

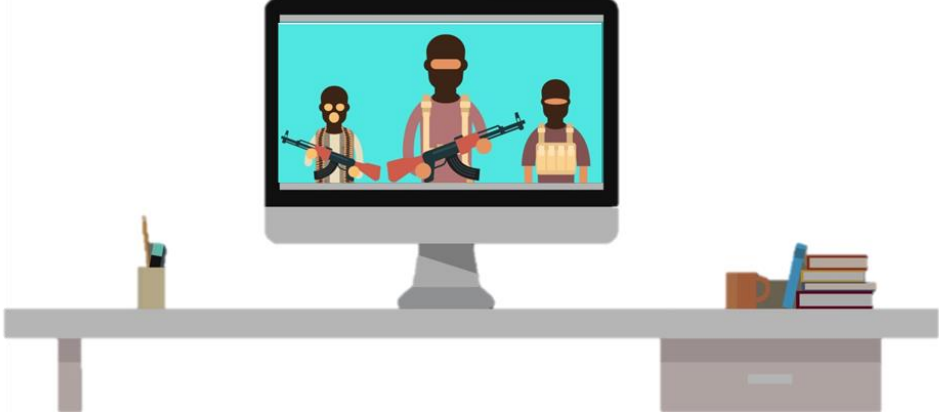
## Жоба туралы қысқаша ақпарат

Жоба аты	AP19676342 «Қазақ тіліндегі кибер экстремизмнің идеологиялық бағыттарын жасанды интеллект әдістері көмегімен мультиклассификациялау»
Жоба өзектілігі	<p>Экстремистер Интернетті өз идеяларын тарату, жаңа қолдаушыларды тарту, үйлестіру, тіпті қылмыстарды жоспарлау және оларды орындау үшін қолдана алады. Интернет оларға байланыс пен үгіт-насихат үшін кең мүмкіндіктер береді, бұл мазмұнды бақылау және онлайн ортадағы экстремистік топтарға қарсы тұру үшін маңызды рөл атқарады.</p> <p>Экстремистік тіл идеологиясының анықтамасы қазіргі әлемде, әсіресе онлайн-коммуникация контекстінде үлкен маңызға ие. Бұл анықтаманың өзектілігі бірқатар факторларға байланысты. Үкіметтер мен қоғамдық ұйымдар экстремизм мен терроризмге қарсы тұруға тырысады және бұл үшін олардың мазмұнындағы идеологияларды нақты анықтау маңызды болып табылады. Интернетте экстремистік идеялардың таралуы радикалдануға және жағымсыз салдарға әкелуі мүмкін. Мұндай мәтіндерді анықтау және онымен күресу қауіпсіз онлайн кеңістікті құруға көмектеседі. Экстремизмге қарсы күресте елдер мен халықаралық ұйымдар арасындағы ынтымақтастық бірыңғай стандарттар мен тәсілдерді қамтамасыз ету үшін идеологияларды нақты анықтауды талап етеді. Осылайша, экстремистік тіл идеологияларын анықтаудың өзектілігі қазіргі ақпараттық қоғамда маңызды орын алады.</p>
Жоба мақсаты	Жобаның мақсаты – әлеуметтік желілерде және мессенджерлерде жасанды интеллект әдістері көмегімен деструктивті сипаттағы кибер насихаттың таралуын зерттеу, мәтіндік, аудио және бейне жарияланымдарында экстремистік мазмұнның идеологиялық бағыттарындағы мультиклассификациялау модельдерін құру және діни мазмұнды ең белсенді кибернасихаттауды, қауымдастықтарды анықтаудың үлгілері мен әдістерін құру.
