

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ

«Л.Н. ГУМИЛЕВ АТЫНДАҒЫ ЕУРАЗИЯ ҰЛТТЫҚ УНИВЕРСИТЕТІ» КЕАҚ

**Студенттер мен жас ғалымдардың
«GYLYM JÁNE BILIM - 2023»
XVIII Халықаралық ғылыми конференциясының
БАЯНДАМАЛАР ЖИНАҒЫ**

**СБОРНИК МАТЕРИАЛОВ
XVIII Международной научной конференции
студентов и молодых ученых
«GYLYM JÁNE BILIM - 2023»**

**PROCEEDINGS
of the XVIII International Scientific Conference
for students and young scholars
«GYLYM JÁNE BILIM - 2023»**

**2023
Астана**

УДК 001+37
ББК 72+74
G99

**«GYLYM JÁNE BILIM – 2023» студенттер мен жас ғалымдардың
XVIII Халықаралық ғылыми конференциясы = XVIII
Международная научная конференция студентов и молодых
ученых «GYLYM JÁNE BILIM – 2023» = The XVIII International
Scientific Conference for students and young scholars «GYLYM JÁNE
BILIM – 2023». – Астана: – 6865 б. - қазақша, орысша, ағылшынша.**

ISBN 978-601-337-871-8

Жинаққа студенттердің, магистранттардың, докторанттардың және жас ғалымдардың жаратылыстану-техникалық және гуманитарлық ғылымдардың өзекті мәселелері бойынша баяндамалары енгізілген.

The proceedings are the papers of students, undergraduates, doctoral students and young researchers on topical issues of natural and technical sciences and humanities.

В сборник вошли доклады студентов, магистрантов, докторантов и молодых ученых по актуальным вопросам естественно-технических и гуманитарных наук.

УДК 001+37
ББК 72+74

ISBN 978-601-337-871-8

**©Л.Н. Гумилев атындағы Еуразия
ұлттық университеті, 2023**

**МАШИНАЛЫҚ ОҚЫТУҒА НЕГІЗДЕЛГЕН МӘТІНДІКЛАССИФИКАЦИЯЛАУ
АЛГОРИТМДЕРІН ТАЛДАУ**Рахимова Дана Акылбековна nadara022414@gmail.comҚазақстан, Астана, Л.Н.Гумилев атындағы ЕҰУ «Ақпараттық жүйелер» мамандығының 2 курс
магистранты

Ғылыми жетекші – Тусупов Д.А.

Абстракт

Мәтінді классификациялау – бұл табиғи тілді өңдеудің ең маңызды саласы, онда мәтіндік деректер алдын-ала анықталған кластар жиынтығы бойынша автоматты түрде сұрыпталады. Мәтінді жіктеу спамды сүзу, шешім қабылдау, шикі деректерден ақпараталу және басқа да көптеген коммерциялық жұмыстарда кеңінен қолданылады. Мәтінді жіктеу көптеген кәсіпорындар үшін маңызды, өйткені ол деректерді қолмен жіктеу қажеттілігінен босатады, бұл қаражатты және уақытты қажет ететін механизм. Бұл мақалада мәтіндердің жіктелуіне салыстырмалы талдау жасайды, оның барысында әртүрлі мәліметтер жиынтығындағы машиналық оқытудың әртүрлі алгоритмдерінің тиімділігі талданады және салыстырылады. Тірек векторлық машина (SVM), k-ең жақын көрші (k- NN), логистикалық регрессия (LR), көпмүшелік аңғал Байес (MNB) және кездейсоқ орман (RF) – бұл қазіргі кезде кеңінен қолданылатын машиналық оқытуға негізделген алгоритмдер. Осы алгоритмдерді салыстырмалы талдау үшін екі түрлі мәліметтер жиынтығы қолданылады. Бұл мақалада өнімділік көрсеткіштеріне, атап айтқанда дәлдікке, еске түсіруге және F1 бағалауға негізделген мәтінді жіктеу үшін қолданылатын машиналық оқыту әдістері талданады. Нәтижелері логистикалық регрессия мен тірек векторлық машинаның IMDb деректер жиынындағы басқа модельдерден, ал kNN ұсынылған жүйе арқылы алынған нәтижелерге сәйкес E-mail хабарламалар деректер жиынының басқа үлгілерінен асып түсетінін көрсетеді.

1. Кіріспе

Қазіргі уақытта өнеркәсіп салалары құрылымдалмаған мәтін көздерінен пайдалы құрылымдық деректерді алу үшін автоматты жүйелерді әзірлеуден үлкен үлес көруге болады. Зерттеушілер мен сала мамандары құрылымдық ресурсты пайдалана отырып, өндірістік жұмыстарға қатысты барлық ақпаратты алу үшін жеткілікті қарапайым сұраныстарды орындайтын еді. Біз бұл машиналық оқыту классификаторларын қоршаған орта саласында пайдалана аламыз. Тұрақтылық пен климаттың өзгеруіне қатысты деректер әртүрлі көздерден жиналады делік. Бұл жағдайда бұл деректерге пәндік сала туралы білімді алу үшін машиналық оқытудың әртүрлі әдістерін қолдануға болады [1]. Бұл бізге алдағы уақытта шешім қабылдаудың әртүрлі салаларында көмектеседі, сонымен қатар қолда бар ресурстарды қалай ұтымды пайдалану керектігі туралы түсінік береді. Біз сондай-ақ климаттың өзгеруіне байланысты адамдардың проблемалары туралы білеміз делік. Біз сондай-ақ климаттың өзгеруі мен тұрақтылығы туралы хабардар болу үшін әртүрлі платформаларда нәтижелерді жариялай аламыз. Мәтінді талдау – қажетті ақпаратты алудың маңызды аспектілерінің бірі. Мәтінді жіктеу – бұл мәтіндік доменге негізделген мәтінді әртүрлі кластарға бөлу. Бұл мәтіндік деректерді жіктеуге арналған құралдар бар табиғи тілді өңдеудің негізгі процесі. Мәтінді автоматты түрде классификациялау сандық құжаттар пайда болғаннан бері әрқашан маңызды зерттеу тақырыбы болды. Қазіргі таңда мәтінді жіктеу бүкіл әлемде күнделікті жасалынатын мәтіндік құжаттардың көптігіне байланысты өте қажет. Мәтіндік аналитика мәтінді сандарға түрлендіреді, құрылымдық деректерді береді және трендтерді анықтауды жеңілдетеді [2]. Деректер неғұрлым құрылымдалған болса, талдау соғұрлым жақсы болады және шешімдер соғұрлым жақсы қабылданады. Осы мақсатта машиналық оқыту (ML) қолданылады, ол жасанды интеллект (AI) бөлімі болып табылады, ол компьютерлерге бағдарламаланбаған болса да жұмыс істеуге және оқуға мүмкіндік береді [3]. Бұл зерттеу мәтінді классификациялау үшін таңдалған машиналық оқытудың әртүрлі әдістерін қолданады. Осы әдістерден басқа, мәтінді классификациялаудың әртүрлі тәсілдері бар, бірақ сол тәсілдердің көбі тиімді нәтиже беретін машиналық оқыту әдістерімен салыстырғанда мәтіндік деректерді дәл

