Жоба туралы қысқаша ақпарат

Жоба аты	AP19576868 «Жастар экстремизмін анықтау және заманауи ақпараттық кеңістікте жастардың қауіпсіздігін қамтамасыз етуге арналған модельдер мен әдістерді әзірлеу»
Жоба өзектілігі	Желідегі экстремистік мәтіндер экстремистік әрекеттерді, соның ішінде террористік актілерді жоспарлау және жүргізу құралы бола алады. Мұндай тілді анықтау қауіпсіздіктің ықтимал қауіптерін анықтауға мүмкіндік береді және олардың орындалуының алдын алуға көмектеседі. Интернет көбінесе экстремизм мен радикалдану идеяларын тарату үшін қолданылады. Экстремистік тілдік анықтау радикалдану белгілерін ерте анықтауға және ықтимал жағымсыз салдардан сақтандыруға мүмкіндік береді. Экстремистік тілді анықтау сөз бостандығы мен қоғамды ықтимал қауіптен қорғау арасындағы тепе-теңдік үшін қажет. Бұл заңды мәлімдемелер мен қоғамдық қауіпсіздікке қауіп төндіретін сөздерді ажыратуға мүмкіндік береді. Көптеген елдерде экстремистік қызмет пен тілді реттейтін заңдар бар. Экстремистік мәтінді анықтау заңдарды сақтауға және заңсыз әрекеттердің алдын алуға көмектеседі. Экстремистік мәтіннің анықтамасы пайдаланушылар үшін, әсіресе экстремистік идеялардың әсерінен осал болуы мүмкін жастар үшін қауіпсіз онлайн кеңістігін құруға ықпал етеді. Экстремистік тіл әлеуметтік шиеленісті, жанжал тудыруы мүмкін. Мұндай мәтіндерді анықтау жеккөрушіліктің таралуын болдырмауға көмектеседі және үйлесімді қоғам құруға ықпал етеді. Интернеттегі экстремистік тілді анықтау елдер мен ұйымдар арасындағы ынтымақтастықты қажет етеді. Бұл жаһандық қауіп-қатерлерді тиімді бақылауға және қарсы тұруға көмектеседі. Осы аспектілердің барлығы қоғамның қауіпсіздігін қамтамасыз ету, радикалданудың алдын алу және сөз бостандығы мен қоғамдық қауіпсіздікті қамтамасыз ету міндеті арасындағы тепе-теңдікті сақтау үшін интернеттегі экстремистік тілді анықтаудың маңыздылығын көрсетеді.
Жоба мақсаты	Жобаның мақсаты жастар арасында зорлық-зомбылық, ұлттық экстремизм, нәсілшілдік, буллингтің таралуын айқындау және оған қарсы іс-қимыл жасау үшін семантикалық талдау модельдері мен әдістерін, жастар арасында құқыққа қайшы сипаттағы идеологияның таралуына қарсы іс-қимыл жасау үшін желідегі трафикті мониторингілеу және талдау әдістерін зерттеу және әзірлеу, жастар үшін ықтимал қауіпті веб-ресурстардың тізімін жасау, қазақ тілі үшін психоэмоционалдық талдау әдістерін бейімдеу болып табылады
Жоба міндеттері	1. Жастарға бағытталған ұлттық, зорлық-зомбылық экстремизмі, қорқыту және нәсілшілдік мәтіндерін анықтау үшін жаңа модельдер мен әдістерді әзірлеу 1.1 Таңдалған бағыт бойынша қол жетімді мәтіндерді талдау және негізгі ақпарат көздерін анықтау

