

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ
БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РЕСПУБЛИКИ КАЗАХСТАН



ҚазҰТЗУ ХАБАРШЫСЫ _____

_____ **ВЕСТНИК КазННТУ**

VESTNIK KazNRTU _____

№ 5 (141)

Главный редактор
И. К. Бейсембетов – ректор

Зам. главного редактора
А.Х. Сыздыков – проректор по науке

Отв. секретарь
Н.Ф. Федосенко

Редакционная коллегия:

З.С. Абишева- акад. НАН РК, Л.Б. Атымтаева, Ж.Ж. Байгунчечков- акад. НАН РК, А.Б. Байбатша, А.О. Байконурова, В.И. Волчихин (Россия), К. Дребенштед (Германия), Г.Ж. Жолтаев, Г.Ж. Елигбаева, Р.М. Исаков, С.Е. Кудайбергенов, Б.У. Куспангалиев, С.Е. Кумекон, В.А. Луганов, С.С. Набойченко – член-корр. РАН, И.Г. Милев (Германия), С. Пежовник (Словения), Б.Р. Ракишев – акад. НАН РК, М.Б. Панфилов (Франция), Н.Т. Сайлаубекон, А.Р. Сейткулов, Фатхи Хабаши (Канада), Бражендра Мишра (США), Корби Андерсон (США), В.А. Гольцев (Россия), В. Ю. Коровин (Украина), М.Г. Мустафин (Россия), Фан Хуаан (Швеция), Х.П. Цинке (Германия), Е.М. Шайхутдинов-акад. НАН РК, Т.А. Чепуштанова

Учредитель:

Казахский национальный исследовательский технический университет
имени К.И. Сатпаева

Регистрация:

Министерство культуры, информации и общественного согласия
Республики Казахстан № 951 – Ж “25” 11. 1999 г.

Основан в августе 1994 г. Выходит 6 раз в год

Адрес редакции:

г. Алматы, ул. Сатпаева, 22,
каб. 609, тел. 292-63-46
Nina. Fedorovna. 52 @ mail.ru

Мусина Э. Б.

«Карцемент» АҚ цемент кәсіпорны үлгісінде өндірістік жарақаттанудың себептерін талдау және алдын алу

Түйіндемесі. Мақала «Карцемент» АҚ мысалын қолдана отырып цемент зауытындағы авариялардың себептерін жан-жақты зерттеуге арналған. Жұмыста жарақаттану динамикасы, саны, жиілігі, ауырлығы және себептері туралы жүйелі мәліметтер ұсынылды. Жазатайым оқиғалар саны зардап шеккендердің жас санаттары, кәсіптері, диагноздары және өндірістік жарақатқа әкеп соқтырған авариялар түрлеріне сәйкес бөлінді. Өндірістегі жазатайым оқиғаларды талдау нәтижелері өндірістегі жазатайым оқиғаларды болжауға негіз бола алады және оны болдырмауға бағытталған нақты іс-шараларға нұсқаулық болады. оның туындауының басқарылатын себептерін жою, кәсіби тәуекелдерді төмендету, алдын-алу шараларынан ең төменгі шығындардан жоғары тиімділікті қамтамасыз ету. Мақаланың соңында автор кәсіпорындағы жазатайым оқиғалардың санын азайту шараларын ұсынды.

Түйін сөздер: еңбекті қорғау, цемент кәсіпорны, жазатайым оқиға, өндірістік жарақаттану деңгейі, статистикалық және топтық әдістер, жиілік пен ауырлық коэффициенттері.

ӘОЖ 81.93.29

Ш.Ж. Мусиралиева, Б.С. Омаров, Ж.Б. Медетбек, Ғ.Р. Қараман, А.К. Бекетова

IDENTIFICATION OF MESSAGES CONTAINING ELEMENTS OF EXTREMISM IN SOCIAL SYSTEMS USING MACHINE LEARNING METHODS

Abstract: Along with the growth of social networks, the number of religious hatred and racism on the Internet is growing. In addition, the activity of radical groups on the Internet calling for violence and extremism is one of the most important issues of public safety. Because for such structures the main tool for information exchange, recruitment and promotion is the Internet, in particular, web resources, social networks, social messengers, etc. In this regard, it is necessary to identify individual users, groups and Internet communities, who create and distribute terrorist and extremist information on the Internet, as well as prevent the spread of extremist materials.

This work is devoted to the study and development of machine learning methods aimed at solving the problem of identifying extremist text in social systems. In addition, models and methods for identifying extremist text are presented, which are used for in-depth linguistic analysis and statistical processing of texts. To classify text as extremist or non-extremist based on social media posts left by a user on the internet, we create a text classification system using sentiment analysis techniques based on machine learning.

Keywords: social network, online extremism, radical text, linguistic analysis, machine learning, big data, vector model, logistic regression, naive Bayes, deep learning.