классификацияламайды [3]. Мәтіндерді классификациялаудің бірнеше тиімді тәсілдері жасалғанына қарамастан, мәтіндерді классификациялау тиімділік тұрғысынан жақсартуға көптеген мүмкіндіктер бар күрделі міндет болып қалуда [5]. Алайда ұйымдар мен кәсіпорындар өздерінің өндірістік және мемлекеттік қызметтерін бақылау үшін мәтіндік құжаттарды пайдаланады [4,6]. Мәтінді классификациялау жүйесінде классификатор негізгі бөлік болып табылады. Классификатордың жұмыс сапасы мәтінді классификациялаудің тиімділігі мен әсерімен тікелей байланысты. Классификаторлардың көпшілігі мәтінді классификациялау мақсатында енгізілген ақпаратты іздеу әдістеріне және машиналық оқыту алгоритмдеріне негізделген [6]. Дегенмен, жақсы мәтіндік классификатор көп функциялы оқу деректерінің үлкен жиынтығы үшін тиімді жұмыс істейді [7]. Нысандарда жоғары өлшем мен шудың болуына байланысты мәтін санатталған жағдайда тек маңызды нысандарды таңдау өте маңызды. [9]. Бұл мақалада әртүрлі мәліметтер жиынтығында машиналық оқыту әдістерін қолдана отырып, мәтіндерді классификациялауға негізделген салыстырмалы талдау жасалынады. Мәселе мынада, мәтіндік деректерді қолмен классификациялау процесі жалықтырып жіберуі мүмкін және көп уақытты қажет етеді. Сондықтан процесті автоматтандыру және деректерге негізделген шешімдер қабылдауды жақсарту өте маңызды [8]. Бұл зерттеуде машиналық оқыту алгоритмдері әртүрлі деректер жиынында ең жақсы өнімділік пен тиімділікке қол жеткізу үшін қолданылады және салыстырылады. Мәтінді классификациялау моделіндегі құжаттар әр түрлі кезеңдерден өтеді, атап айтқанда: (i) негізгі құжатты қарапайым мәтінге, толық кіші құжатқа айналдыру, тоқтату сөздерін жою, әр түрлі сөздерді түбір сөз деп аталатын бір сөзге қысқарту, пайдалы емес сөздерді жою, стемминг пен лемматизацияны қолдану және (ii) оқыту және тестілеу үшін деректерді таңдау, классификаторды жасау, содан кейін классификаторды әртүрлі деректер жиынына орналастыру [9]. Сонымен қатар, Машиналық оқыту әдістерін пандемия уақытындағы шешім қабылдауды өлшеу және халық арасындағы қате түсініктерді анықтау үшін классификациялау мәселелерін шешу үшін де қолданылған, бұл бізге денсаулық сақтау ұйымдарын ақпараттандыруға көмектеседі, содан кейін халыққа білім берудің тиімді әдістерін жасауға болады.

Машиналық оқыту – индустрия 4.0-ді кәсіпорындар мен өнеркәсіп салаларына енгізуге мүмкіндік беретін маңызды техникалық әзірлемелердің бірі. 4.0 индустриясына жасанды интеллект пен машиналық оқытуды енгізу өндірістік ұйымдар үшін айтарлықтай өзгерісті білдіреді, бұл бизнестің жаңа перспективаларына және өнімділікті арттыру сияқты артықшылықтарға әкелуі мүмкін. Жасанды интеллект, машиналық оқыту, терең оқыту, денсаулық сақтау, қаржы, ақылды кәсіпорындар сияқты салаларда да кеңінен қолданылады 4.0 [10].

2. Негізгі бөлім

Мәтінді классификациялау – табиғи тілді өңдеудегі (NLP) негізгі міндеттердің бірі [10]. Интернет қосымшаларының қарқынды өсуіне байланысты онлайн мәтіндердің көлемінің өсуі құжаттарды автоматты түрде ұйымдастыра алатын және жіктей алатын жетілдірілген мәтінді автоматты талдау классификаторларын қажет етеді. Көптеген машиналық оқыту алгоритмдері жіктелген оқу құжаттарының жиынтығына негізделген оқыту арқылы автоматты мәтіндік классификатор жасау үшін қолданылды. Ағылшын, орыс, француз, испан, қытай және басқа да көптеген тілдерге арналған мәтіндерді классификациялаудің бірнеше модельдері бар [8,9]. Тірек векторлық машина (SVM) - екі топтық классификациялау есептері үшін классификациялау алгоритмдерін пайдаланатын бақыланатын машиналық оқыту үлгілерінің бірі. Бұл жұмыс мәтінді интеллектуалды талдауда қолданылатын бірқатар мәтіндік классификациялаушыларды қолданады және салыстырады [8]. Әдетте бақыланатын және бақыланбайтын – бұл мәтінді классификациялау үшін қолданылатын классификаторлардың екі санаты. Мәтіндерді классификациялауде, "белгісіз" NLP мәтінін оқыту процесінде ақпарат жинаудың ең маңызды элементі құжаттар пакетін көрсетілген жиынтықтан санаттарға (немесе сыныптарға немесе пәндерге) автоматты түрде классификациялау болып табылады. Әр түрлі құралдар мен әдістер мәтіндерді классификациялауде бірнеше қолданысы бар доменнің туындылары болып табылады. SVM басқа алгоритмдермен де салыстырылады, бірақ SVM әртүрлі зерттеулерде басқалардан жоғары. Құжаттың қай санатқа жататынын анықтау үшін әртүрлі классификаторлар қаралады және талданады. Біз үлкен өлшемді деректерді классификациялау үшін ядро функцияларын қолдана

отырып, сызықтық емес деректерді жіктей аламыз. Тірек векторлық машина жоғары өнімділікті қамтамасыз етеді, бірақ аз жауап береді, бұл тірек векторлық машинаны пайдаланудың шектеулерінің бірі. К

- жақын көршісінде (k-NN) объектіні дұрыс классификациялау үшін көпшілік дауыспен дауыс беру әдісі қолданылады. Демек, бұл тестілеу үшін аз мәтіндік деректерді және оқыту үшін үлкен мәтіндік деректерді қолданатын, ең жақсы өнімділік пен тиімді нәтиже беретін жаңадан енгізілген әдіс. Сондай-ақ, машиналық оқыту моделінің негізгі мақсаты автоматты оқыту және тәжірибеге негізделген модельдің тиімділігін арттыру болып табылады. Ұсынылған классификациялау моделі үш негізгі модульді қамтиды, атап айтқанда шикі деректерді алдын ала өңдеу, машиналық оқытуды пайдалану және классификациялау. Бұл модель модельдің өнімділігін жақсарту үшін өткен деректерден немесе тәжірибелер мен оқыту арқылы үйренеді. Цифрлық құжаттардың пайда болуымен мәтінді автоматты түрде классификациялау зерттеудің маңызды саласына айналды. Машиналық оқыту әдістерінің ішінде SVM классификаторы классификациялау тапсырмаларының көпшілігінде, атап айтқанда ауруларды анықтау және бетті тану үшін жақсы нәтижелерге қол жеткізеді [10]. Сонымен қатар, бүйректің созылмалы ауруын болжау үшін Кездейсоқ орман(Random forest), k-жақын көршісінде(kNN) және Аңғал Байес(Naïve Bayes) сияқты машиналық оқытудың үш алгоритмі қолданылды. Кездейсоқ орман ең жақсы нәтиже беретіні дәлелденді. Нақты болжамдар мен жақсы жалпылауға кездейсоқ іріктеу стратегиялары мен әдістеме көмегімен қол жеткізуге болады. Машиналық оқыту әдістері медициналық сектордағы медициналық жазбалардан ауруларды талдауда да маңызды рөл атқаруы мүмкін. Машиналық оқыту әдістері күн сайын өсіп келе жатқан дәрілерді анықтау үшін де қолданылады. Жаңа талаптарды болжау үшін пайдаланылған дәрі-дәрмектерге негізделген алдыңғы деректерді талдау оңай. Денсаулық сақтау ұйымдары мен денсаулық сақтау мамандары үшін олардың алаңдаушылығын, күйзелісін және жарақатын не тудыратыны туралы жалпы жұртшылықтың пікірін түсіну, содан кейін машиналық оқыту әдістерін пайдалана отырып, қолжетімді деректерге негізделген тиімді саясат пен емдеу әдістерін әзірлеу маңызды. Қазіргі уақытта кез-келген жаңалық немесе ақпарат көптеген әлеуметтік медиа арналарда тез таралады және біз оған сенетін болсақ та, оның шынайы немесе шынайы емес екенін білуге мүмкіндік жоқ. Көптеген адамдар бұл платформаны саяси, діни немесе басқа себептермен қоғамдық пікірді манипуляциялау үшін қару ретінде пайдаланады, бірақ біз жаңалықтардың немесе ақпараттың шынайы немесе жаңылыстыратын ақпарат екенін анықтау үшін машиналық оқыту алгоритмдерін пайдалана аламыз [11]. Әлеуметтік медиа сайттарында қол жетімді мәтіндік деректерден эмоцияларды, пікірлерді және көзқарастарды алудың дәстүрлі әдісін машиналық оқыту алгоритмдерін қолдану арқылы да жүзеге асыруға болады. Modelssto жақсы өнімділікпен жақсы нәтиже беруі үшін деректерді алдын ала өңдеудің әртүрлі әдістері де қажет. TF-IDF – бұл енгізілген мәтіннен маңызды ақпаратты алудың статистикалық көрсеткіші, бірақ ол теңгерімі жоқ үлестірімдерде тиімсіз. Дегенмен, біз модельдің тиімділігі мен сенімділігін арттыру үшін TF-IDF жаңартылған нұсқасын қолдана аламыз. Мәтінді классификациялауға арналған caps-net негізіндегі көп тапсырмалы оқыту архитектурасы сияқты мәтіндік деректерді классификациялаудың көптеген қосымша әдістері бар, бірақ сол мәселені шешуге арналған машиналық оқыту тәсілдерімен салыстырғанда, машиналық оқыту әдістерін қолданған кезде нәтижелер айтарлықтай жақсы болады. Бұл машиналық оқытудың әртүрлі тәсілдерін қолдана отырып, жалпы жұртшылықтың қазіргі және болашақ қажеттіліктерін талдау үшін медициналық жазбаларды жинауды қиындатады. Сонымен қатар, мәтіндерді классификациялау үшін машиналық оқытудың тиімділігіне қатысты әдебиеттерге жүйелі шолу 1-кестеде келтірілген.

