

А.Б. Тоқтарова^{1*}, Ж.Ж. Ажибекова², Г.Н. Казбекова¹, Ф.Е. Темірбекова³, С. Е. Алдешов⁴

¹Қожа Ахмет Ясауи атындағы Халықаралық Қазақ -Түрік Университеті, Түркістан қ., Қазақстан

²С. Ж. Асфендияров атындағы Қазақ ұлттық медицина университет, Алматы қ., Қазақстан

³М.Ауезов атындағы Оңтүстік Қазақстан университеті, Шымкент қ., Қазақстан

⁴Оңтүстік Қазақстан Мемлекеттік Университеті, Шымкент қ., Қазақстан

*e-mail: toktar.aigerim@list.ru

ҚАЗАҚ ТІЛДІ БЕЙӘДЕП СӨЗДЕР КЛАССИФИКАЦИЯСЫ ЖӘНЕ МАШИНАЛЫҚ ОҚЫТУ ӘДІСІН ОЛАРДЫ АНЫҚТАУҒА БЕЙІМДЕУ

Аңдатпа

Бұл мақалада «ғадауат» сөзінің мағынасы және оның зерттелу аймақтары қарастырылған. Сонымен қатар, қолдану және кездесу аймақтарына байланысты олардың жіктелуі көрсетілген. Яғни, ғадауат сөздер құрамына дискриминация (адам құқығын шамадан тыс бейәдеп сөздер арқылы шектеу), қорлау тілі (жағымсыз сөздер арқылы кемсіту, қорлық көрсету), кибербуллинг (интернет желісінде адамды қорлау немесе қудалау), экстремизм (саяси және діни бағыттағы шектен шыққан жағымсыз пікірмен бөлісу) және радикализм (өмір сүріп отырған саяси ортаны сөзбен сынға алу) кіретіндігі анықталып, қазақ қоғамындағы дискриминация олардың қолданылу аймақтары мен кездесу жолдары баяндалған. Мақалада онлайн контенттегі ғадауат сөздер түрлері мен адамға тигізетін әсері туралы баяндалған. Зерттеу нәтижесінде онлайн контенттен пікірлерді жинақтаушы көмегімен «Youtube» және жаңалықтар сайттарынан мәтіндік деректер жинақталған, мәтіндерді «ғадауат тілді» және «бейтарап» деп екі сыныпқа қолмен жіктеу жасалған. Сондай – ақ, модельді одан әрі оқыту үшін бейәдеп пікірлерден тұратын корпус құрылып, ғадауат тілді және бейтарап сипаттағы мәтіндердің, сөздік бұлттарды пайдалана отырып, корпусқа талдау жүргізілген.

Түйін сөздер: ғадауат сөздер, кибербуллинг, аккаунт, пікір, онлайн әлеуметтік желі, машиналық оқыту.

Аннотация

А.Б. Тоқтарова¹, Ж.Ж. Ажибекова², Г.Н. Казбекова¹, Ф.Е. Темірбекова³, С. Е. Алдешов⁴

¹Международный Казахско - Турецкий Университет им. Ходжа Ахмет Ясауи, г.Туркестан, Казахстан

²Казахский Национальный медицинский университет имени С. Д. Асфендиярова, г.Алматы, Казахстан

³Южно Казахстанский Университет имени М.Ауезова, г.Шымкент, Казахстан

⁴Южно-Казахстанский государственный педагогический университет, г.Шымкент, Казахстан

КЛАССИФИКАЦИЯ НЕЦЕНЗУРНЫХ СЛОВ КАЗАХСКОГО ЯЗЫКА И АДАПТАЦИЯ МЕТОДА МАШИННОГО ОБУЧЕНИЯ К ИХ ОБНАРУЖЕНИЮ

В статье рассматривается значение языка ненависти и области его изучения. Кроме того, показана их классификация, в зависимости от области применения и назначения. То есть в состав языка ненависти входят дискриминация (ограничение прав человека чрезмерно небрежными словами), оскорбительный язык (дискриминация отрицательными словами, унижение), киберзапугивание (оскорбление или преследование человека в сети интернет), экстремизм (обмен чрезмерно негативными мнениями политической и религиозной направленности) и радикализм (словесное оскорбление и критика политической среды, в которой он живет). Были выявлены данные классификации, зоны их применения в казахском обществе и пути встреч. В статье рассказывается о типах слов языка ненависти в онлайн-контенте и о влиянии на человека. В результате исследования были собраны текстовые данные с «Youtube» и новостных сайтов с помощью сборщика мнений из онлайн-контента, а тексты были вручную классифицированы на два класса: «нецензурные речи» и «нейтральные». Также для дальнейшего обучения модели был создан корпус, состоящий из нецензурных комментариев, и проведен анализ корпуса с использованием словесных облаков оскорбительных и нейтральных текстов.

Ключевые слова: язык ненависти, кибербуллинг, аккаунт, комментарий, онлайн социальная сеть, машинное обучение.

Abstract

CLASSIFICATION OF UNCENSORED WORDS OF THE KAZAKH LANGUAGE AND ADAPTATION OF THE MACHINE LEARNING METHOD FOR THEIR DETECTION

Toktarova A.¹, Azhibekova Zh.², Kazbekova G.¹, Temirbekova F.³, Aldeshov S.⁴

¹Khoja Akhmet Yassawi International Kazakh - Turkish University, Turkistan, Kazakhstan

²Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan

³South Kazakhstan University named after M.Auezov, Shymkent, Kazakhstan

⁴South Kazakhstan State Pedagogical University, Shymkent, Kazakhstan

This article shows the meaning of hate speech and its areas of study. In addition, their classification shown, depending on the scope and purpose. That is, hate speech includes discrimination (restriction of human rights with excessively sloppy words), offensive language (discrimination with negative words, humiliation), cyberbullying (insulting or harassing a person on the Internet), extremism (exchange of excessively negative opinions of a political and religious orientation) and radicalism (verbal abuse and criticism of the political environment in which he lives). Classification data, areas of their application in Kazakh society and ways of meeting identified. The article talks about the types of hate speech words in online content and the impact on a person. In the result of the study, text data collected from Youtube and news sites using an opinion collector from online content, and the texts were manually classified into two classes: "hate speech" and "neutral". Also, for further training of the model, a corpus consisting of obscene comments was created, and the corpus was analyzed using word clouds of offensive and neutral texts.