Жоба міндеттері	<ol style="list-style-type: none"><li>1. Мәліметтер жинау модулін әзірлеу.</li><li>2. Экстремистік мазмұндағы және идеологиялық бағытты мультиклассификациялау корпусын құру.</li><li>3. Әлеуметтік желілер мен мессенджерлердің мәтіндік жарияланымдарындағы экстремистік мазмұнның идеологиялық бағытының (деструктивті діни ағымдарды насихаттау, радикалдандыру және экстремистік және террористік ұйымдарға тарту) мультиклассификациясының үлгісін құру.</li><li>4. Әлеуметтік желілер мен мессенджерлердің аудио және бейне жарияланымдарындағы экстремистік мазмұнның идеологиялық бағытының мультиклассификациялық үлгілерін құру</li><li>5. Әлеуметтік желілер мен мессенджерлердің мәтіндік, аудио, видео жарияланымдарындағы экстремистік(діни экстремизм,</li></ol>

	<p>саяси экстремизм, ксенофобия) мазмұнның мультиклассификациясының үлгісін құру.</p> <p>6. Әлеуметтік желілер мен мессенджерлердегі ең белсенді кибернасихаттарды анықтаудың гибридті моделін құру</p> <p>7. Берілген параметрлер жиынтығы негізінде әлеуметтік желілердегі қауымдастықтарды анықтау моделін, алгоритмдерін және әдістерін құру.</p> <p>8. Дін мәселелері бойынша кеңес беру үшін қазақ тілінде диалог жүргізетін чат-бот құру.</p> <p>9. Жасалған әдістер мен модельдерді жүзеге асыратын бағдарламалық жабдықтама әзірлеу.</p>
<p>Күтілетін және қол жеткізілген нәтижелер</p>	<p>Қол жеткізілген нәтижелер:  Мәліметтерді жинау модулі әзірленді. Мәліметтерді жинау модулі Телеграм, ВКонтакте, Твиттер, Youtube әлеуметтік желілерде мәліметті іздеу үшін API технологияларын пайдалу көмегімен әзірленді. Таңдалған әлеуметтік медиа желілерінде деструктивті нанымдардың белгілері бар 400 ден астам топтар сарапталып кілттік сөздер жиыны әзірленді. Мәліметтерді жинау модулінде кілттік сөздер тізіміне сәйкес және нақты таңдалған уақыт аралығы, топ идентификаторы сұранысы негізінде мәлімет қорын жинақтау және толықтыру функционалы жүзеге асырылды. Экстремистік мазмұндағы және идеологиялық бағытты мультиклассификациялау корпусы құрылды. Әлеуметтік желілер мен жаңалықтар сайттарынан мәтіндік деректерді жинайтын идеологиялық экстремистік бағытты анықтау үшін экстремистік және бейтарап мәтіндердің ережелері мен категориялары анықталды. Деструктивті діни ағымдарды насихаттау, радикалдандыру және экстремистік және террористік ұйымдарға тарту, бейтарап кластарына сараптама жүргізіліп, әр класс үшін кілттік сөздер тізімі, ережелер, деректерді жинау үшін топтар тізімі құрастырылды. Корпус ережелерге сай қолмен сұрыпталып белгіленді. Аннотация нәтижесінде кластарға келесідей белгілер тағайындалды : Propaganda (0), Radicalization (1), Recruitment (2), Neutral (3). Құрастырылған корпус машиналық оқыту әдістерін қолдану мақсатында 80% және 20% қатынасында оқыту және тестілеу жинақтарына бөлінді. Әр класста 2000 астам мәтіндер жинақталған. Корпус мәліметтеріне токенизация, пунктуациядан тазалау, көп кездесетін сөздерден тазалау, стоп сөздерден тазалау сияқты препроцессинг алгоритмдері қолданылды. Корпус статистикасы және визуализациясы құрастырылды. Word2vec, bag of words және n-gram алгоритмдері мен модельдері қолданылып, корпус машиналық оқытуға дайындалды.</p> <p>Әлеуметтік желілер мен мессенджерлердің мәтіндік жарияланымдарындағы экстремистік мазмұнның идеологиялық бағытының (деструктивті діни ағымдарды насихаттау, радикалдандыру және экстремистік және террористік ұйымдарға тарту) мультиклассификациясының үлгісі құрылды. Корпус мәліметтеріне word2vec, tf-idf алгоритмдер қолданылып,</p>