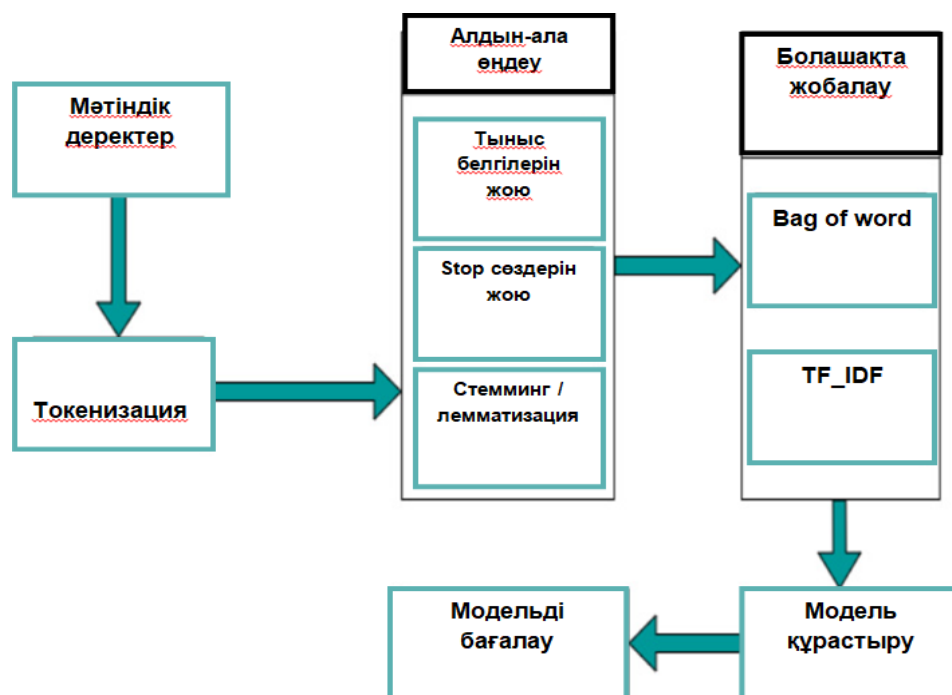
Кесте 1. Әдебиетке жүйелі шолу.

Атауы	Автор(жыл)	Әдістемелер	Қорытындылар
-------	------------	-------------	--------------

Classification of Customer Reviews Using Machine Learning Algorithms [14]	Behrooz Noori.(2021)	(SVM), Жасанды нейрондық желі (ANN), аңғал Байес (NB), шешім ағашы(DT), k-жақын көрші (K-NN)	Пайдаланушының керемет тәжірибесіне әсер ететін ең маңызды факторлар шешім ағашы алгоритмі көмегімен бөлінді
A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services [15]	Manuel J. Sánchez-Franco, Antonio Navarro-García, Francisco Javier Rondán-Cataluña.(2018)	аңғал Байес	Аңғал Байес алгоритмі қонақүй жайлы кері байланыстарды жоғары дәлдікпен және жауаптылықпен, сондай-ақ төмен есептеу шығындарымен классификациялайды.
COVID-19 World Vaccination Progress Using ML Classification Algorithms [16]	Abdulkareem, N. M., Abdulazeez, A. M., Zeebaree, D. Q., & Hasan, D. A.(2021)	Шешім ағашы, K-NN, Аңғал Байес Кездейсоқ орман	Шешім ағашы уақыт пен дәлдіктің басқа алгоритмдерінен асып түседі
Speech emotion recognition based on SVM and KNN classifications fusion [17]	Al Dujaili, M. J., Ebrahimi-Moghadam, A., & Fatlawi, A.(2021)	SVM K-NN	Сөйлеушінің тікелей сезімдерін анықтау мәселесін шешу үшін жеті сезімді классификациялауға арналған сезім моделі ұсынылған
Applications of SVM Learning in Cancer Genomics [18]	Huang. S., Cai. N., Pacheco.P., Narrandes. S., Wang. Y., Xu. W. (2018)	SVM	Қатерлі ісік геномикасын анықтаудағы SVM күшті жақтарын түсіну
Efficient English Text Classification Using Selected Machine Learning Techniques [19].	X. Luo. (2021)	SVM, аңғал Байес, Логистикалық регрессия	Дәлдік және F1 мәні классификаторды бағалау үшін есептеледі, онда SVM екі деректер жиынтығында екіншісінде жоғары, ал логистикалық регрессия бір деректер жиынтығында жоғары
Classification of movie reviews using term frequency-inverse document frequency and optimized machine learning algorithms[20].	Naeem MZ, Rustam F, Mehmood A, Mui-Zzud-Din, Ashraf I,Choi GS. (2022)	тірек векторлық машиналар (SVM), аңғал Байес классификаторы, кездейсоқ орман және градиентті күшейтетін классификаторлар	SVM TF-IDF функцияларын қолдана отырып барлық классификаторлар арасында ең жоғары дәлдікке жетті.

3. Ұсынылған жүйе және әдістеме

Бұл зерттеу жұмысында қолданылатын әдістемелер машиналық оқыту әдістеріне негізделеді, атап айтқанда тірек векторлық машина (SVM), k-ең жақын көрші, аңғал Байес, Гаусс (GNB), көпмүшелік аңғал Байес (MNB) және логистикалық регрессия (LR). Машиналық оқытуға негізделген классификациялау модельдері әр модельдің дәлдігі тұрғысынан әр түрлі мәліметтер жиынтығында салыстырылады. Классификациялау моделін жасамас бұрын кіріс деректерін және алдын ала өңделген деректерді алдын ала өңдеу үшін әртүрлі әдістер қолданылады, содан кейін олар оқыту және тестілеу мақсатында пайдаланылады [11]. Деректердің бір бөлігі оқыту үшін алынады, ал қалғаны тестілеу үшін пайдаланылады, бірақ оқыту және тестілеу үшін өңделген деректер оқытуда қолданылатын әдістеме негізінде бөлінеді. Бұл машиналық оқытуға негізделген мәтінді классификациялау процесі Сурет 1-де көрсетілген.