Keywords: hate speech, cyberbullying, account, comment, online social net, ML.

Кіріспе

Жалпы алғанда «ғадауат» сөзінің мағынасына, қолданылу аймақтарына және зерттеушілердің берген анықтамаларына тоқталдық. Ғадауат тіліне әр тұлға өзінше әртүрлі анықтама береді. Сол себепті, ең алдымен «ғадауат» (өшпенділік) және «тіл» сөздерінің мазмұнын анықтап алу қажеттілігі туындап отыр.

Журналистика мектебінің оқытушысы Оразай Қыдырбаев: «ғадауат — қандай да бір ақпаратты қабылдау негізінде адам бойында көрінетін әлдекімге жеккөрініш, адамдарға жаны ашымаушылық, тіпті әлдекімге залал келтіруді қалау сезімі» -, деп көрсетеді.

Орталық Азия елдеріндегі бітімгершілік және медиатеchnологиялар мектебінің директоры, медиасарапшы ретінде танылған Инга Сикорскаяның пайымдауынша: «Ғадауат тілі дегеніміз — медиада немесе онлайн контентте қандай да бір нәсіл, этникалық, гендерлік, діни, әлеуметтік немесе саяси топ өкілі үшін жағымсыз, сонымен қатар оны қорлайтын визуалды контент немесе жағымсыз баға берілген кез келген сөз формасы, өшпенділік риторикасы» -, деп тұжырымдайды.

Сонымен қатар, «ғадауат» сөзі күнделікті қолданыста көп кездеспегендіктен, мағынасын, шығу түбін анықтауды жөн көрдік. Ә.Т.Қайдаровтың «арабша – қазақша түсіндірме сөздігінде», яғни қазақ тіліндегі араб сөздері сөздігінің II томында: «ғадауат - 1) дұшпандық, қастық, жаулық; 2) ынтымақсыз, қырғи – қабақтық, жек көрушілік.» деген анықтама берілген. Ал, «Ғадауат» сөзі ^{арабша}-ағылшынша сөздікте «a state of deep-seated ill-will» – тереңге бойлаған қиянатшылдық күйі, тікелей мағынасын берер болсақ, «анти-махаббат» деп түсінсек болады.

Ал, ағылшын тіліне «ғадауат сөзін» тікелей аударма жасасақ, «hate speech» деген сөз тіркесі шығады, «Cambridge Dictionary» сөздігінде "public speech that expresses hate or encourages violence towards a person or group based on something such as race, religion, sex, or sexual orientation" -деп, көрсетілген. Қазақ тіліне аударма жасайтын болсақ, нәсіліне, дініне, жынысына немесе жыныстық бағдарына негізделген тұлғаға немесе топқа қатысты өшпенділік білдіретін, зорлық-зомбылық жасауға шақыру.

Онлайн контенттегі ғадауат сөздердің тигізетін әсері және түрлері

Интернеттегі әлеуметтік медиа біздің өмірімізге оң және теріс әсер ететін әлеуметтік өмірімізде өте маңызды рөл атқарады. Интернет желілері адамдарға бір-бірімен желіде сөйлесу платформаларын ұсынады, онда адамдар үйде отырып-ақ барлық жаңалықтардан хабардар болады. Адамдарға қызықты, көңілді және танымдық сәттерді көпшілік алдында бөлісу өте ыңғайлы. Дегенмен, жек көруді білдіретін немесе зорлық-зомбылықты қолдайтын, жеке адамдарға немесе ұйымдарға әртүрлі дәрежеде зиян келтіре отырып, көпшілік алдында өшпенділік көрсетіп сөйлеу жағдайлары көп. [1]

Желілік (виртуалды) байланыстың өз ерекшеліктері бар. Ерекшеліктерді коммуникацияның виртуалдылығы және оның анонимділігі, жалған атаумен желілік коммуникацияда өзін көрсету және бір уақытта бірнеше «лақап аттарды» қолдану деп түсінуге болады. Желілік (виртуалды)

коммуникацияда агрессивті коммуникациялық стратегиялар мен тактикалар, сонымен қатар ауызша агрессия құралдары өте кең қолданылады, нәтижесінде «hate speech» деп аталып жүр, яғни, қазақшалағанда «ғадауат сөздер».

Онлайн контенттегі ғадауат түсінігін және оның түбірін ескере отырып, Интернетте жиі кездесетін *холивар, троллинг, флуд және флейм* сияқты төрт агрессивті коммуникация тактикасын атап өткен жөн.

Холивар (ағылшынша Holy war – қасиетті соғыс) – шешімі жоқ мәселеге түбегейлі қарама-қайшы көзқарастағы адамдар арасындағы дау. Мысалы, қайсысы дәмдірек: шабдалы немесе шабдалы шырыны.

Троллинг - интернет-коммуникациядағы әлеуметтік арандатудың немесе қорлаудың бір түрі, оны өзін көбірек танытуға, жариялауға, шектен шығуға мүдделі атын жасырмайтын жеке қатысушылар да, сонымен, қатар анықтау мүмкіндігінсіз анонимді пайдаланушылар да пайдаланады.