Ш.Ж. Мусиралиева, Б.С. Омаров, Ж.Б. Медетбек, Ғ.Р. Қараман, А.К. Бекетова

(Казахский национальный университет имени аль-Фараби,

Алматы, Республика Казахстан,)

E-mail: medetbek.zhanar@gmail.com

МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІН ҚОЛДАНУ АРҚЫЛЫ ӘЛЕУМЕТТІК ЖҮЙЕЛЕРДЕГІ ЭКСТРЕМИЗМ ЭЛЕМЕНТТЕРІН ҚАМТИТЫН ХАБАРЛАМАЛАРДЫ АНЫҚТАУ

Аннотация: Әлеуметтік желілердің өсуімен қатар, сәйкесінше, желідегі діни жеккөрушілік пен нәсілшілдік саны артып келеді. Сондай-ақ, зорлық-зомбылыққа, экстремизмге шақыратын радикалды топтардың желідегі белсенділігі қоғам қауіпсіздігі үшін маңызды проблемалардың бірі болып табылады. Өйткені мұндай құрылымдар үшін ақпарат алмасудың, рекрутингтің және насихаттаудың негізгі құралы Интернет желісі, атап айтқанда веб ресурстар, әлеуметтік желілер, әлеуметтік мессенджерлер және т.б. болып табылады. Осыған байланысты интернетте террористік және экстремистік ақпаратты тудыратын және тарататын жекелеген пайдаланушыларды, топтарды және желілік қауымдастықтарды анықтау, сондай-ақ экстремистік материалдардың таралуының жолын кесу міндеті туындайды.

Бұл жұмыс әлеуметтік жүйелерде экстремистік мәтінді анықтау мәселесін шешуге бағытталған машиналық оқыту әдістерін зерттеуге және әзірлеуге арналған. Сонымен бірге, экстремистік мәтінді анықтаудың модельдері мен әдістері ұсыналады, олар мәтіндерді терең лингвистикалық талдау мен статистикалық өңдеу үшін қолданылады. Интернет желілерінде қолданушы қалдырған әлеуметтік медиа жазбаларына сүйене отырып,

мәтінді экстремистік немесе экстремистік емес деп жіктеу үшін машиналық оқытуға негізделген пікірлерді талдау (sentiment analysis) әдістерін қолдана отырып, мәтінді жіктеу жүйесін жасаймыз.

Кілттік сөздер: әлеуметтік желі, онлайн экстремизм, радикалды мазмұндағы мәтін, лингвистикалық талдау, машиналық оқыту, үлкен көлемді деректер, векторлық модель, логистикалық регрессия, аңқау Байес, терең оқыту.

I. КІРІСПЕ

Соңғы жылдары қылмыстық іс-әрекеттер, әсіресе терроризм күшейе түсті. Бүгінде экстремистердің көпшілігінде әлеуметтік желілерде аккаунттары мен чаттары бар, олар виртуалды қоғамдарға өмірге белсене қатысады. Әлеуметтік желілерді радикалды идеялар мен экстремистік қауіптерді тарату үшін қолдану мәселесі соңғы 10 жылда осы саланы зерттеушілердің назарын аударып келеді. Интернет – бұл, ашық ресурс болғандықтан, оны пайдалану - кез-келген ақпаратты жылдам, әрі жасырын түрде таратуға, әлеуметтік желілер мен форумдардың аудиториясымен тікелей байланысуға мүмкіндік береді. Бүгінгі таңда әйгілі әлеуметтік желілердің бүкіл әлемде кең ауқымы және аудиториясы бар: Facebook айына 2,5 миллиардтан астам белсенді қолданушыларды, ал YouTube ай сайын 2 миллиардқа жуық қолданушыларды қолдайды. Сол сияқты Twitter-де айына орта есеппен 386 миллион белсенді қолданушысы бар. Сонымен қатар, танымал Instagram және WhatsApp әлеуметтік желілерінде миллиардтаған пайдаланушылар тіркелген [0].

Зерттеу жұмысында көрсеткендей **[Ошибка! Источник ссылки не найден., Ошибка! Источник ссылки не найден.]**, 2015 жылы Twitter-де террористік әрекеттерге байланысты 125000-нан астам аккаунттар табылған. Твиттердегі экстремистік мазмұндағы посттардың өсуіне байланысты, компания әр түрлі елдердегі (Америка Құрама Штаттары (АҚШ), Ирландия және т.б.) кәсіпқойлар тобын құрды, олар күдікті аккаунттарды қадағалап, олар табылған кезде оларды бұғаттайды.

Желі арқылы таратылатын ақпараттың үлкен көлеміне, оның тілдік әртүрлілігіне және нақты уақыттағы мониторингке қойылатын талаптарға байланысты, ықтимал қауіпті пайдаланушыларды анықтау, экстремистік материалдарды уақытылы өшіру, экстремистер немесе олардың шабуылдары туралы барлық ақпаратты талдау үшін мәтінді талдаудың автоматтандырылған процедураларын қолдану қажет.

Машиналық оқыту қолданылу аумағының кең спектріне байланысты, үлкен көлемді деректерді өңдей отырып, болжау мақсаттарында жиі қолданылады.