Сурет 1 Ұсынылған Модельдің технологиялық схемасы.

3.1. Машиналық оқыту әдістері

Барлығымыз білетіндей, мәтіндік деректердің көлемі экспоненциалды түрде артады. Сондықтан оны қолмен классификациялау оңай емес, сондықтан қысқа мерзімде деректердің үлкен көлемін классификациялаудың әртүрлі мүмкін жолдарын анықтаған жөн. Классификациялауден кейін жасалған деректер ақпарат деп аталады, содан кейін бұл ақпарат болашақта бизнесті және өнеркәсіптік қосымшаларды жоспарлау үшін қолданылады. Бұл жұмыс Кесте 2-де көрсетілгендей мәтінді классификациялау үшін әртүрлі машиналық оқыту алгоритмдерін қолдануды ұсынады. Қандай машиналық оқыту алгоритмі деректер жиынтығының қай түріне жоғары дәлдік беретінін анықтау қажет. Бұл салыстырмалы талдау әртүрлі машиналық оқыту алгоритмдерінің тиімділігін зерттейді, содан кейін қандай алгоритм қандай да бір деректер түріне жақсырақ сәйкес келетінін анықтайды, өйткені біз әртүрлі машиналық оқыту алгоритмдері мәтіндік деректерді басқаша жіктейтінін білеміз. Нәтижесінде деректер жиынтығының белгілі бір түріне қандай тәсіл сәйкес келетінін анықтау өте маңызды. Машиналық оқытудың барлық қолданылатын әдістерінің егжей-тегжейлі анықтамалары келесі бөлімде келтірілген.

Кесте 2. Машиналық оқыту алгоритмдерінің қысқаша сипаттамасы.

Әдістер	Артықшылықтары	Әлсіз тұстары	Қосымшалар
Тірек векторлық машина (SVM)	Жоғары ретті деректерді ядро функциялары арқылы классификациялауға болады.	Үлкен деректер жиынтығына сәйкес келмеуі мүмкін және ядро функциясын таңдау қиын.	Қолжазбаны тану, мәтін мен гипермәтінді классификациялау, кескіндерді классификациялау.
К-ең жақын көрші (KNN)	Шулы деректер үшін оңтайлы. Енгізу өте қарапайым, оқуға уақытты қажет етпейтін оқу деректерінің үлкен көлеміндегі деректерге тиімді нәтижелер береді.	Есептеу құны жоғары, k мәнін табу қиын.	Денсаулық сақтау саласы, сегменттеу, тұтынушыларға қызмет көрсету, алаяқтықты анықтау.
Көпмүшелік аңғал Байес алгоритмі (MNB)	Жүзеге асыру оңай. Ең жақсы нәтижелер көптеген алдыңғы оқыту кезінде алынды, аз мөлшерде деректері қажет болады.	Ықтималдықтар дәл емес болғандықтан функциялар арасындағы өзара әрекеттесуге қол жеткізу мүмкін емес	Нақты уақыттағы болжау, деректерді сүзу, көңіл-күйді талдау.
Логистикалық регрессия (LR)	Деректерді жылдам өңдеу, категориялық деректермен жақсы жұмыс істейді, параметрлерді қарапайым бағалау, сызықтық деректер үшін жақсы.	Үлкен өлшемдерді қажет ететін емес сызықтық емес деректер үшін тиімді емес	Медицина, мәтінді өңдеу, қонақүйлерге бронь жасау, қаржылық болжам жасау.
Кездейсоқ орман (RF)	Үлкен деректер жиынтығында қолдануға жақсы нәтижелер береді, айнымалыны өзара әрекеттесу кезінде анықтаудың эксперименттік әдісі	Көп мәнді және анықталмаған атрибуттар үшін күрделі болып келеді, бұл үлкен өңдеу қуатын қажет етеді.	Банк секторы, денсаулық сақтау секторы, клиенттерді талдау, маркетингтік деректер секторы.

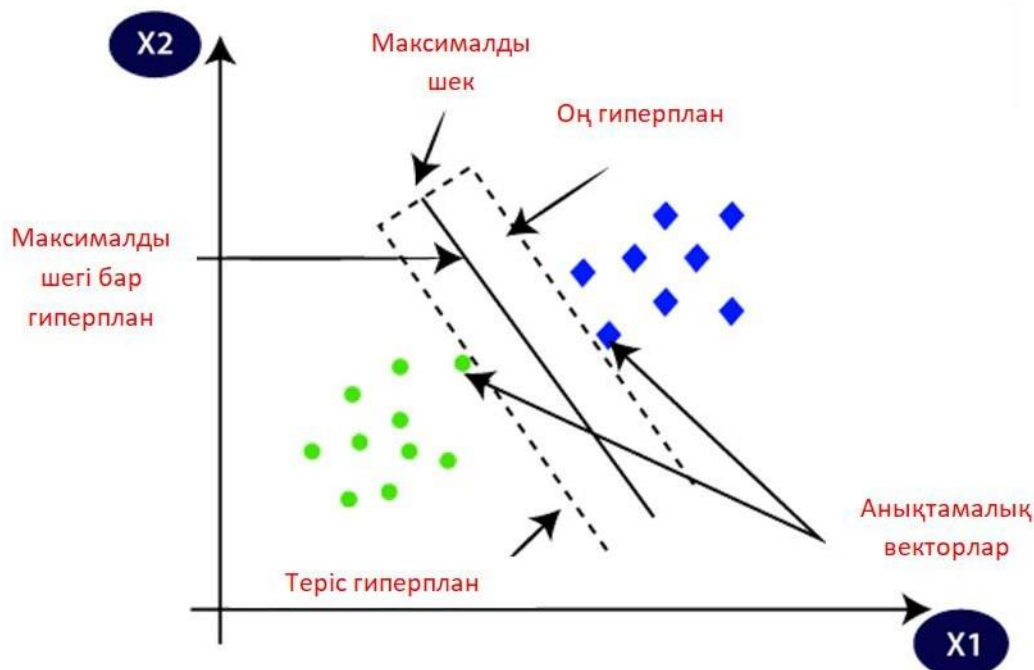
3.1.1. Тірек векторлық машина

Тірек векторлық машина – регрессия үшін де, классификация үшін де қолдануға болатын машиналық оқыту әдісі, бірақ ол классификациялау есептері үшін ең қолайлы [12]. Ол сызықтық деректерді анықтамалық векторлар деп аталатын деректер нүктелері арасындағы қашықтық максималды болатын максималды өріс гиперпланының (ММН) көмегімен жіктей алады. Деректерді бөлетін екі параллель сызық оң және теріс гиперпландар деп аталады, өйткені біз бірнеше сурет сала аламыз. Сызықтық емес деректер үшін ядро функциясын классификациялау үшін көп өлшемді гиперплан қалыптастыру үшін пайдалануға болады. Классификациялау мақсатында көптеген ядро функциялары бар. Зерттеушілер жол тізбегінің ядросы (SSK) және жуықтайтын ядролар (АК) сияқты ядро функцияларын пайдаланды. Бұл екі ядро мәтіндік деректерді жоғары дәлдікпен жіктей алатын классификатор жасайды. Анықтамалық векторлық машина есептеуде тиімді, кейбір шектеулер оның шағын деректер жиынтығы үшін өнімділігін төмендетеді. Деректерді классификациялаудың екі түрі бар SVM (i) деректердің сызықтық жіктелуі

және (ii) деректердің сызықтық емес жіктелуі.