Флуд (ағылшынша Flood – су тасқыны) – пайдаланушылар талқылау тақырыбына қатысы жоқ мазмұны бойынша мағынасыз хабарламалармен алмасатын Интернетте әңгіме жүргізу тәсілі. Әдетте, флуд форумдар мен чаттарда назар аудару немесе тролль үшін қолданылады. Мұндай хабарламаларды жіберетін пайдаланушылар флудерлер деп аталады.

Flame (ағылшын тілінен Flame – алау, от, жалын) «дауласу үшін аргумент», көп қолданушы желілік байланыс орындарында хабар алмасу (мысалы, интернет-форумдарда, чат бөлмелерінде, әлеуметтік желілерде және т.б.), бұл ауызша соғыс, көбінесе даудың бастапқы себебімен байланысы жоқ. [2]

Вербальды (сөз арқылы қарым - қатынас) агрессияны зерттеуші Ю.В. Щербинина, «жалпы мағынада «жек көрушілік» - бұл адамдарды агрессияға бейсаналық немесе айқын мағыналайтын сөздер мен өрнектер және оның іске қосу механизмі». Автор «ғадауат сөзінің» «басқа», әдеттен тыс, адамдардың белгілі бір әлеуметтік немесе мәдени қауымдастықтың басым стереотиптерінен ерекшеленетін барлық нәрсеге деген өшпенділік сезімін білдіруге қызмет ететініне ерекше назар аударады. «Ғадауат сөзінің» қолданылу аясы өте кең екенін атап өткен жөн: ол әртүрлі тәжірибелерде қолданылады – ресми, кәсіби, күнделікті өмірде де кездестіре аламыз. Саясаткерлердің сөзінде де, зейнеткерлердің әңгімесінде де, студенттердің, мектеп оқушыларының, ерлі-зайыптылардың әңгімесінде де байқай аламыз.

Зерттеу әдіснамасы. Ғадауат сөздер классификациясы.

Онлайн медиа пайдалануды зерттеушілерінің пікірінше, «ғадауат сөзінің» пайда болуына және таралуына қоғамның экономикалық, әлеуметтік және саяси өміріндегі дағдарыс әсер етуі мүмкін. Бұл виртуалды кеңістікте, әлеуметтік желілерде, интернетте бірден көрінеді. Офлайн кеңістіктегі "жек көрушілік" тез желіге ауысады. Жүргізілген бақылау мен зерттеу нәтижелері интернет кеңістігінде «ғадауат сөздерінің» көбірек екенін растайды. Мысалы, Medianet халықаралық журналистика орталығы жүргізген «Қазақстандағы әлеуметтік желілердегі бейәдеп сөздерінің мониторингі» зерттеуіне сәйкес, Facebook желісінің қазақтілді сегментінде «қатаң» этникалық немесе діни кемсітушілік сөздерінің салыстырмалы түрде жоғары таралуы байқалады. [3]

Ғадауат пікірлерді өршітудің мақсаты – дискурсивті кеңістікті арттыру, белгілі бір топқа өшпенділік тудыру, қоғамда агрессивті көңіл-күй тудыруға ықпал ету. Арандатушылық, кемсітушілік, экстремистік сипаттағы жеке хабарламалардың белгілі бір сипаттағы бұқаралық коммуникативті әрекеттерге айналғанын байқауға болады.

Қазіргі таңда Media Marketing Index статикасы бойынша ұялы телефонмен Интернет қолданушылар статистикасы көрсетілген. 2021 жылға Қазақстанның 100 мыңнан астам халқы бар қалаларында тұратын 15 жастан асқан интернет пайдаланушылардың шамамен 60%-ы Интернетке кіру үшін тек мобильді құрылғыларды пайдаланады. 2020 жылмен салыстырғанда бұл көрсеткіш 7,3%-ға, ал 2018 жылмен салыстырғанда 30%-дан астамға өсті. Сонымен қатар, тек жұмыс үстелі интернетін пайдаланушылардың пайызы төмендеуде – соңғы 4 жылда олардың саны 4,9%-дан (2018 ж.) 1,1%-ға (2021) дейін азайды. 2021 жылға арналған Интернетке кіру үшін мобильді және жұмыс үстелі құрылғыларын ортақ пайдалану 37% жетеді [4].

2022 жылдың ақпан-сәуір айларын қоса алғанда, «Қолжетімділігі бойынша үздік 20 мобильді веб-сайт» рейтингінің бестігіне Google* (осы холдингке қатысты барлық жобалар) – 77%, Яндекс* (осыған қатысты барлық жобалар) сияқты ірі холдингтер кіреді. холдинг) - 56,4% және VK (Mail.Ru Group) * (осы холдингке қатысты барлық жобалар, соның ішінде ВКонтакте, Одноклассники және MoiMir әлеуметтік желілері) - 54,6%; Сондай-ақ, бұл «Facebook» әлеуметтік желісі (63%) және «Wikipedia.org»

электронды энциклопедиясы (48%). Қазақстандық ресурстарға 1xbet.kz букмекерлік кеңсесі, жергілікті базарлар, электронды үкімет веб-сайттары және ірі жаңалықтар порталдары кіреді.

Сонымен қатар, әлеуметтік желілерді пайдалану деңгейі өскен сайын онлайн платформадағы әлімжеттік көрсету деңгейі де өсе бастағанын байқауымызға болады. Өкінішке қарай, Қазақстан желі қолданушыларынан кибербуллинг бойынша шағымдар статикасы жүргізілмеген. Дегенмен, 2022 жылы наурыз айында құрылған BalaQorgau оқушылар мен ата – аналарына көмек жобасына 600 ден астам өтінім келіп түскен, оның ішінде 100 – ге жуығы мектеп оқушыларына әлеуметтік желілерде кибербуллинг жасалған. Яғни, бала өз парақшасына жүктеген видеолар мен суреттер астына кемсіту пікірлері жазылғанын алға тартса, кейбір шағымдарда баланың фотосын фотошоппен жағымсыз етіп өңдеп, мазақ етіп әлеуметтік желілерге таратқан [5].