Үлкен көлемді деректер [4] термині, әдеттегі деректер қорының мүмкіндіктерінен асып түсетін деректер жиынтығының өлшемін ғана емес, сонымен қатар өңдеу және талдау алдында дәстүрлі алгоритмдер әлсіз болып табылатын құрылымдалмаған ақпараттар есептеледі. Машиналық оқытудың кейбір заманауи алгоритмдері үлкен ауқымды мәліметтер жиынтығын оқыта отырып, машиналық оқытудың жақсы алгоритмдерін жасауға мүмкіндік береді.

Бұл мақалада әр түрлі жіктеу әдістерінің жұмыс сапасын алгоритмнің дәлдігі, толықтығы, уақыты, алгоритмнің өсу режимінде жұмыс істеу қабілеті, жіктеуге қажетті алдын-ала ақпараттың көлемі, тілдің тәуелсіздігі сияқты сипаттамалары бойынша талдау және салыстыру әрекеттері жасалды.

2. ӘДЕБИЕТТЕРГЕ ШОЛУ

Бұл бөлімде біз әлеуметтік желілердегі экстремистік мәтінді жіктеу бойынша жүргізілген зерттеулерге шолу жасаймыз.

Зерттеулер көрсеткендей, экстремистік топтар негізінен қате хабарламалар, жағымсыз сөздер мен дөрекі пікірлер таратады. Көптеген жек көрушілікті насихаттайтын топтар өздерінің идеологиясын насихаттау үшін әсіресе танымал әлеуметтік желілерді қолданады, олардың экстремистік мазмұнын көрермендеріне таратады [5]. Олар әлеуметтік желілерді өз тобына жаңа мүшелерді тартуды жеңілдету үшін құрал ретінде пайдаланады, біртіндеп әлемдік аудиторияны қамтып, басқаларды зорлық-зомбылық пен терроризмге итермелейді [6].

Интернеттегі экстремизмді анықтау мәселесі ХХІ ғасырдың басынан бастап зерттеушілердің назарын аударған. Бүгінгі күні осы тақырып бойынша көптеген мәліметтер мен әдебиеттер бар, сондай-ақ осы саладағы қолданыстағы шешімдер мен әдістерде жетерлік.

Мәтіндік классификация - экстремистік контентті анықтаудың ең танымал әдістері болып табылады. Мәтіндік классификация лингвистикалық ерекшеліктерді зерттейді және машинаны оқыту классификаторының негізінде шешім қабылдайды. Бұл тәсіл біздің зерттеу жұмысымызда

қолданылатын әдіске сәйкес келеді. Машиналық оқудағы тапсырмалардың шамамен 70% -ы классификация міндеттеріне жатады.

Соңғы онжылдықтарда табиғи тілді өңдеу (NLP) және пікірлерді талдау (sentiment analysis) саласы жылдан-жылға дамуда. KNN, Naive Bayes әдісі, EDA, мәліметтерді кластерлеу, Decision Tree, Gradient Boosted Decision Tree (GBDT) және Deep Neural Network (DNN) - әлеуметтік желілерде радикалдықты анықтайтын ең кең таралған әдістер болып табылады.

Әлеуметтік желілердегі жазбаларға сүйене отырып, Azizan мен Aziz [**Ошибка! Источник ссылки не найден.**] машиналық оқыту әдістерін қолдана отырып, экстремистік мәтінді анықтау үшін зерттеу жүргізді, дәлірек айтсақ, Naive Bayes алгоритмін қолданды. Ол басқа машиналық оқыту классификаторларымен салыстырғанда жақсы нәтижелер көрсетті. Деректерді жинауда Twitter streaming API пайдаланылған. Бұл зерттеу жұмысында деректер мына критерийлер бойынша, яғни “ISIS”, “Muslim”, “bomb”, “terrorist” және т.б. террористтік кілттік сөздері бойынша жинақталды. Сонымен қатар, олар экстремистік топтарға қатысты пайдаланушылардың пікірлерін оң (positive sentiment) және теріс (negative sentiment) деп жіктеді. Алайда, твиттерді оң және теріс класстар бойынша жіктеу экстремистік және экстремистік емес твиттерді ажыратудың тиімді әдісін ұсынбайды.

Біздің мәселеге жақын мәселені шешу тәсілдері келесі жұмыста [**Ошибка! Источник ссылки не найден.**] қарастырылған. Мұнда Twitter-де жиһадшыл топтар шығарған радикалды мазмұнды мәтінді автоматты түрде анықтау тәсілдері ұсынылған. Ол үшін SVM, AdaBoost, Naive Bayes әдістерін қолдана отырып, радикалды және радикалды емес деп жіктеу нәтижелерін салыстырып көрсеткен.