(i). Деректердің сызықтық жіктелуі

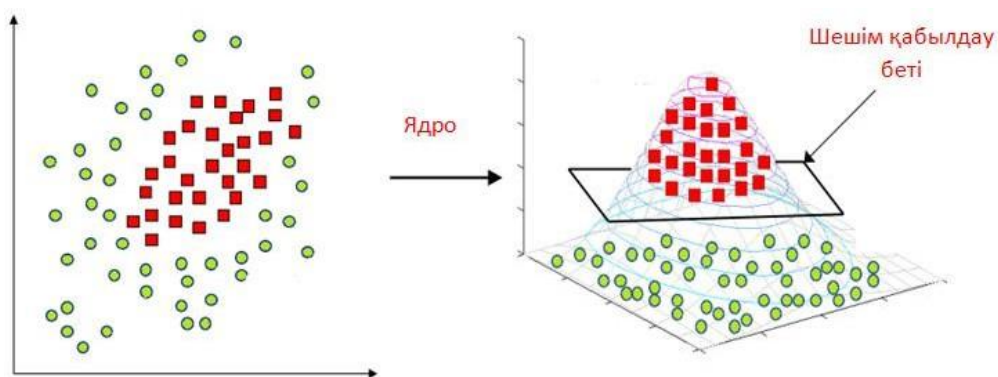
SVM көмегімен сызықтық деректерді классификациялау үшін максималды өрісгиперпланы (ММН) көптеген гиперпландарды салу үшін екі деректер нүктесін бөлу үшін қолданылады. Векторлық нүктелер арасындағы ең үлкен қашықтықты тапқан жөн, бұл суретте көрсетілгендей деректер нүктелерін дәл классификациялауға мүмкіндік береді. Сурет 2 – де көрсетілгендей, оң және теріс гиперпландар бар, ал оң гиперплан оң деректер нүктесінің жағында сызылған. Керісінше, теріс гиперплан теріс деректер нүктесінің жағында сызылады [62]. Оң және теріс гиперпландар арасындағы максималды мәнді алу үшін гиперпландарды салған дұрыс.



Сурет 2 SVM көмегімен деректерді сызықтық классификациялау.

(ii). Деректердің сызықтық емес жіктелуі

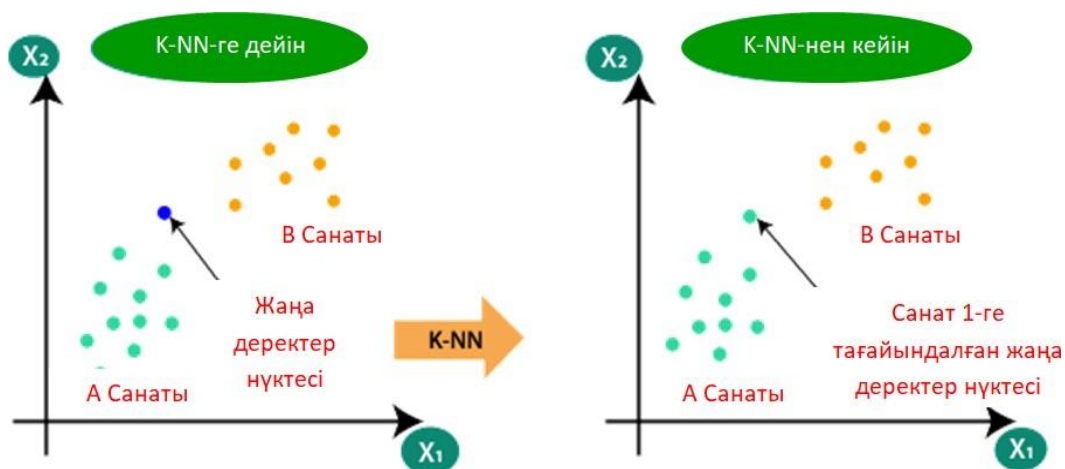
Анықтамалық векторлық Машина (SVM) ядро функциясы арқылы сызықтық емес деректерді де жіктей алады. Ол Сурет 3 те көрсетілгендей классификациялауді орындау үшін деректерді жоғары өлшемдерге түрлендіреді. Біз классификациялау мақсатында қолдануға болатын ядро функцияларының әртүрлі түрлері бар [13]. Бұл әдіс деректер нүктелерін сәйкес классификациялау үшін дұрыс ядро функциясын анықтауы керек. Деректер қорындағы нүктелерді классификациялау үшін ядро функциясы пайдаланылған кезде, ол бір деректер класын жоғары деңгейдегі өлшемге дейін түрлендіреді және деректер қорындағы нүктелерді классификациялау үшін шешім қабылдау беті алынады.



Сурет 3 SVM көмегімен деректерді сызықтық емес классификациялау.

3.1.2. К-жақын көршілер классификаторы (KNN)

К-жақын көрші алгоритмі – классификациялау және регрессия мәселелерін шешу үшін пайдалануға болатын қарапайым, оңай орындалатын, басқарылатын Машиналық оқыту алгоритмі [13]. Бұл алгоритм қол жетімді және жаңа деректердің ұқсастығын табады, ал жаңа деректер ұқсастығы жоғары санатқа жіктеледі [12, 13]. К мәнін талдау қиын, сондықтан k-NN бойынша классификациялау уақыты ұзағырақ. Оны жалқау оқыту алгоритмі деп те атайды, өйткені ол кенеттен оқу деректерінен сабақ алмайды, бірақ Сурет 4 те көрсетілгендей классификациялау кезінде әрекет етеді.



Сурет 4 K-NN жұмысы.

3.1.3. Көпмүшелік аңғал Байес алгоритмі (MNB)

MNB классификациялау алгоритмі дискретті белгілерді классификациялау үшін қолданылады (мысалы, мәтінді классификациялауға арналған сөз жиілігі) [12]. Көпмүшелік үлестіру кезінде белгілердің бүтін саны қажет, бірақ іс жүзінде TF-IDF сияқты жағдайларда бөлшек мәндер де жұмыс істей алады [14].

"Сөз қапшығы" тәсілі қолданылады және әр сөз сөз тәртібі маңызды емес функцияны білдіреді. Аңғал Байес Байестің шартты ықтималдық ережесіне негізделген [15] және MNB төмендегі теңдеуде математикалық түрде анықталады.

$$P(h|x) = \frac{P(x|h)P(h)}{P(x)}$$

Мұндағы h -гипотеза, ал x -атрибут.

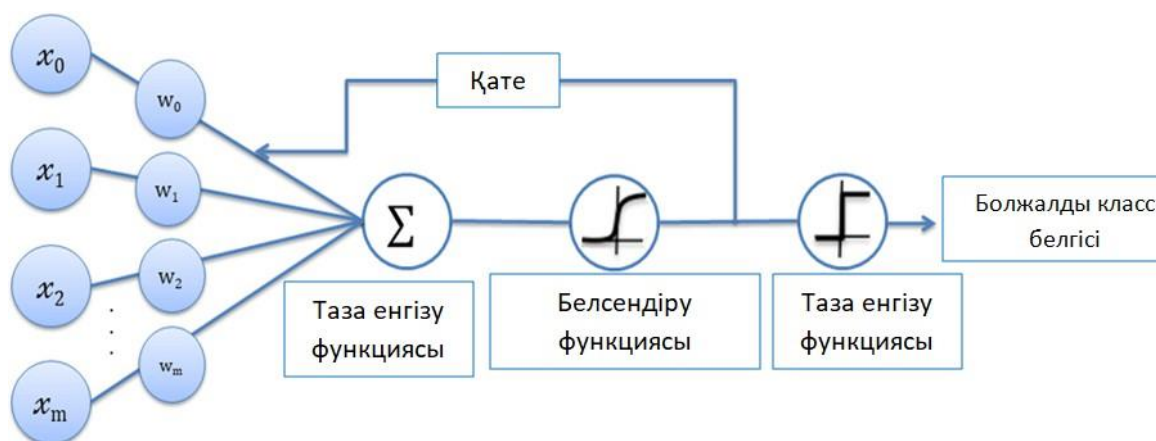
3.1.4. Логистикалық регрессия (LR)

Логистикалық регрессия – бұл классификациялау мақсатында қолданылатын машиналық оқытудың бақыланатын алгоритмі. Ол деректер екілік түрінде ұсынылған кезде қолданылады, яғни 0 және 1, бұл кластың бір санатқа немесе екіншісіне жататындығын білдіреді. Біз екілік мәндер үшін екі функцияны қолдана аламыз, атап айтсақ логистикалық функция және сигмоидтық функция [15]. Логистикалық регрессия, оны классификациялау алгоритмі деп те атайды, Сурет 5-те көрсетілген. Логистикалық регрессияны төменде көрсетілгендей санаттарға қарай классификациялауға болады.