Сонымен, ғадауат сөздер құрамына *дискриминация қорлау тілі, кибербуллинг, экстремизм және радикализм* (өмір сүріп отырған саяси ортаны сөзбен сынға алу) кіретінін көруге болады.

Дискриминация - (латын тіліне «*discriminatio*» - *оқшаулау, алалау*) әлдебір белгілерге байланысты адамдарды бөле-жара қарау, алалау немесе оған негативті қарым-қатынас жасап, құқығынан айыру және *бостандығын шектеу*.

Қазақ қоғамындағы өзіне тән дискриминация бар деп айтуға болады, яғни, ру сұрасу. Ру сұраған адам өзіне ұнамайтын ру өкілін алшақтауға, кемсітуге қандай да бір шектеу көрсетуге тырысып жататын кездерін байқауға болады. Сонымен қатар, жерлес адамдардың басқа елді мекеннің адамдар қатыранан бөлініп шығу қасиеті бар екенін ескере кетуіміз қажет.

Кибербуллинг – (ағылшын тілінен аударғанда «*bullying*» - қысым көрсету, қорлау және қысым көрсету) әлеуметтік желі мессенджерлерін пайдалану арқылы сөзбен қорқыту және үркіту кей жағдайда қысым көрсетуге дейін бару. Кибербуллинг адамның жасын, ұлтын, нәсілін, әлеуметтік жағдайын таңдамайтындығы белгілі болып отыр. Мысалы, интернет желілерді біреуді қасақана ренжіту, қорлау немесе ызасын келтіру үшін пайдалану, бұл – кибербуллинг. Кейде, біреу туралы жағымсыз постқа лайк басу немесе ренішті постты жіберу, бұл да – кибербуллинг [6].

The Oxford Advanced Learner's Dictionary, Merriam-Webster Dictionary, Collins Dictionary сөздіктерін пайдалану арқылы кибербуллингке қандай сөздер топтамасы енетінін анықтадық:

1. *Тыйым салынған сөздер (табу) және ұятсыз тіл* (бір және бірнеше сөздер құрамынан тұрады) - жәбірленушінің әлеуметтік беделін түсіру, оның кемшілігін көрсету арқылы оны манипуляциялау мақсатында агрессия құралы ретінде қолданылады. Бұл сөздіктің құрамында анық түрде балағаттаушы сөздер кездеседі. Мысалы, м*лғұн, т*пас, б*йбақ, м*нғұрт және ж*леп сияқты сөздер қатары.

2. *Адамның интимдік өміріндегі* сөздер – жәбірленушіні ренжіту, оны қаралау және қорлау үшін жиі қолданады. Көптеген тілдерде жыныстық органдардың атаулары мен жыныстық әрекеттердің түрлері тыйым салынған лексикаға жатады сондықтан, көбінесе адресаттың агрессивтілігін білдіру үшін қолданылады. қ*тақжегіш, *мжалақ, қ*тақбас, с*гіп *лтіру сияқты сөздер тізбегі.

3. *Жыныстық ориентацияға және сексизмге қатысты ұғымдарға сілтеме жасайтын сөздер* - көбінесе «құрбандарға» гей, лесбиян, транс, педофил/педо, сексистік сөздерді қолдана отырып, олардың дәстүрлі емес сексуалдық бағдары айтылады.

4. *Зұлымдық пен өлім тілеу* – мұндай лексикалық сөз тіректерінің қолданылуы буллардың жәбірленушіні жақтырмауынан, тіпті жәбірленушіні жек көруінен туындайды. Мысалы, *сылып өліп қал, т*нысың бітіп қалсын, *лгенің артық, *лсейші, *ліп қалшы, сені жер қалай көтереді және т.б. сияқты лексикалық тіркестерінің қолданылуы буллардың жәбірленушіден өзінің артықшылығын атап көрсеткісі келетінін, ұлылығын дәлелдегісі келетінін көрсетеді;

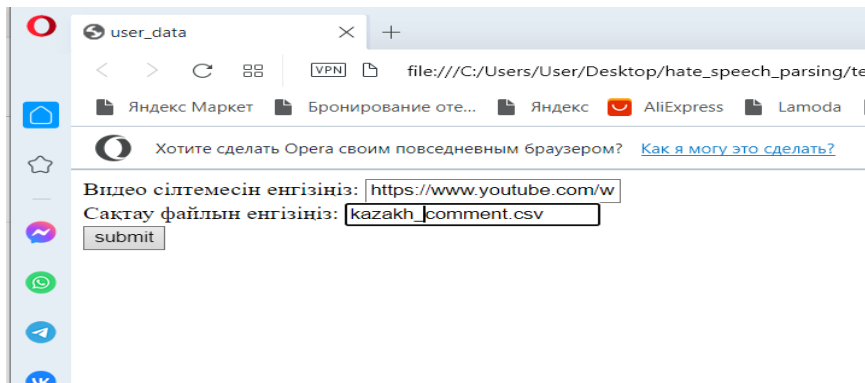
5. *Адамды қорлауды және қорлауды білдіретін сөздер* – «бірнеше баскесер – бір құрбан» схемасы бойынша электронды қудалау орын алған кезде жиі қолданылатын сөздер. Жәбірленушінің өзін-өзі бағалауын төмендетуге бағытталған дөрекі сөздер. Мысалы, сұм, сүмелек, надан, топас, ақымақ, т.б. лексикалық сөз тіркестер. Осылайша, Буллар жәбірленушіден күшті екенін көрсетуге тырысады.