	1.2 Веб-ресурстардан деректерді жинау үшін талдаушыны әзірлеу 1.3 Жастарға бағытталған ұлттық, зорлық-зомбылық экстремизм, буллинг және нәсілшілдік мәтіндерінің корпусын құру 1.4 Корпустағы деректерді өңдеу 1.5 Веб-ресурстардағы ұлттық, зорлық-зомбылық экстремизмін, қорқытуды және нәсілшілдікті анықтау міндетін жақсарту үшін белгілер жиынтығын анықтау 1.5 Қазақ тіліндегі жастарға бағытталған ұлттық, зорлық-зомбылық экстремизм, буллинг және нәсілшілдік мәтіндерін айқындау үшін семантикалық талдаудың жаңа модельдері мен әдістерін әзірлеу 1.6 Қазақ тіліне арналған мәтіндерді психоэмоционалды талдау әдістерін бейімдеу 2. Желілік трафикті талдау мен бақылаудың жаңа әдістерін әзірлеу 2.1 Желілік деректерді жинау модулін әзірлеу. 2.2 Өңделген трафик журналдарын талдау модулін әзірлеу. 2.3 Машиналық оқыту негізінде желілік трафикті талдау және бақылау әдісін әзірлеу. 2.4 Жастарға қауіпті веб-сайттардың тізімін жасау 3. Жастар арасында зорлық-зомбылық, зорлық-зомбылық экстремизмі, нәсілшілдік және қорқытудың таралуын анықтау және оған қарсы іс-қимыл бағдарламалық қамтамасыз етуді әзірлеу 3.1 Сәулет дизайны 3.2 Серверлік және алдыңғы бөлігін іске асыру 3.3 Бағдарламалық өнімді сынау
Күтілетін және қол жеткізілген нәтижелер	Қол жеткізілген нәтижелер: Жастарға бағытталған ұлттық, зорлық-зомбылық экстремизмі, қорқыту және нәсілшілдік мәтіндерін анықтау үшін жаңа модельдер мен әдістер әзірленді. Веб-ресурстардағы экстремистік мәтіндерді анықтау бойынша отандық және шетелдік басылымдарда жаңа әдебиеттерге шолу жасалды. Жастарға бағытталған ұлттық, зорлық-зомбылық экстремистік, буллингтік және нәсілшіл мәтіндерді анықтаудың жаңа модельдері мен әдістері жасалды. Экстремистік мәтіндерді анықтау үшін тірек векторлық машиналар, аңғал Байес классификаторлары, кездейсоқ ағаш әдістері, шешім ағашы, жақын көршілердің k алгоритмі, логистикалық регрессия, градиентті күшейту сияқты Машиналық оқыту әдістерін қолдануға арналған бір мақала жарияланды. Интернеттегі жастарға бағытталған ұлттық, зорлық-зомбылық экстремизмі, буллинг және нәсілшілдікке қатысты

	мәтіндерді жіктеудің қолданыстағы әдістеріне кең шолу жасалды. Шолу Springer, Elsevier және Web of Science және Scopus дерекқорларына енгізілген басқалар сияқты жоғары бағаланған ғылыми журналдарда жарияланған соңғы жарияланымдарды қамтиды. Әдебиеттерді талдау осы саладағы зерттеулердің қазіргі жағдайын анықтауға және жобамыздың өзекті бағыттарын анықтауға көмектесті. Бұл талдау және экстремистік мәтіндерді анықтау әдістерінің қазіргі жағдайына шолу осы салада жұмыс істейтін зерттеушілер мен инженерлер қауымдастығы үшін пайдалы болады. Бұл оларға осы әдістерді өз жұмыстарында тиімдірек қолдануға және осы маңызды саланың дамуына үлес қосуға мүмкіндік береді. Интернеттегі ұлттық, зорлық-зомбылық экстремизміне, қорқытуға және нәсілшілдікке қатысты мәтіндерді анықтау үшін машиналық оқытудың дәстүрлі әдістері, трансформаторларға негізделген әдістер мен модельдер