	<p>визуализация жасалды. Веб-ресустардағы мәтіндер негізінен құрылымданбаған күйде болатындықтан және әр түрлі қолданушылар арқылы толтырылатындықтан, ол жақта орфографиялық қателіктер көптеп кездеседі. Сол себепті ең алдымен қазақ тіліндегі орфографиялық қателерді түзетуге арналған Spell Checker негізіндегі әдіс ұсынылды. Аталған әдіс кез-келген іздеу жүйесінің өте пайдалы функциясы болып табылады. Ең қарапайым шешім - сөздіктің барлық позицияларын редакциялық қашықтықтың өсуіне қарай сұрыптау және тек алғашқы бірнеше позицияларды көрсету. Редакциялық қашықтық ретінде Левенштейн қашықтығын алуға болады. Левенштейн қашықтығы - бастапқы жолды мақсатты жолға түрлендіру үшін қажет таңбаларды енгізу/жою/өзгерту әрекеттерінің ең аз санын көрсетеді. Мәтіндегі идеологиялық бағытты анықтау үшін Logistic Regression, KNN, SVM, Naive Bayes, Decision Tree, Random Forest, Gradient Boosting сияқты классикалық машиналық алгоритмдері үшін салыстырмалы талдау жүргізілді. Spellchecker+Stemming+TF-idf+LSTM+BERT негізінде жаңа модель құрылды. BERT моделінің гиперпараметрлері: кіріс = 128 сөз немесе токен, сызықтық классификация = 4, bert_model_name = 'bert-base-multilingual-uncased', num_classes = 4, max_length = 128, batch_size = 64, num_epochs = 10, learning_rate = 2e-5. Сонымен қатар, идеологиялық мәтіндерге мультиклассификация есебін шешу үшін 2 терең оқыту алгоритмі (BERT және LSTM, Bert+linear) біріктіріліп қолданылды. LSTM үшін модель мәтін тізбегін кіріс ретінде әр тізбектің сәйкес ұзындықтарымен бірге қабылдайды. Ол мәтінді ендіреді (кірістірілген өлшем = 20), оны LSTM қабаты арқылы өңдейді (өлшемі = 64), соңғы жасырын күйді ReLU белсендірулері бар толық қосылған қабаттар арқылы жібереді және соңында жалғыз шығыс мәнін алу үшін сигма тәрізді белсендіруді қолданады. Біріккен модельдің гиперпараметрлері: кіріс = 128 сөз немесе токен, BERT = 768, LSTM = 256, Dropout = 0.2, Linear Classification = 4, bert_model_name = 'bert-base-multilingual-uncased', num_classes = 4, max_length = 128, batch_size = 64, num_epochs = 20, learning_rate = 2e-5.</p> <p>Күтілетін нәтижелер:  - - әлеуметтік желілер мен мессенджерлердегі мәтіндік, аудио және бейнематериалдардың экстремистік бағытын мультиклассификациялауға арналған машиналық және терең оқыту үлгілері (діни экстремизм, саяси экстремизм, ксенофобия т.б.), әлеуметтік желілерде және мессенджерлерде ең белсенді кибер насихаттауды анықтауға арналған, белгілердің түріне қарай таңдалатын әртүрлі терең нейрондық желілердің комбинациясы болып табылатын жаңа гибриді модель.</p>
Зерттеу тобы мүшелерінің аты-жөні, идентификаторлары (Scopus Author ID, Researcher ID, ORCID, бар болса)	1. Мусиралиева Шынар Женисбековна, ORCID: <a href="https://orcid.org/0000-0001-5794-3649">https://orcid.org/0000-0001-5794-3649</a> , Scopus-тағы профайл сілтемесі: <a href="https://www.scopus.com/authid/detail.uri?authorId=57202216979">https://www.scopus.com/authid/detail.uri?authorId=57202216979</a> , Web of Science профайл сілтемесі: <a href="https://www.webofscience.com/wos/author/record/2394890">https://www.webofscience.com/wos/author/record/2394890</a>

<p>және сәйкес профильдерге сілтемелер</p>	<ol style="list-style-type: none"> <li>2. Болатбек Милана Асланбекқызы, ORCID: <a href="https://orcid.org/0000-0002-2153-180X">https://orcid.org/0000-0002-2153-180X</a> , Scopus-тағы профайл сілтемесі: <a href="https://www.scopus.com/authid/detail.uri?authorId=57202834055">https://www.scopus.com/authid/detail.uri?authorId=57202834055</a> , Web of Science профайл сілтемесі: <a href="https://www.webofscience.com/wos/author/record/GZL-7318-2022">https://www.webofscience.com/wos/author/record/GZL-7318-2022</a></li> <li>3. Байсылбаева Қымбат Данияровна, ORCID: <a href="https://orcid.org/0000-0001-9753-0398">https://orcid.org/0000-0001-9753-0398</a> , Web