6. *Ұлтқа және нәсілшілдікке қатысты ұғымдарды атайтын сөздер*, мысалы, негр, өзбек, сарт, ит қытай және т.б. - агрессорға жәбірленушіге моральдық ғана емес, сонымен қатар әлеуметтік зиян келтіруі мүмкін;

7. *Жануарлар атымен атайтын сөздер* – қорлау ретінде қолданылады, мұндай лексикалық тіркестер дәстүрлі балағат сөздермен қатар инвективті сөздер санатына өтеді. Спыр, есек, шошқа, қой, құрт, жылан т.б. сияқты жануарларды салыстыру арқылы жәбірленушінің қасиеттерін жануарлардың қасиеттерімен теңестіреді, бұл оның қоғамдағы мәртебесін төмендетеді.

8. *Физикалық және ақыл-ой кемістігі бар адамдарды кемсітетін сөздер қатары* – бұл сөздік қорды қолданудың негізгі мақсаты жәбірленушінің ар-намысы мен қадір-қасиетін қорлау, оның әлеуметтік беделін төмендету, тұлға ретінде жою. Буллер жәбірленуші өзінен ақылдырақ болғанда жәбірленушінің шектеулі ақыл-ой қабілетін көрсететін тілді жиі қолданады. Бұл топқа керек, мүгедек, топас, ақымақ, мәңгүрт, әлсіз әлжуаз, т.б. сөздер тізбегі кіреді [7-9].

Жоғарыда көрсетілген 8 классты ескере отырып, қазақ тілді әлеуметтік желілердегі пікірлерден тұратын мәліметтер қорын жинақтауда Python бағдарламалау тілінде жазылған парсерді қолдандық Сурет 1 - де көрсетілген.



Сурет 1. Youtube желісі үшін пікірлерді жинақтау (парсинг)

Пікірлерді жинақтау (парсинг) интерфейсі ұсынылып отыр, яғни онлайн түрде қажетті файлдың сілтемесін енгізу арқылы, екінші жолаққа сақтау файлы құрау арқылы видеоның астындағы пікірлерді, аз уақыт ішінде өзімізге жүктеп алуға болатындығын байқадық. Келесі қадам, жинақталған пікірлерді екілік классқа бөлу, яғни ғадауат тілді пікірлерді «1» атрибутымен, ал бейтарап пікірлерді «0» деп Сурет 2 – де көрсетілген бойынша нөмерледік.

	A	B
4493	Кайыршы шенеуликтер. Сұмдықкой , көздері мен колдары топыракка толсын.	1
4494	Спортшыларға бір жылға жуық айлықтарын бермеді 2022 жылы не сұмдық құрсын сонда ұлттық банктен бөлінген қаржыны мына жалмауыздар жеген екен ғой желке лерінен тесіп шықырлар...масқара ғой.	1
4495	Жемқорлықтан бас алмаған мүсәпірлер!	1
4496	Нартай осындай өтіркіші қотақбастарды шығармашы қырғыздар сөзбен сігіп жатыр ана жақта	1
4497	Котин айратын Адам екенгой Бекболат деген Адам екенгой	1
4498	Самрық қазына қорының ақшасын Назарбаевтардың бауырлары әбден белден басып келген екен ғой ұрылар.	1
4499	Әбдрахманов қотақбас шсгн	1
4500	Солай жағымпаз болып енді билік Тен Нұрсұлтан кеткенсоң екі жүзділік пен жамандап ал енді ақ болып отырсың Ар ұяттан жұрдай болып кетіпсің	1
4501	Зорға шыдап отырм отиригине еи деп солемек туры тилин котине тыгып отрасын ⁹	1
4502	Ешакбас Серик екенсин	1
4503	Серік Абдрахманов кандай екі жүзді адамсыз Назарбевтың қасынд жүргенде мақтап қолдап жүрдің Сол кезде халықтың сөзің сөйлеген халық азаматтар қайда кетті кілен халықпен санаспайтын жемқорлар қалған	1
4504	Назарбаев деген кім өзі қазақты біріне бірін айдап салып бітті ғой сол генералдардың бәрі ұрғашығой еркее бодса қайда елге жаны ашығандығы өтіріктің арты бір тұтам еркек болса қайда елдің алдына шығып жауап бермеидіма ұрғашы	1
4505	Енді не деп отырсың, құйыршық, тиын-тебенге жолатбай қойған соң сатып отырсың да .	1
4506	Назарбевтің тупойлығы ұры екендігі сонша халықтың алдына шыға алмағаны.жәнеде Астананың жақсы атын Нұрсұлтан деп ерсі қойғаны тупойлығы	1
4507	Мына шал не деп кыртып отыр бұрында айтпай ма пиздес халыкка жаны ашыса енди келе калып жамандап отыр котилаш	1
4508	Ұры назарбаев ты бари билет неге туқымымен камамайды токаев	1
4509		1

Сурет 2. Қазақ тілді ғадауат тілді пікірлер мәліметтер қоры

Дайындалған жинақтаушы көмегімен «Youtube» және жаңалықтар сайттарынан мәтіндік деректер жинақталып, мәтіндерді «ғадауат тілді» және «бейтарап» деп екі сыныпқа қолмен жіктеу жасалды. Келесі қадам, модельді одан әрі оқыту үшін корпус құрып, ғадауатт тілді және бейтарап сипаттағы мәтіндердің, сөздердің бұлттарының таралуын пайдалана отырып, корпусқа талдау жүргізу болып табылады.