жасалды. Іздеу машиналарының көмегімен әртүрлі веб-ресурстарда (Вконтакте, Twitter, YouTube, Telegram әлеуметтік желілері, блогтар, форумдар, жаңалықтар мақалалары) жалпыға қолжетімді мәтіндер анықталды және зерттелді. Осы зерттеу нәтижесінде ұлттық, зорлық-зомбылық экстремизміне, қорқытуға және нәсілшілдікке қатысты негізгі тіркестер мен мәтіндердің бастапқы көздері анықталды. Анықталған кілт сөздерді пайдалана отырып, Вконтакте, Twitter, YouTube, Telegram әлеуметтік желілерінің веб-ресурстарынан мәтіндерді жинау үшін талдаушы әзірленді. Талдаушының кірісіне дереккөздің домендік атауы, сақтау орны және қарау мерзімі беріледі. Нәтижесінде аталған веб-ресурстардың мәтіндік мазмұны жүктеледі. API технологиялары қолданылды. Құрастырылған талдаушының нәтижесінде Вконтакте, Twitter, YouTube, Telegram әлеуметтік желілеріндегі топтар мен арналардың мазмұнынан жиналған мәтіндік корпус құрылды. Корпусқа 5 санат кіреді: ұлттық экстремизм, зорлық-зомбылық экстремизмі, нәсілшілдік, қорқыту және бейтарап санаттағы мәтіндер, әр санатқа тиісті белгілер беріледі (0-ден 4-ке дейін). Корпус мәтіндеріне лингвистикалық және статистикалық талдау жүргізілді. Корпустың жалпы көлемі шамамен 10 000 мәтінді құрайды. Жиналған мәтіндік корпустың мәтіндері үшін алдын ала өңдеу алгоритмдері орындалды: токенизация, мәтінді морфологиялық талдау, стемминг, тыныс белгілерін жою, мәтіндегі сандық мәндер мен гиперсілтемелерді жою, тоқтату сөздерін жою. TF-idf, tf-IDF-bigram, bag-of-words сияқты жастарға бағытталған ұлттық, зорлық-зомбылық экстремизмі, қорқыту және нәсілшілдік мәтіндерін анықтаудың дәлдігін арттыратын белгілер анықталды. Бұл белгілер мәтіндегі деструктивті мазмұнды анықтауға қатысты модельдер мен әдістерді құрастыруда қолданылады. Decision Trees, Random Forest, logistic Regression, Naïve Bayes, Support Vector Machine, LSTM, BiLSTM Машиналық оқыту әдістерінің негізінде ұлттық, зорлық-зомбылық экстремизмін, қорқытуды және жастарға бағытталған нәсілшілдікті анықтау үшін деректерді семантикалық талдаудың жаңа әдістері мен модельдерін әзірлеу бойынша жұмыс жүргізілді. Stemming+TF-IDF+BERT негізінде модель жасалды. Сонымен қатар, осы санаттағы мәтіндерді анықтаудың дәлдігін арттыру үшін психоэмоционалды талдау моделі

	құрылды. Distilbert және Roberta сияқты трансформаторлар негізінде модель жасалды. Роберта оқудың гиперпараметрлерін мұқият және ақылға қонымды оңтайландыру арқылы БЕРТТІ жақсартады. RoBERTa моделі Pytorch-та жасалған. Модельдің гиперпараметрлері: Input = 128 сөз немесе токен, RoBERTa = 1280 vector, Linear = 768, DropOut = 0.1, linear Classification = 5. model_name = "xlm-roberta-base", num_classes = 5, max_length = 128, batch_size = 64, num_epochs = 20, learning_rate = 2e-5, val_size=0.2, test_size=0.2 ұлттық, зорлық-зомбылық мәтіндерін