(i) Биномдық: мақсатты айнымалыда тек екі түрін қабылдауы мүмкін: "0" немесе "1" "жеңіс" пен "жеңілісті" салыстырғанда, "сәттілік" пен "сәтсіздікті" салыстырғанда, "өлген" мен "тіріні" салыстырғанда және т. б.

(ii) Көпмүшелік: мақсатты айнымалыларда үш немесе одан да көп типтер болуы мүмкін, олар реттелмеген (яғни түрлердің сандық мәні жоқ), мысалы "А вирусы" және "В вирусы" және "С вирусы".

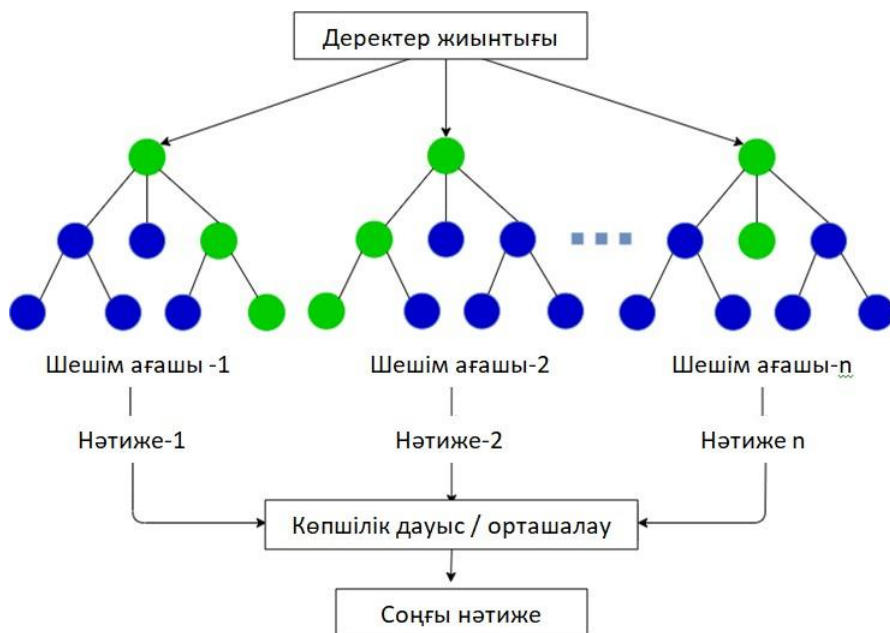
(iii) Реттік нөмір: мақсатты айнымалыдағы реттелген кластар; мысалы, баллды "өте жақсы", "жақсы", "жаман" және "өте жаман" деп классификациялауға болады. Мұнда әр санатқа 0, 1, 2, 3 типті балл берілуі мүмкін немесе керісінше.



Сурет 5 Логистикалық регрессия.

3.1.5. Кездейсоқ орман

Деректерді классификациялаудың бірнеше алгоритмдері бар, бірақ кездейсоқ орман (Сурет 6) машиналық оқытудағы ең жақсы классификациялау алгоритмдерінің бірі. Ол регрессия әдісі ретінде де қолданыла алады, бірақ көбінесе әртүрлілігі мен қарапайымдылығына байланысты классификациялау үшін қолданылады. Бұл қорытынды нәтижені [12] арттыратын оқу үлгілерінің комбинациясы. Бұл машиналық оқыту техникасында көптеген ағаштар кездейсоқ орман жасау үшін бірігеді. Егер бізде корреляцияланбаған ағаштар көп болса, біз жоғары дәлдікке ие боламыз [10]. Жетіспейтін мәндерді кездейсоқ орман арқылы толтыруға болады [11]. Сонымен қатар, шешімдердің ағаш классификаторлары керемет өнімділігімен танымал, өйткені кездейсоқ орман шешім ағаштарының жиынтығы екені белгілі, сондықтан ол сенімдірек және күштірек болады. Классификациялау есептері үшін қолданылатын қарапайым шешім ағашы жоғары дәлдікпен жақсы нәтиже береді [69].



Сурет 6 Кездейсоқ орманды қолдану арқылы классификациялау.

4. Нәтижелер

Бұл бөлімде екі бөлек деректер жиынтығына қолданылған әртүрлі Машиналық оқыту алгоритмдерінің нәтижелері қарастырылады. Әрбір алгоритм дәлдік, дәлдік, еске түсіру және F1 ұпайы сияқты әртүрлі өнімділік көрсеткіштерін пайдалана отырып, Машиналық оқыту алгоритмінің тиімділігін анықтау үшін бөлек қолданылды. Біз шынымен теріс, жалған оң және жалған теріс осы шынымен оң бағалау көрсеткіштері туралы білмес бұрын бағалау көрсеткіштерінің негізгі блоктарын түсінуіміз керек.

Шынайы позитивті: позитивті сыныпты модель немесе классификатор дұрыс болжайды. Мұны TP ұсынуы мүмкін.

Шын теріс: модель немесе классификатор теріс класты дұрыс болжайды. Бұл TN арқылы ұсынылуы мүмкін.

Жалған оң нәтиже: модель немесе классификатор оң қоңырауды дұрыс болжамайды. Бұл FP арқылы ұсынылуы мүмкін.

Жалған теріс: модель немесе классификатор теріс класты дұрыс болжамайды. Бұл FN көмегімен ұсынылуы мүмкін.

Дәлдік: бұл машиналық оқыту моделін бағалаудың бір көрсеткіші, оның көмегімен классификациялау іштің деректерді қаншалықты дәл классификациялайтынын айта аламыз. Біз осы формуланы қолдана отырып дәлдікті есептейміз:

$$Accuracy = \frac{TP + TN}{P + N}$$

Болжам дәлдік біздің модельдің қаншалықты дәл екенін айтады (қанша оң анықталған кластар дұрыс болды). Біз осы формуланы қолдана отырып дәлдікті есептейміз:

$$Precision = \frac{TP}{TP + FP}$$

Кері байланыс: бұл біздің модельдің қаншалықты толық екенін айтады (нақты оң нәтижелер қаншалықты дұрыс анықталды). Біз осы формуланы пайдаланып кері байланысты есептейміз:

$$Recall = \frac{TP}{TP + FN}$$

F1-бағалау: дәлдік пен еске түсірудің орташа гармоникалық мәні дәлдік пен ескетүсірудің теңдестірілген нәтижесін береді. Біз f1 ұпайын осы формула арқылы есептейміз:

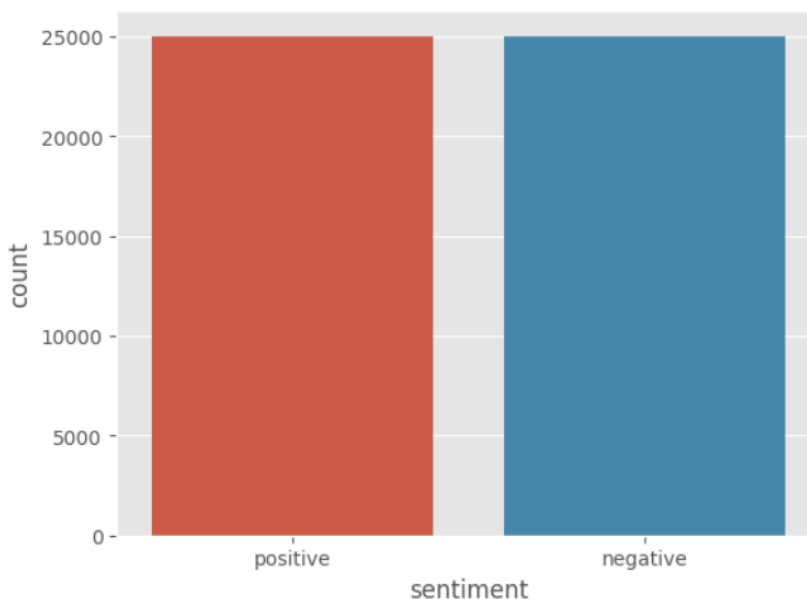
$$F1 - Score = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

4.1. Пайдаланылған деректер жиынтығы

Бұл жұмыста біз екі деректер жиынтығын қолдандық Және бұл деректер жиынтығы онлайн репозиторийлерден жиналды. Деректер жиынтығы әр алгоритмнің тиімділігін талдау үшін әртүрлі Машиналық оқыту алгоритмдері арқылы талданады. Деректер жиынының сипаттамасы төмендегі бөлімде берілген.

4.1.1. IMDb деректер жиынтығы

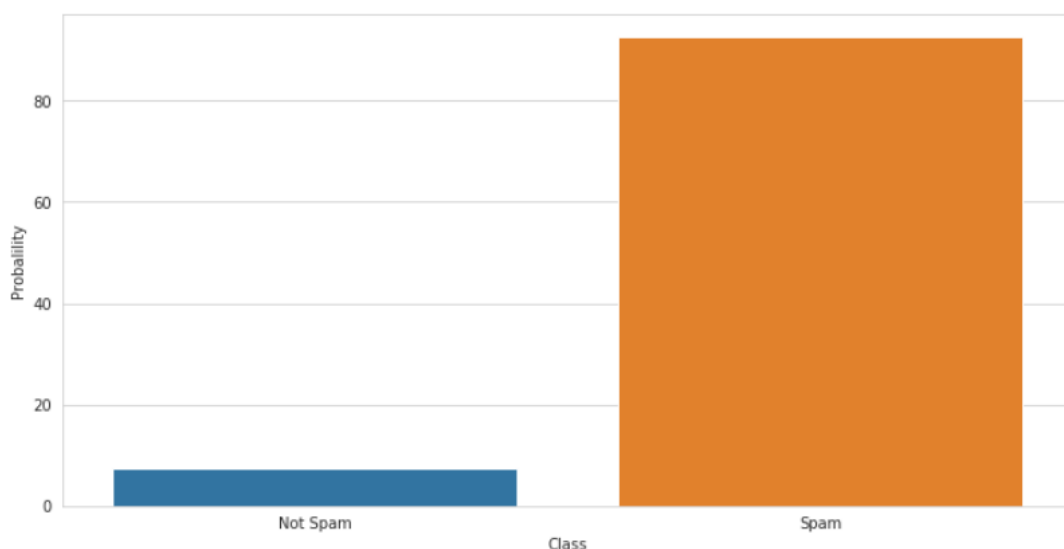
Бұл деректер жиынтығы Интернетте қол жетімді фильмдерді қарастырады, оларда бізде екі атрибуты бар 52 мың жазба бар. Олардың бірі – шолу. Екіншісі – сурет 7 де көрсетілгендей көңіл-күй. Бұл деректер жиынында оң және теріс көңіл-күйлердің саны тең, сондықтан бұл деректер жиыны теңдестірілген деректер жиыны деп те аталады, яғни деректер бұрмаланбаған.



Сурет 7 Көңіл-күй атрибуттарының графикалық көрінісі.

4.1.2. E-mail хабарламалар туралы мәліметтер жиынтығы

Бұл деректер жиынтығы бізде екі белгісі бар E-mail хабарламаларға арналған. Деректер электрондық пошта мәтінінен және мәтінді классификациялауға арналған белгілерден тұрады: 1) хабарлама мәтін 2) белгілер. Бұл деректер жиынтығында Сурет 8 де көрсетілгендей екі атрибуты, белгісі және хабарламасы бар 51572 жазба бар. Бұл деректер жиынтығында бізде "спам" және "спам емес" белгілердің саны әртүрлі, сондықтан бұл деректер жиынтығы теңгерімсіз деректер жиынтығы деп те аталады, яғни деректер бұрмаланған болып табылады.

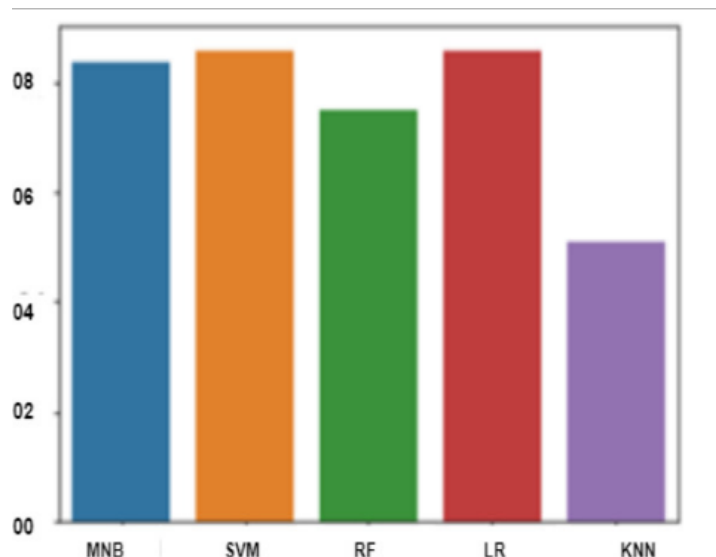


Сурет 8 Белгі атрибуттарының графикалық көріністері.

IMDb деректер жинағында бізде оң және теріс көңіл-күй бірдей. Алдымен біз тыныс белгілерін, тоқтату сөздерін, жиі сөздерді жою, стемминг және лемматизация сияқты әртүрлі алдын ала өңдеу қадамдарын пайдаланып деректерді тазаланды [70]. Алдын ала өңдеуден кейін мәтін "сөз қаптау", "термин-жиілік" моделін, құжат терминдерінің кері жиілік моделін және соңында матрицаны бағалаудың әртүрлі әдістерін қолдана отырып, алгоритмнің тиімділігін пайдаланып векторларға айналдырылды. Екінші жағынан, E-mail хабарламалар деректер жинағында әртүрлі белгілер бар, мысалы, спам жазбаларының саны, бұл спам емес жазбаларының санынан көп. Бұл деректер жиынтығы үшін F1 дәлдігін, еске түсіруді және бағалауды қолдана отырып, классификациялау ішкі бағалау қажет. 3-кестеде IMDb және E-mail хабарламалар дерекқорында алгоритмді іске асырудың дәлдігі, қайтарып алу және F1-балл көрсеткіштері көрсетілген. Графикалық көріністі қолдана отырып, осы алгоритмдердің өнімділігі қай машиналық оқыту алгоритміекіншісінен жоғары екендігі туралы нақты түсінік береді. SVM және логистикалық регрессия Сурет 9 да көрсетілгендей сәйкесінше 85,5 және 85,9% дәлдікке ие.

Кесте 3. Машиналық оқыту алгоритмдерін салыстырмалы талдау.