Нәтижені талқылау

Қазақ тіліндегі мәтіндерді өңдеу кезінде машиналық оқытудың басқа тілдерінен біршама айырмашылығы бар. Әрине, тілдің өзіндік ерекшеліктерімен қатар заңдылықтары да аз емес. Қазақ тілінде негізгі мәселелердің бірі – негізгі түбір сөзді табу, сонымен қатар жұрнақтар мен жалғауларды сөздің түбірі мен мағынасын сақтай отырып алып тастау.

Тоқтау сөздер (stop words) – мәтінде ерекше рөл атқармайтын, бірақ жалпы мәтіннің мағынасын жақсы толықтыратын сөздер. [10]Бұл жердегі мәселе әртүрлі тақырыптық мәтіндерде және ауызекі сөйлеуде тым жиі кездеседі. Олардың машиналық оқытуда болуы объектіні оқытуда және жіктеуде біршама қателерді тудыратын үлкен мәселе тудырады. Бұл тоқтау сөздерінің көпшілігі *Сурет 3 - те* көрсетілгендей nltk кітапханасынан алынған. Қазақ тілінде экстремизм тақырыбына қатысты 300 сөз, сонымен қатар өз сөздері бар. Нәтижесінде олар осындай тақырыптық мәтіндерді жақсырақ анықтауға көмектесті.

```
@author: User
import nltk
from nltk.corpus import stopwords
print(stopwords.words('kazakh'))
```

Сурет 3. Nltk кітапханасындағы қазақша тоқтау сөздерді шақыру

Nltk кітапханасындағы қазақша тоқтау сөздерді шақырту, яғни мәтінде ерекше рөл атқармайтын, бірақ жалпы мәтіннің мағынасын жақсы толықтыратын *Сурет 4- те* көрсетілген сөздерді жинақталған деректер қорынан алып тастау арқылы, машиналық оқыту әдістерін пайдалана отырып жоғары дәлдікке жетуге болады.

```
In [8]: runfile('C:/Users/User/anaconda3/lib/site-packages/pandas/io/untitled15.py', wdir='C:/Users/User/anaconda3/lib/site-packages/pandas/io')
['ах', 'ох', 'эх', 'ай', 'эй', 'ой', 'тағы', 'тағыда', 'әрине', 'жоқ', 'сондай', 'осындай', 'осылай', 'солай', 'мұндай', 'бұндай', 'мен', 'сен', 'ол', 'біз', 'біздер', 'олар', 'сіз', 'сіздер', 'маған', 'оған', 'саған', 'біздің', 'сіздің', 'оның', 'бізге', 'сізге', 'оларға', 'біздерге', 'сіздерге', 'оларға', 'менімен', 'сенімен', 'онымен', 'бізбен', 'сізбен', 'олармен', 'біздермен', 'сіздермен', 'менің', 'сенің', 'біздің', 'сіздің', 'оның', 'біздердің', 'сіздердің', 'олардың', 'маған', 'саған', 'оған', 'менен', 'сенен', 'одан', 'бізден', 'сізден', 'олардан', 'біздерден', 'сіздерден', 'олардан', 'айтпақшы', 'сонымен', 'сондықтан', 'бұл', 'осы', 'сол', 'анау', 'мынау', 'сонау', 'осынау', 'ана', 'мына', 'сона', 'әні', 'міне', 'ей', 'үйт', 'бүйт', 'біреу', 'кейбіреу', 'кейбір', 'қайсыбір', 'әрбір', 'бірнеше', 'бірдеме', 'бірнеше', 'әркім', 'әрне', 'әрқайсы', 'әрқалай', 'әлдекім', 'әлдене', 'әлдеқайдан', 'әлденеше', 'әлдеқалай', 'әлдеқашан', 'әлдақашан', 'әш', 'әшкім', 'әшбір', 'әштеме', 'дәнене', 'әшқашан', 'әшқандай', 'әшқайсы', 'емес', 'бәрі', 'барлық', 'барша', 'бар', 'күллі', 'бүкіл', 'түгел', 'өз', 'өзім', 'өзің', 'өзінің', 'өзіме', 'өзіне', 'өзімнің', 'өзі', 'өзге', 'менде', 'сенде', 'онда', 'менен', 'сенен', 'тонан', 'одан', 'ау', 'па', 'ей', 'эй', 'е', 'уа', 'уау', 'уай', 'я', 'пай', 'ә', 'о', 'оно', 'ой', 'ие', 'ана', 'ау', 'беу', 'массаған', 'бәрекедлі', 'әттегенай', 'жаракімалла', 'масқарай', 'астапыралла', 'япырмай', 'ойпырмай', 'кәне', 'кәнеки', 'ал', 'айда', 'кәні', 'міне', 'әні', 'сорап', 'қош-қош', 'пйша', 'піше', 'құрау-құрау', 'шәйт', 'шек', 'моң', 'төк', 'құрау', 'құр', 'кә', 'кәң', 'күшім', 'күшім', 'мышы', 'пырс', 'әүкім', 'алақай', 'паң-паң', 'бәрекедлі', 'ура', 'әттен', 'әттеген-ай', 'қап', 'түге', 'пішту', 'шіркін', 'алатау', 'пай-пай', 'үшін', 'сайын', 'сияқты', 'туралы', 'арқылы', 'бойы', 'бойымен', 'шамалы', 'шақты', 'қаралы', 'ғұрлы', 'ғұрлым', 'шейін', 'дейін', 'қарай', 'шаман', 'салым', 'тарта', 'жуық', 'таяу', 'гері', 'бері', 'кейін', 'соң', 'бұрын', 'бетер', 'қатар', 'бірге', 'қоса', 'арс', 'ғұрс', 'дүрс', 'қорс', 'тарс', 'тырс', 'ырс', 'барқ', 'борт', 'күрт', 'кірт', 'морт', 'сарт', 'шырт', 'дүңк', 'күңк', 'қыңк', 'мыңк', 'маңк', 'саңк', 'шаңк', 'шіңк', 'тыңк', 'тыңқ', 'ыңк', 'болл', 'былл', 'жалл', 'желл', 'қолл', 'ірк', 'ырқ', 'сарт-сұрт', 'тарс-
```

Сурет 4. Nltk кітапханасындағы қазақша тоқтау сөздер

Стемминг - қажетсіз жұрнақтар мен жалғауларды алып тастап, соңында түбірді табады. Бұл бөлім меншікті алгоритмдермен орындалды. Онда алгоритм екі нұсқа негізінде жұмыс істейді. Біріншісі – дерекқордан бірегей сөздерді іздеп, мәтінде оларды ауыстыру. Базада ондай сөздер болмаса, қазақ тілінің ережелері бойынша орындалады. Stemming Python тілінде жазылған.