анықтауға арналған семантикалық талдау моделі distilbert негізінде қазақ тіліндегі экстремизм, буллинг және нәсілшілдік бағытталды өлшемді азайту және жылдамдықты арттыру арқылы оқуды оңтайландыру үшін мұның бәрі өнімділікті сақтау мақсатында жасалды. Модельдің гиперпараметрлері: Input = 128 сөз немесе токен, DistilBERT = 768 Вектор, Linear = 768, DropOut = 0.1, linear Classification = 5. model_name = "distilbert-base-uncased", num_classes = 5, max_length = 128, batch_size = 64, num_epochs = 20, learning_rate = 2e-5, val_size=0.2, test_size=0.2. Сондай-ақ, жастарға бағытталған ұлттық, зорлық-зомбылық экстремизмін, қорқытуды және нәсілшілдікті анықтау үшін MLM моделі (маскирленген тіл моделі - "маскаланған тілді модельдеу") құрылды. Модельдің гиперпараметрлері: Input = 128 сөз немесе токен, MLM = 1280 Вектор, Linear = 768, DropOut = 0.1, linear Classification = 5. model_name = "xlm-mlm-100-1280", num_classes = 5, max_length = 128, batch_size = 64, num_epochs = 20, learning_rate = 2E-5, val_size=0.2, test_size=0.2 Мәтіндерді психоэмоционалды талдаудың белгілі әдістері қазақ тіліне бейімделді, жастарға бағытталған ұлттық, зорлық-зомбылық экстремизм, буллинг және нәсілшілдік мәтіндеріндегі психоэмоционалды лексемалардың анализаторы жасалды. Зерттеу жұмысында экстремистік лингвистикалық корпус сөздерді санау стратегиясы мен liwc жабық сөздік әдісін қолдана отырып талданды. Ұсынылған әдістің міндеті-мәтіндік мәліметтер жиынтығындағы психологиялық категорияларға қатысты сөздерді іздеу және санау. Барлығы 80-нен астам санат бөлінді. Бағдарламадағы мәтіндік файлды өңдеудің нәтижесі келесі Шығыс айнымалылары болып табылады: сөздердің саны, жиынтық тілдік айнымалылар (аналитикалық ойлау, әсер ету, мәтіннің ерекшелігі және эмоционалды тон) және мәтіндегі сөздердің пайызы, бұл мәтіндегі сөздердің пайызын білдіреді (мысалы, есімдіктер, мақалалар, көмекші етістіктер және т.б.). д.), психологиялық құрылымдарға әсер ететін категориялар (мысалы, аффект, таным, биологиялық процестер, импульстар), жеке қызығушылық категориясы (мысалы, жұмыс, үй, демалыс), бейресми тіл маркерлері (мысалы, қарғыс). Нәтижелер сөздерді қолдану тәсілдерінде тұлғаның даму деңгейлерін білдіру туралы жалпы теориялық білімнің пайдасын көрсетеді. Әзірленген әдіс арқылы ұлттық экстремизм, зорлық-зомбылық экстремизмі, нәсілшілдік және буллин мәтіндері бойынша жіктеу үшін LSTM негізіндегі модель жасалды. Құрастырылған модульге авторлық куәлік алынды. Күтілетін нәтижелер:

	Жастар үшін қауіпті веб-ресурстарды анықтау мақсатында желілік трафикті талдау және мониторингілеу үшін жаңа әдістер әзірленетін болады. Жастарға бағытталған ұлттық, зорлық-зомбылық экстремизмі, қорқыту және нәсілшілдік мазмұны бар веб-ресурстарды анықтауға мүмкіндік беретін бағдарламалық қамтамасыз ету әзірленеді.
