиялар университетінің қауымдастырылған профессоры Омаров Батырхан Сұлтанұлына және шетелдік кеңесші – Ыстамбұл техникалық университетінің профессоры Ешреф Адалыға (Түркия, Ыстамбұл қаласы) зор алғысын білдіреді.