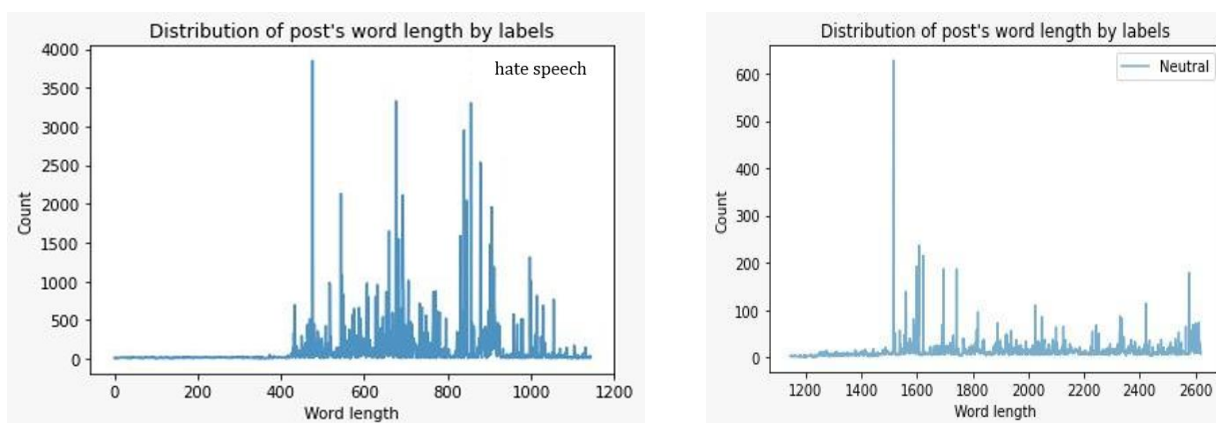
Statistics - бұл сөздің немесе таңбаның жіктеу тапсырмаларында мәтіндерді оқытуға қаншалықты жақсы әсер ететіндігімен анықталатын *Сурет 5 – те* көрсетілген мәтіндік статистика. [11-12]

$$feature_1 = \frac{\text{unique words}}{\text{word counts}}$$

$$feature_2 = \text{word counts}$$

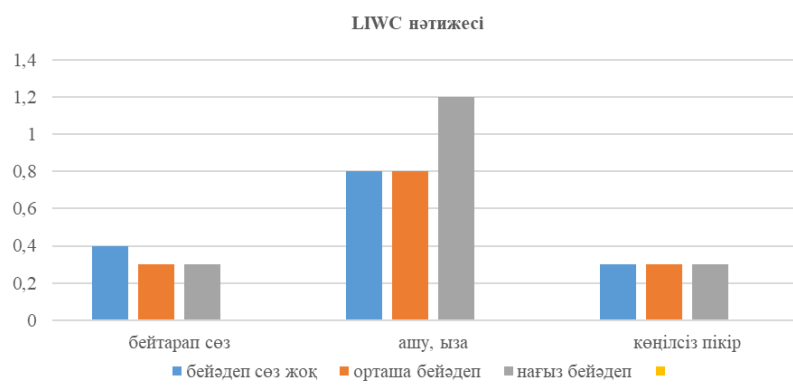
$$feature_3 = \text{symbol counts}$$

$$feature_{alphabet} = \sum_{i=0}^{n=41} x_i \text{ in text}$$



Сурет 5. Бейәдеп және бейтарап тілді мәтіндік статистика.

LIWC (*Linguistic Query and Word Count*) – мәтінді талдау бағдарламасы. Ол мәтінде *Сурет 6 – да* көрсетілгендей сөздердің әртүрлі санаттарының қаншалықты қолданыста екендігін есептейді.



Сурет 6. Мәтінді талдау бағдарламасы

Сөздік бұлт. Деректерді көрнекі түрде көрсету үшін сөздік бұлттар пайдаланылды. Кемсіту ықтималы немесе өшпенділік идеялары бар пайдаланушы жазбалары *Сурет 7 –де* бөлек көрсетілген. Көріп отырғанымыздай, ғадауат жазбалар пайдаланушылардың кемсітіп, қорлау немесе қорқыту мақсатында екенін тікелей көрсететін «ж.леп», «м.қтабас құл», «тып тастау керек» сияқты сөздерді жиі пайдаланады.



Сурет 7. Гадауат сөздер және қалыпты сөздер сөздік бұлтты

Жинақталған декрекетер қорының жалпы мағыналы көрінісін көру үшін сөздік бұлттар пайдаланылды. Гадауат тілді немесе өшпенділік бағытты қамтитын сөздер тіркестері бар пайдаланушылар пікірлері көрсетілген.

Қорытынды

Қорытындылай келе, интернетті пайдаланушылар саны күн сайын артуы адамды ақпаратпен жылдам қамтамасыз етумен қатар, онлайн контенттен келетін «қауіп» түрі де артуда. Физикалық буллинг түрі белгілі бір уақытта тоқтауы мүмкін, ал кибербуллинг көрсетуге 24/7 мүмкіндік бар екенін ескере отырып, оның түрлеріне қандай жағыдайда онлайн контенттегі хабарламалар арқылы қауіп келетіндігін атап өттік.