Зерттеу тобы мүшелерінің аты-жөні, идентификаторлары (Scopus Author ID, Researcher ID, ORCID, бар болса) және сәйкес профильдерге сілтемелер	1. Болатбек Милана Асланбекқызы, ORCID: https://orcid.org/0000-0002-2153-180X , Scopus-тағы профайл сілтемесі: https://www.scopus.com/authid/detail.uri?authorId=57202834055 , Web of Science профайл сілтемесі: https://www.webofscience.com/wos/author/record/GZL-7318-2022 2. Байсылбаева Кымбат Данияровна, ORCID: https://orcid.org/0000-0001-9753-0398, Web of Science профайл сілтемесі: https://www.webofscience.com/wos/author/record/N-9664-2017 3. Сағынай Мөлдір, ORCID: https://orcid.org/0009-0004-1377-5742 4. Елтай Жастай ЫбрайұлыResearcher ID: https://www.webofscience.com/wos/author/record/JNR-6763-2023 , ORCID: https://orcid.org/my-orcid?orcid=0000-0002-9275-7582 Scopus author ID: https://www.scopus.com/authid/detail.uri?authorId=57237959800 5. Ахмед Гүлмарал Жалғасбекқызы, ORCID: https://orcid.org/0000-0002-4464-9544 6. Мейрбекова Бибинур Калдыбаевна, ORCID: https://orcid.org/0000-0001-9215-9382 , Scopus-тағы профайл сілтемесі: https://www.scopus.com/authid/detail.uri?authorId=57212476113 , Web of Science профайл сілтемесі: https://www.webofscience.com/wos/author/record/ABD-4499-2021 7. Шайзат Медет Жанболатұлы, ORCID: https://orcid.org/0000-0002-1651-8205 , Scopus-тағы профайл сілтемесі: https://www.scopus.com/authid/detail.uri?authorId=57216968174 8. Райымкулова Алима Мухамбеткалиевна
Жарияланымдар тізімі (URL, DOI көрсетілген)	1. Scientific Journal of Astana IT University ISSN (P): 2707-9031 ISSN (E): 2707-904X VolUmE 14, JUNE 2023, COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORTMS TO IDENTIFY EXTREMIST TEXTS IN THE KAZAKH LANGUAGE,DOI: 10.37943/14DKRN4681, Shynar Mussiraliyeva , Milana Bolatbek ,Aigerim Zhumakhanova ,Zhanar Medetbek , Moldir Sagynay https://journal.astanait.edu.kz/index.php/ojs/article/view/344 2. Болатбек М.А., Сағынай М., Мусиралиева Ш.Ж., Байсылбаева К.Д., Шайзат М.Ж. Қазақ тіліндегі мәтінге психо-эмоционалдық талдау жүргізуге арналған әдісті құру және

	зерттеу, VIII — Международная научно-практическая конференция «Информатика и прикладная математика» https://conf.iict.kz/wp-content/uploads/2023/11/collection_CSAM_VIII_2023_2.pdf 3. Shynar Mussiraliyeva, Milana Bolatbek, Aygerim Zhumakhanova, Moldir Sagynay, Development of a software module for collecting and analyzing web content to determine extremist direction in the text принята к публикации в 17th International Conference on Information Technology and Applications (ICITA2023) https://link.springer.com/book/9789819983230 4. М.А.Болатбек, К.Д.Байсылбаева, М.Сағынай, Ш.Ж. Мусиралиева, А.Н.Жумаханова, Интернет кеңістігіндегі жастарға бағытталған деструктивті мәтіндерді жинақтауға қажетті парсер бағдарламасын әзірлеу, Известия НАН РК. Серия физико-математическая, №4, 2023 г. https://journals.nauka-nanrk.kz/physics-mathematics/article/view/5925 5. Bolatbek, Milana, and Shynar Mussiraliyeva. "Detection of Extremist Messages in Web Resources in the Kazakh Language." Lodz Papers in Pragmatics, vol. 19, no. 2, Dec. 2023, pp. 415–425, doi:10.1515/lpp-2023-0020. https://journals.scholarsportal.info/details/18956106/v19i0002/415_doemiwritkl.xml
Патент туралы ақпарат	-
 	


мұсылман
шабуыл
бол
мұжахид
пәк
американдық
ауған
болған
талиб
үкімет
айт
адам
соғыс
кез
көп
ислам
күш
он
сом
ақ
баста
қала
ел өлтір
басқа
ем
күрес
халық
қарсы
аймақ
жағдай
джихад
жаңа
қазір
келе
сия
дей
топ
деген
кел
армия
ретінде
алла
жиһад
әскери
әскер
қауіпсіздік
байланысты