Орыс тілінде экстремистік мәтінді анықтаудың модельдері мен әдістері [**Ошибка! Источник ссылки не найден.**] зерттеу жұмысында көрсетілген. Сондай-ақ, орыс тіліндегі террористік, діни жеккөрушілік, нәсілшілдік және басқа радикалды мәтіндер жиынтығын және осы деректер қорымен жүргізілген тәжірибелердің нәтижелері ұсынылған. Орыс тілінде дайын экстремистік корпус болмағандықтан, авторлар зерттеу үшін қолдан корпус құрған. Корпуста 493 мәтін бар (650 000 сөз), олардың 368-і экстремистік, 125-і нейтралды. Экстремизм ұғымы кең, сондықтан олар экстремистік мәтінді 7 санатқа жіктеді: Терроризм (27 мәтін, 3296 сөз), Идеологиялық мәтіндер (26 мәтін, 21 131 сөз), Діни өшпенділік (55 мәтін, 16 697 сөз), Сепаратизм (7 мәтін, 852 сөз), Ұлтшылдық (208 мәтін, 19 399 сөз), Агрессия және көтеріліске үндеулер (43 мәтін, 6757 сөз), Фашизм (13 мәтін, 2059 сөз). Зерттеуде 2 мәселе қарастырылады: Экстремистік мәтіндерді анықтау, Экстремистік тақырыпты анықтау. Экстремистік мәтіндерді анықтауда бинарлы классификация бойынша шешу, яғни экстремистік немесе нейтралды мәтін деп бөлу ұсынылады. Ал экстремистік тақырыпты анықтауда жоғарыда көрсетілген 7 санат бойынша жіктеу ұсынылады. Екі тапсырма бойынша да логистикалық регрессия, SVM, Random Forest, Gradient boosting машиналық оқыту әдістерін қолданған.

Мәтінді классификациялау мәселесін шешуге арналған ұқсас тәсіл [**Ошибка! Источник ссылки не найден.**] жұмыста ұсынылған. Деректерді жинауда Twitter Streaming API пайдаланылды. Корпус 2018 жылы Ауғанстандағы Кундуз медресесіне шабуылдарға байланысты Твиттердегі посттардың негізінде құрылған. Корпус 3380 посттан тұрады. Мәтіндік классификация үшін Naive Bayes, SVM, Decision Tree, Random Forest, KNN, Bagging, Boosting машиналық оқыту әдістері пайдаланылған.

Келесі зерттеу жұмысында [**Ошибка! Источник ссылки не найден.**] терең оқыту (Deep learning) әдістерін пайдаланып классификация жасалынған. Twitter-дегі пайдаланушылардың хабарламаларына негізделген экстремистік және экстремистік емес класстарға жіктеу жүйесі ұсынылған. Корпус 25000 посттан тұрады. Оның 80%-ы жүйені жаттықтыруға бағытталған, ал 20%-ы тестілеу үшін. Модельді жаттықтыру үшін 12754 “экстремистік” және 8432 “экстремистік емес” деп таңбаланған мәтіндер қолданылған. Авторлар өздерінің LSTM+CNN терең оқыту моделін ұсынып, басқа ML және DL әдістерімен салыстырған. Нәтижесінде, авторлар ұсынған модель басқа модельдерге қарағанда ең жоғарғы дәлдікті көрсеткен.

3. Корпус құру

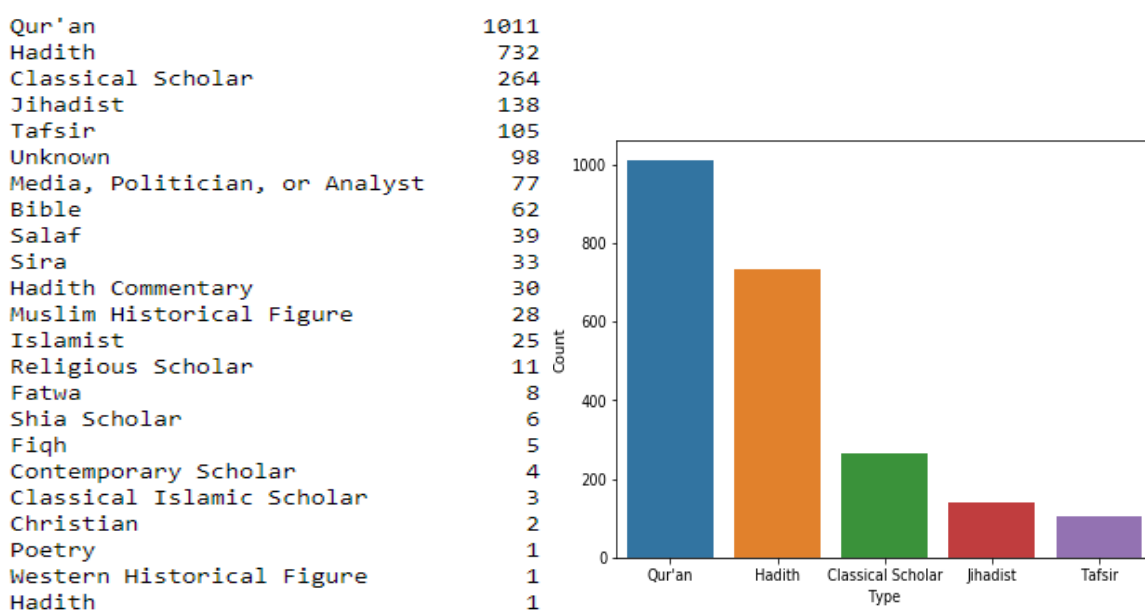
Бұл жұмыста Dark Web Portal Project әр түрлі тілдегі 28 форум мәліметтерін қамтитын жобаның бір форумына талдау жасалған [8]. Аталған форумдарда экстремистік және жалпы діни мәселелер талқыланады, олардың бірқатары радикалды ислами сипаттағы мәтіндер болып келеді. Аталған форумдардың ішіндегі Ansar1 форумына тоқталған. Ол жерде барлығы 29942 хабарлама бар. Хабарламалар 2008-2010жж. мерзімі аралығында жинақталған. Қолданушылардың жалпы саны - 382. Негізгі тілі ағылшын тілі болса да, арасында араб, орыс, түрік және т.б. ұлттарының тіліндегі хабарламалар кездеседі. Талдаудың дәлдігін арттыру мақсатында келесі тапсырмалар орындалды: -