Python бағдарламалау тілінде құрастырылған жинақтаушы (парсинг) көмегімен «Youtube» және жаңалықтар сайттарынан мәтіндік деректер жинақталып, мәтіндерді «гадауат тілді» және «бейтарап» деп екі сыныпқа қолмен жіктеу жасалды. Сонымен қатар, модельді одан әрі оқыту үшін корпус құрып, гадауат тілді және бейтарап сипаттағы мәтіндерден тоқтау сөздерді алып тастау арқылы, стеминг жасау арқылы деректер қорын «тазалау» жұмыстары жүргізілді, сөздік бұлттарының таралуын пайдалана отырып, корпуска талдау жүргізілді.

Пайдаланылған әдебиеттер тізімі:

- 1 R. Alshalan and H. Al-Khalifa, "A deep learning approach for automatic hate speech detection in the Saudi Twittersphere", *Appl. Sci.*, vol. 10, no. 23, pp. 1-16, 2020.
- 2 A. Al-Hassan and H. Al-Dossari, "Detection of hate speech in social networks: A survey on multilingual corpus", *Proc. Comput. Sci. Inf. Technol. (CS IT)*, pp. 83-100, Feb. 2019.
- 3 A. Schmidt and M. Wiegand, "A survey on hate speech detection using natural language processing", *Proc. 5th Int. Workshop Natural Lang. Process. Social Media*, pp. 1-10, 2017.
- 4 TNS Central Asia | *Media Marketing Index (kantar.kz)*
- 5 UNICEF Kazakhstan
- 6 A. Alrehili, "Automatic hate speech detection on social media: A brief survey", *Proc. IEEE/ACS 16th Int. Conf. Comput. Syst. Appl. (AICCSA)*, pp. 1-6, Nov. 2019.
- 7 M. A. Al-Garadi, M. R. Hussain, N. Khan, G. Murtaza, H. F. Nweke, I. Ali, et al., "Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges", *IEEE Access*, vol. 7, pp. 70701-70718, 2019.
- 8 A. Rodriguez, C. Argueta and Y.-L. Chen, "Automatic detection of hate speech on Facebook using sentiment and emotion analysis", *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIC)*, pp. 169-174, Feb. 2019.
- 9 G. Weir, K. Owoeye, A. Oberacker and H. Alshahrani, "Cloud-based textual analysis as a basis for document classification", *Proc. Int. Conf. High Perform. Comput. Simul. (HPCS)*, pp. 629-633, Jul. 2018.
- 10 J. Cheng, C. Danescu-Niculescu-Mizil and J. Leskovec, "Antisocial behavior in online discussion communities", *Proc. 9th Int. Conf. Web Soc. Media (ICWSM)*, pp. 61-70, 2015, 2015.
- 11 Alkomah, F., & Ma, X. (2022). A Literature Review of Textual Hate Speech Detection Methods and Datasets. *Information*, 13(6), 273.
- 12 Mühlhoff, R. (2020). Human-aided artificial intelligence: Or, how to run large computations in human brains? *Toward a media sociology of machine learning. new media & society*, 22(10), 1868-1884.

References:

- 1 R. Alshalan and H. Al-Khalifa, "A deep learning approach for automatic hate speech detection in the Saudi Twittersphere", *Appl. Sci.*, vol. 10, no. 23, pp. 1-16, 2020.
- 2 A. Al-Hassan and H. Al-Dossari, "Detection of hate speech in social networks: A survey on multilingual corpus", *Proc. Comput. Sci. Inf. Technol. (CS IT)*, pp. 83-100, Feb. 2019.
- 3 A. Schmidt and M. Wiegand, "A survey on hate speech detection using natural language processing", *Proc. 5th Int. Workshop Natural Lang. Process. Social Media*, pp. 1-10, 2017.
- 4 TNS Central Asia | Media Marketing Index (kantar.kz)
- 5 UNICEF Kazakhstan
- 6 A. Alrehili, "Automatic hate speech detection on social media: A brief survey", *Proc. IEEE/ACS 16th Int. Conf. Comput. Syst. Appl. (AICCSA)*, pp. 1-6, Nov. 2019.
- 7 M. A. Al-Garadi, M. R. Hussain, N. Khan, G. Murtaza, H. F. Nweke, I. Ali, et al., "Predicting cyberbullying on social media in the big data era using machine learning algorithms: Review of literature and open challenges", *IEEE Access*, vol. 7, pp. 70701-70718, 2019.
- 8 A. Rodriguez, C. Argueta and Y.-L. Chen, "Automatic detection of hate speech on Facebook using sentiment and emotion analysis", *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIIC)*, pp. 169-174, Feb. 2019.
- 9 G. Weir, K. Owoeye, A. Oberacker and H. Alshahrani, "Cloud-based textual analysis as a basis for document classification", *Proc. Int. Conf. High Perform. Comput. Simul. (HPCS)*, pp. 629-633, Jul. 2018.
- 10 J. Cheng, C. Danescu-Niculescu-Mizil and J. Leskovec, "Antisocial behavior in online discussion communities", *Proc. 9th Int. Conf. Web Soc. Media (ICWSM)*, pp. 61-70, 2015, 2015.
- 11 Alkomah, F., & Ma, X. (2022). A Literature Review of Textual Hate Speech Detection Methods and Datasets. *Information*, 13(6), 273.
- 12 Mühlhoff, R. (2020). *Human-aided artificial intelligence: Or, how to run large computations in human brains? Toward a media sociology of machine learning.* *new media & society*, 22(10), 1868-1884.