of Science профайл сілтемесі: <a href="https://www.webofscience.com/wos/author/record/N-9664-2017">https://www.webofscience.com/wos/author/record/N-9664-2017</a></li> <li>4. Елтай Жастай Ыбрайұлы, Researcher ID: <a href="https://www.webofscience.com/wos/author/record/JNR-6763-2023">https://www.webofscience.com/wos/author/record/JNR-6763-2023</a> , ORCID: <a href="https://orcid.org/my-orcid?orcid=0000-0002-9275-7582">https://orcid.org/my-orcid?orcid=0000-0002-9275-7582</a> Scopus author ID: <a href="https://www.scopus.com/authid/detail.uri?authorId=57237959800">https://www.scopus.com/authid/detail.uri?authorId=57237959800</a></li> <li>5. Нарбаева Салтанат Муратбековна, ORCID: <a href="https://orcid.org/0000-0001-5230-3781">https://orcid.org/0000-0001-5230-3781</a> , Scopus-тағы профайл сілтемесі: <a href="https://www.scopus.com/authid/detail.uri?authorId=57216484412">https://www.scopus.com/authid/detail.uri?authorId=57216484412</a> , Web of Science профайл сілтемесі: <a href="https://www.webofscience.com/wos/author/record/KBA-1599-2024">https://www.webofscience.com/wos/author/record/KBA-1599-2024</a></li> <li>6. Медетбек Жанар ORCID: <a href="http://orcid.org/0000-0001-7536-5889">http://orcid.org/0000-0001-7536-5889</a></li> <li>7. Беккожаева А.</li> </ol>
<p>Жарияланымдар тізімі (URL, DOI көрсетілген)</p>	<ol style="list-style-type: none"> <li>1. Ш.Ж. Мусиралиева, Р.Қ. Оспанов, М.А. Болатбек, Ж.Б. Медетбек Профилерование пользователей социальных сетей на основе демографических данных. АЭБУ Хабаршысы, Том 3 № 62(2023), 133-144 б. <a href="https://vestnik.aues.kz/index.php/none/issue/view/95/116">https://vestnik.aues.kz/index.php/none/issue/view/95/116</a></li> <li>2. Мусиралиева Ш.Ж., Байспай Г.Б., Болатбек М.А., Сағынай М., Терейковский И.А. Веб-ресурстардағы экстремисттік мәліметтерді анықтауға арналған машиналық әдістерді оқыту және сынау үшін қазақ тіліндегі мәтіндер корпусын құру. ҚарМТУ университеті еңбектері, №3, 2023, 453-458 б. Рекомендовано КОКСОН, <a href="http://tu.kstu.kz/archive">tu.kstu.kz/archive</a></li> <li>3. Ш.Ж. Мусиралиева, М.А. Болатбек, М. Сағынай, Ж.Ы. Елтай, К.Б. Багитова. Экстремисттік мәліметтер түсінігі және экстремизмге қарсы күрес жобаларына жүйелік шолу. (КОКСОН) . Журнал NEWS OF THE NATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN PHYSICO-MATHEMATICAL SERIES, ISSN 1991-346X Volume 3. № 347 (2023). 112–130. <a href="https://journals.nauka-nanrk.kz/physics-mathematics/article/view/5792">https://journals.nauka-nanrk.kz/physics-mathematics/article/view/5792</a></li> <li>4. Ш.Ж.Мусиралиева, Ж.Б.Медетбек , М.А.Болатбек , Жумаханова А.Н., Ж.Ы.Елтай. ӘЛЕУМЕТТІК МЕДИАДАҒЫ ЭКСТРЕМИСТІК МАЗМҰНДЫ АНЫҚТАУ ЖӘНЕ ЗЕРТТЕУ. «Информатика и прикладная математика» конференция</li> </ol>

	<p>еңбектері, 26 – 27 қазан 2023 ж., 265-272 б., Алматы, Қазақстан</p> <p>5. Shynar Mussiraliyeva, Milana Bolatbek, Moldir Sagynay, Aygerim Zhumakhanova, Zhastay Yeltay, and Zhanar Medetbek. 2024. Identifying Cyber-Threatening Texts in the Kazakh Segment of Web Resources. In Proceedings of the 2023 7th International Conference on Advances in Artificial Intelligence (ICAAI '23). Association for Computing Machinery, New York, NY, USA, 68–72. <a href="https://doi.org/10.1145/3633598.3633610">https://doi.org/10.1145/3633598.3633610</a></p>
Патент туралы ақпарат	-
 An illustration of a desk setup. In the center is a computer monitor displaying three stylized figures wearing balaclavas and holding rifles. To the left of the monitor is a pen holder with a pen. To the right is a stack of books. The desk is a simple grey surface with a single leg on the right side.	