экстремистік мазмұндағы хабарламалардан жеке корпусы құрылды, машиналық оқыту әдістерін оқыту үшін және аталған әдістер арқылы жіктеу жұмыстарын жүргізу үшін пайдаланылады; - корпустағы мәтіндердің орфографиялық қателері түзетілді, барлық тыныс белгілері мен сілтемелер өшірілді; - корпустағы мәтіндерге стемминг алгоритмі қолданылды, яғни мәтіндегі әр сөздің жалғаулары мен жұрнақтары алынып тастап, тек негізі ғана қалдырылады; - нәтижесінде корпустағы мәтіндердің жалпы саны мыңнан асады [9].

Мәтінді классификация жасау үшін ең алдымен корпус құратырдық. Біз экстремистік мазмұндағы мәтіндерден және нейтралды мәтіндерден тұратын корпус құрастырылды. Экстремистік мәтін ретінде ағылшын тіліндегі ИГИЛ экстремисттік тобы қолданған религиялық мәтіндерді [10] алдық. Бұл діни мәтіндер ИГИЛ идеологиясында, үгіт-насихаттауда және адамдар тартуда шешуші рөл атқарады. Корпустағы экстремистік мәтіндердің жалпы саны, жоғарыдағы дерек көздерін қосқанда - 2700 шамасында.

Біздің жұмысымызда деректерді зерттеу үшін төрт әдісті қолдандық: TF-IDF және сөз қапшығы (Bag of Words) негізінде логистикалық регрессия (logistic regression), тірек вектор әдісі (SVM), және кездейсоқ орман алгоритмі (Random Forrest), аңқау Байес әдістері (naive Bayes classifier).

Мәтіндер типі бойынша келесі санаттарға бөлінген:



1-сурет. Корпустағы радикалды сөздердің типтік санаттары

4. Экстремистік мәтінді анықтау әдістері

4.1 Мәтіндерді таңбалау

Мәтінді классификация жасайтын модельді оқыту үшін ең алдымен оны таңбалау керек. Біз нейтралды мәтінді 0, экстремистік мазмұндағы мәтінді 1 деп таңбаладық.

4.2 Деректерді алдын-ала өңдеу

Деректер таңбалар, сілтемелер және тыныс белгілері түріндегі «шу» немесе «керексіз деректерден» тұратын өңделмеген түрінде болады. Деректердегі керек емес артық символдар модель үшін пайдасыз және классификаторлардың көрсеткішін төмендетуі мүмкін. Мұндай қажетсіз символдарды корпуста алып тастау үшін, біз төменде сипатталған деректерді өңдеудің бірнеше тапсырмаларын орындаймыз.

Токенизация – мәтінді сөздердің сегментіне түрлендіру. Модель үшін қажетсіз символдарды жою және сөздердің векторын құру үшін әрбір мәтінді токендерге бөлдік.

Сөздерді кіші регистрге ауыстыру – әр түрлі регистрдегі (үлкен немесе кіші) бірдей сөздердің қайталануын болдырмау үшін барлық токенді кіші регистрге түрлендірдік.

Бізге мәтіннің мағыналы бөлігі қажет, сондықтан тыныс белгілерін, алфавитке жатпайтын таңбаларды, сандарды алып тастадық.

Мәтінде тежеуіш сөздер, яғни бұл сөзді алып тастағанымен сөйлемнің мағынасы өзгермейтін сөздер, кездеседі. Бұндай сөздер оқу уақытын ұзартады және модельдің дәлдік көрсеткішін төмендетуі мүмкін. Сондықтан тежеуіш сөздерді алып тастадық.

Стемминг – сөздерден аффикстерді (жұрнақтар мен жалғауларды) жойып, сөзді бір формаға келтіру.

4.3 Сөздерді верторларға түрлендіру

Модельге математикалық және статистикалық зерттеулер жүргізу үшін машинаны оқыту алгоритміне функция векторы түрінде сандық сурет қажет. Векторизациялау кезінде Bag of Words және TF-IDF пайдаланылған. Bag of Words моделінде мәтін грамматика мен тіпті сөздердің ретін ескермей, бірақ сөздердің жиілігін сақтай отырып сөздердің қапшығы ретінде ұсынылады. Бұл модел мәтінді классификациялау әдістерінде жиі қолданылады, мұнда әр сөздің кездесуі (жиілігі) классификаторды оқытатын функция ретінде қолданылады. TF-IDF – мәтін контекстіндегі сөздің маңыздылығын бағалау тәсілі болып табылады.

4.4 Машиналық оқыту әдістерін пайдаланып модель құру

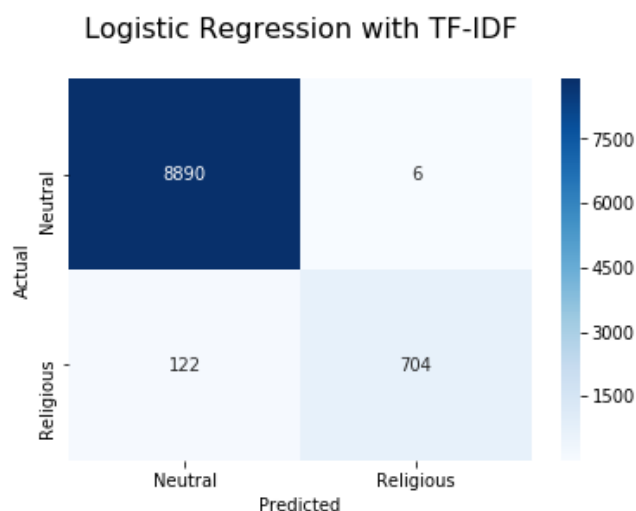
Экстремистік мәтінді анықтау үшін келесі машиналық оқыту алгоритмдері негізінде модель құрылды: Logistic Regression, Support Vector Machine (SVM), Random Forest, Naïve Bayes. Корпустың 0.7 бөлігі модельді жаттықтыруға, ал 0.3 бөлігі тестілеу үшін қолданылды. Бұл алгоритмдердің нәтижелері мен тиімділігі келесі бөлімде көрсетілген.

5. Нәтижелер

Бұл бөлімде алдыңғы бөлімде көрсетілген машиналық оқыту алгоритмдері негізінде құрылған модельдің болжам нәтижелері көрсетілген және осы нәтижелер бойынша салыстырылды.

```
f1_score:0.9426020408163265
```

	precision	recall	f1-score	support
0	0.99	1.00	0.99	8896
1	1.00	0.89	0.94	826
accuracy			0.99	9722
macro avg	0.99	0.95	0.97	9722
weighted avg	0.99	0.99	0.99	9722



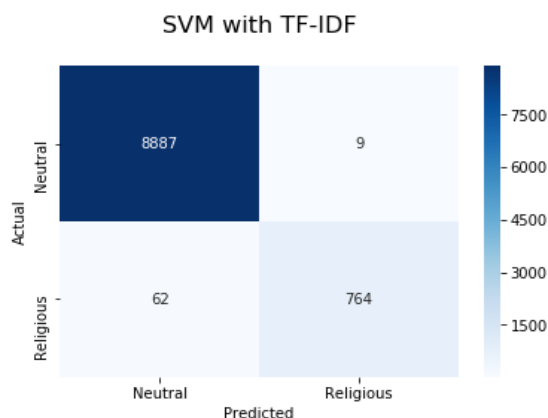
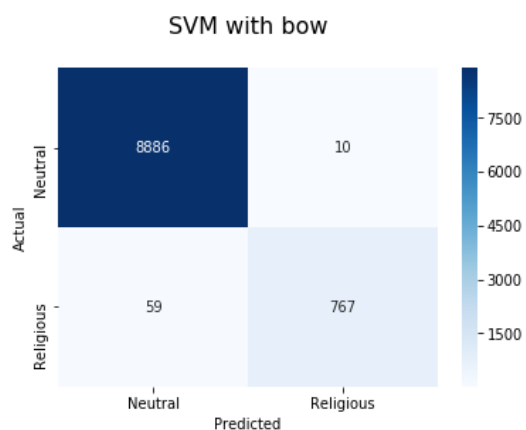
2 - сурет. Logistic Regression әдісінің нәтиже көрсеткіштері

f1_score:0.956955708047411

	precision	recall	f1-score	support
0	0.99	1.00	1.00	8896
1	0.99	0.93	0.96	826
accuracy			0.99	9722
macro avg	0.99	0.96	0.98	9722
weighted avg	0.99	0.99	0.99	9722

f1_score:0.955597248280175

	precision	recall	f1-score	support
0	0.99	1.00	1.00	8896
1	0.99	0.92	0.96	826
accuracy			0.99	9722
macro avg	0.99	0.96	0.98	9722
weighted avg	0.99	0.99	0.99	9722



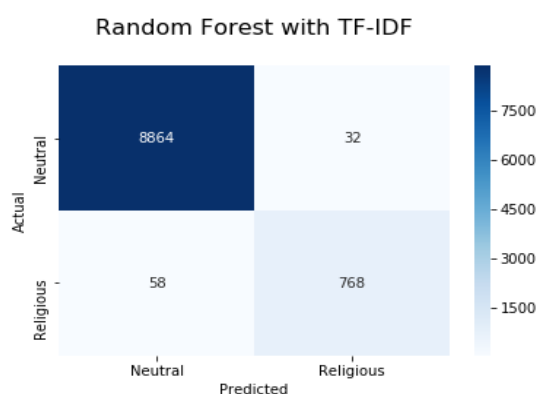
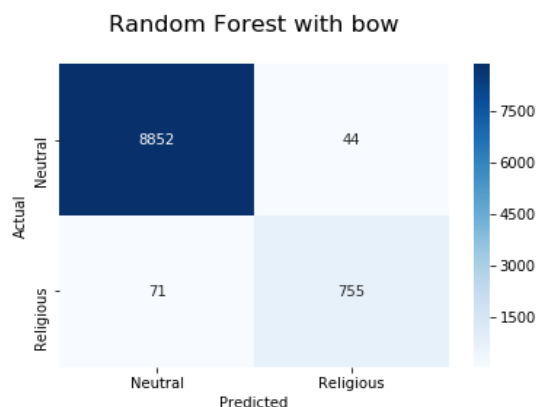
3 - сурет. SVM әдісінің нәтиже көрсеткіштері

f1_score:0.9292307692307692

	precision	recall	f1-score	support
0	0.99	1.00	0.99	8896
1	0.94	0.91	0.93	826
accuracy			0.99	9722
macro avg	0.97	0.95	0.96	9722
weighted avg	0.99	0.99	0.99	9722

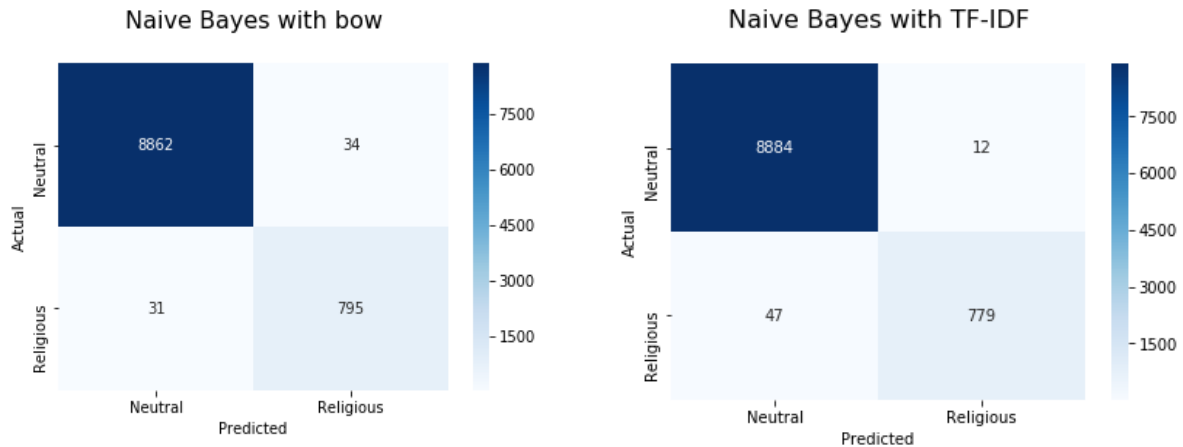
f1_score:0.944649446494465

	precision	recall	f1-score	support
0	0.99	1.00	0.99	8896
1	0.96	0.93	0.94	826
accuracy			0.99	9722
macro avg	0.98	0.96	0.97	9722
weighted avg	0.99	0.99	0.99	9722



4 - сурет. Random Forest әдісінің нәтиже көрсеткіштері

f1_score:0.960725075528701					f1_score:0.9635126777983921				
	precision	recall	f1-score	support		precision	recall	f1-score	support
0	1.00	1.00	1.00	8896	0	0.99	1.00	1.00	8896
1	0.96	0.96	0.96	826	1	0.98	0.94	0.96	826
accuracy			0.99	9722	accuracy			0.99	9722
macro avg	0.98	0.98	0.98	9722	macro avg	0.99	0.97	0.98	9722
weighted avg	0.99	0.99	0.99	9722	weighted avg	0.99	0.99	0.99	9722



5 - сурет. Naive Bayes әдісінің нәтиже көрсеткіштері

Кесте 1. Мәтіндік жіктеу әдістерінің салыстырмалы кестесі

model	f1_score with bow	f1_score with TF-IDF
Logistic Regression	0,94	0,92
SVM	0,95	0,95
Random Forest	0,93	0,94
Naïve Bayes	0,96	0,96

Қорытынды

Бұл жұмыста Интернеттегі ықтимал экстремистік және террористік ақпаратты анықтау үшін машиналық оқыту әдістерін қолданудың мәселесі қарастырылады. Сондай-ақ осы мәселе бойынша қолданыстағы шешімдер мен тәсілдерге шолу жасалды және ақпараттық шуды алып тастай отырып, жаңа өзіндік әдіс ұсынылған. Ұсынылған әдістің қолданылуы мен тиімділігі тәжірибе жүзінде көрсетілді.

Зерттеу жұмысы барысында ВКонтакте әлеуметтік желісінен парсинг жасайтын python тілінде арнайы Parser код жазылды және қазақ тіліне арналған стемминг алгоритмі құрастырылды.

Болашақта осы бағыттағы зерттеулерді жалғастырып, келесі мәселелерді шешуді жоспарлап отырмыз:

- Қазақ тілінді экстремистік корпус құрып, жоғарыда қарастырған алгоритмдер бойынша эксперимент жасау.
- Басқа әлеуметтік желілердің API жұмысымен танысу.
- Классификация дәлдігін арттыру үшін word2vec және алгоритмдерін қолдану.
- Мәтінді классификациялау мәселесі бойынша терең оқыту әдістер қарастыру.

ӘДБИЕТТЕР

- [1] Statistics. Most popular social networks worldwide. The Statistics Portal. 2020. Available online // <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/> (accessed on 2 May 2020).
- [2] Azizan, S.A.; Aziz, I.A. Terrorism Detection Based on Sentiment Analysis Using Machine Learning. J. Eng. Appl. Sci. 2017, 12, 691–698.
- [3] Yadron, D. Twitter Deletes 125,000 ISIS Accounts and expands anti-Terror Teams. Available online // <https://www.theguardian.com/technology/2016/feb/05/twitter-deletes-isis-accounts-terrorism-online>.
- [4] M. Viktor, K. Cukier. Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt, 2013.
- [5] Agarwal, S., Sureka, A.: A focused crawler for mining hate and extremism promoting videos on youtube. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media, pp. 294–296 (2014). ACM.
- [6] Sureka, A., Agarwal, S.: Learning to classify hate and extremism promoting tweets. In: Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint, pp. 320–320 (2014). IEEE
- [7] Ahmad, S., Asghar, M.Z., Alotaibi, F.M. et al. Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. Hum. Cent. Comput. Inf. Sci. 9, 24 (2019) // <https://doi.org/10.1186/s13673-019-0185-6>.
- [8] Көлеңкелі Веб және геосаяси веб-зерттеулер // <https://ai.arizona.edu/research/dark-web-geo-web>. Қаралған күні: 05.11.2017.
- [9] М.А. Болатбек, Ш.Ж. Мусиралиева ЭКСТРЕМИСТІК МӘТІНДЕРДІ МАШИНАЛЫҚ ОҚЫТУ ӘДІСТЕРІ АРҚЫЛЫ АНЫҚТАУ // ВЕСТНИК КазНУ. - 2018г.. - №6 (130). - С. 300-304.
- [10] Kaggle URL: <https://www.kaggle.com/fifthtribe/isis-religious-texts> (дата обращения: 15.02.2020).

Мусиралиева Ш.Ж., Омаров Б.С., Медетбек Ж.Б., Қараман Ғ.Р., Бекетова А.К.

Идентификация сообщений, содержащих элементы экстремизма в социальных системах с использованием методов машинного обучения

Резюме: Вместе с ростом социальных сетей растет количество религиозной ненависти и расизма в сети. Кроме того, активность радикальных групп в Интернете, призывающих к насилию и экстремизму, является одним из важнейших вопросов общественной безопасности. Потому что для таких структур основным инструментом для обмена информацией, найма и продвижения является Интернет, в частности, является веб-ресурсы, социальные сети, социальные мессенджеры и т. д. В связи с этим необходимо выявить отдельных пользователей, группы и интернет-сообщества, которые создают и распространяют террористическую и экстремистскую информацию в Интернете, а также предотвращать распространение экстремистских материалов.

Данная работа посвящена изучению и разработке методов машинного обучения, направленных на решение проблемы выявления экстремистского текста в социальных системах. Кроме того, представлены модели и методы выявления экстремистского текста, которые используются для углубленного лингвистического анализа и статистической обработки текстов. Чтобы классифицировать текст как экстремистский или неэкстремистский на основе сообщений в социальных сетях, оставленных пользователем в Интернете, мы создаем систему классификации текста с использованием методов анализа настроений на основе машинного обучения.

Ключевые слова: социальная сеть, онлайн-экстремизм, текст с радикальным содержанием, лингвистический анализ, машинное обучение, большие данные, векторная модель, логистическая регрессия, наивный Байес, глубокое обучение.

Sh.Zh. Mussiraliyeva, M.Y. Aidyn, R.K. Ospanov
(Al-Farabi Kazakh National University, Almaty, Kazakhstan
Email: {mussiraliyevash, aidynme, ospanov.ruslan.k} @gmail.com)

USER IDENTIFICATION METHOD BASED ON FRIENDSHIP AND DEMOGRAPHIC ATTRIBUTES IN SOCIAL NETWORKS

Abstract. Nowadays, social networks are a platform for a lot of information. With the exception of useful information, social media has become a convenient platform for illegal activities. Suspicious activity is often overshadowed by the lack of threat detection and analysis systems on social media. This article provides a brief overview of approaches for analyzing information from a user profile. Methods of programming interface of social network for intelligence based on open data are considered. The method of identifying the user of the profile based on the analysis of data of friendships and attributes has been tested.

Key words: social network, Vkontakte, data analysis, OSINT, virtual connections.