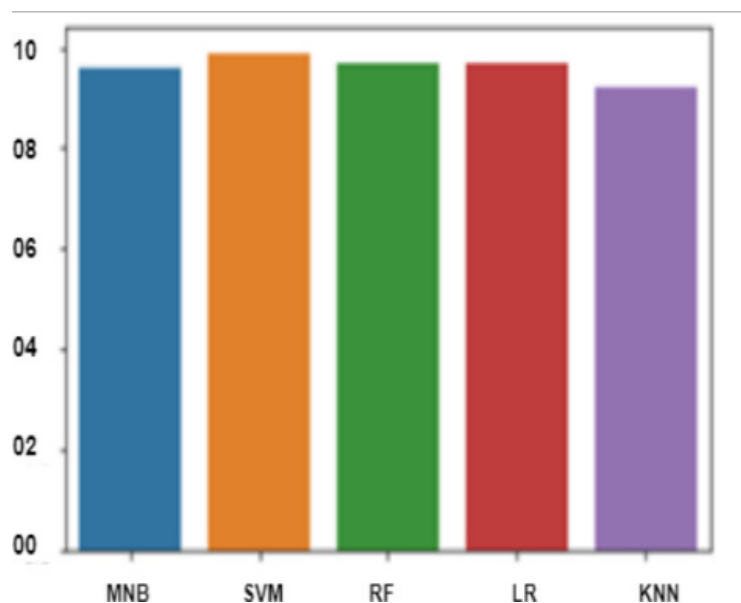
Алгоритм	Дәлдік	Болжам	F1 дәлдігі
SVM - IMDb	86.5	86	87
SVM - E-mail хабарламалар	96.5	97	96
K-nn - IMDb	51.8	51	60
K-nn - E-mail хабарламалар	98.5	99	99
Аңғал Байес - IMDb	85.4	86	87
Аңғал Байес - E-mail хабарламалар	98.4	99	98
Кездейсоқ орман - IMDb	75.9	72	78
Кездейсоқ орман - E-mail хабарламалар	97.5	98	97
Логистикалық регрессия - IMDb	86.8	86	87
Логистикалық регрессия - E-mail хабарламалар	92.9	94	91



Сурет 9 IBM дерекқорында таңдалған машиналық оқыту алгоритмдерінің дәлдігі.

Е-mail хабарламалар деректер жиынтығында тірек векторлық машинасы басқа классификатордан асып түседі. Қалған алгоритмдер тірек векторларының машинасымен бірдей дәлдікке ие болып отыр, яғни 94,5%, k-жақын көрші 97,5, көпмүшелік аңғал Байес алгоритмі 96,4, ал кездейсоқ орман 95,5% құрайды.

Сонымен қатар, логистикалық регрессия 90,9% дәлдікке ие, бұл сурет 10 да көрсетілгендей барлық классификациялаушылардың ішіндегі ең төменгі көрсеткіш. Машиналық оқыту өнімділігінің әртүрлі көрсеткіштері бойынша нәтижелер 3-кестеде келтірілген.



Сурет 10 E-mail хабарламалар жиынтығы үшін таңдалған машиналық оқытуалгоритмдерінің дәлдігі.

5. Шектеулер және болашақта орындалатын жұмыстар

Болашақта бұл зерттеу гиперпараметрлік және ансамбльдік тәсілдермен көбірек алгоритмдерді қосу үшін кеңейтілуі мүмкін. Ақпаратты тиімді анықтауды қамтамасыз ету үшін модельдерді параметрлерді оңтайландырудың жаңа стратегияларын қолдану арқылы да жүзеге асыруға болады. Мәтінді классификациялау саласында ағынды деректерді өңдеу

жеткілікті зерттелмеген және егжей-тегжейлі қарастыруды қажет етеді. Нәтижесінде, дұрыс қолданған кезде ансамбльдік және тексерілген тәсілдер мәтінді классификациялау кезінде оі нәтижелер әкеледі.

6. Қорытынды

Табиғи тілді өңдеудің ең маңызды бөлігі – мәтіндік деректерді қажетті кластар жиынтығы бойынша автоматты түрде жіктейтін мәтінді классификациялау. Мәтінді классификациялау үшін машиналық оқытуға негізделген әдістер қажет. Осылайша, бұл зерттеуде бес алгоритм қолданылады: тірек векторлық машина, k-жақын көршілердің классификациялаушы, логистикалық регрессия, Байес көпмүшелік әдісі, кездейсоқ орман және екі деректер жиынтығы: IMDb және E-mail хабарламалар. Нәтижелерден k-NN моделі 97,5% дәлдікпен E-mail хабарламалар деректер жиынтығындағы басқа модельдерден асып түсетінін көруге болады. Керісінше, логистикалық регрессия моделі ұсынылған жүйені пайдалану кезінде IMDb деректер жиынтығындағы басқа модельдерден 86,8% дәлдікпен асып түседі.

Пайдаланылған әдебиеттер тізімі

1. C.N. Kamath, S.S. Bukhari, A. Dengel, Comparative study between traditional machine learning and deep learning approaches for text classification, Proceedings of the ACM Symposium on Document Engineering 2018 (2018, August), pp. 1-11
2. A. Mohi, U. Din, K. Syed, T. Rabani, Q. Rayees, Machine learning based approaches for detecting COVID-19 using clinical text data, Int. J. Inf. Technol., 12 (3) (2020), pp. 731-739
3. D. Mahesh Matta Meet Kumar Saraf, D. Mahesh Matta, M. Kumar Saraf, S. Memeti, Prediction of COVID-19 using Machine Learning Techniques(2020)
4. H.B. Syeda, M. Syed, K.W. Sexton, S. Syed, S. Begum, F. Syed, ..., F. Yu Jr, Role of machine learning techniques to tackle the COVID-19 crisis: systematic review, JMIR medical informatics, 9 (1) (2021), p. e23811
5. D. Nagar, S. Raghav, A. Bhardwaj, R. Kumar, P.Lata Singh, R. Sindhwan, Machine learning–Best way to sustain the supply chain in the era of industry 4.0, Mater. Today Proc., 47 (2021), pp. 3676-3682, 10.1016/j.matpr.2021.01.267
6. K. Crowston, F. Bolici, Impacts of machine learning on work, Proc. Annu. Hawaii Int. Conf. Syst. Sci., 2019-January (2019), pp. 5961-5970, 10.24251/hicss.2019.719
7. N. Aljedani, R. Alotaibi, M. Taileb, Hmatc: Hierarchical multi-label arabic text classification model using machine learning, Egyptian Informatics Journal, 22 (3) (2021), pp. 225-237
8. I. Ibrahim, A. Abdulazeez, The role of machine learning algorithms for diagnosing diseases, J. Appl. Sci. Technol. Trends, 2 (01) (2021), pp. 10-19, 10.38094/jastt20179
9. M. Elbadawi, S. Gaisford, A.W. Basit, Advanced machine-learning techniques in drug discovery, Drug Discovery Today, 26 (3) (2021), pp. 769-777
10. Z. Jiang, B. Gao, Y. He, Y. Han, P. Doyle, Q. Zhu, Text classification using novel term weighting scheme-based improved TF-IDF for Internet media reports, Mathematical Problems in Engineering, 2021 (2021)
11. R.I. Kurnia, Y.D. Tangkuman, A.S. Girsang, Classification of user comment using word2vec and SVM classifier, Int. J. Adv. Trends Comput. Sci. Eng., 9 (1) (2020), pp. 643-648, 10.30534/ijatcse/2020/90912020
12. A.I. Kadhim, Survey on supervised machine learning techniques for automatic text classification, Artificial Intelligence Review, 52 (1) (2019), pp. 273-292
13. B. Charbuty, A. Abdulazeez, Classification based on decision tree algorithm for machine learning, J. Appl. Sci. Technol. Trends, 2 (01) (2021), pp. 20-28, 10.38094/jastt20165
14. Noori, B. (2021). "Classification of Customer Reviews Using Machine Learning Algorithms." *Applied Artificial Intelligence* 35(8): 567-588.
15. Manuel J. Sánchez-Franco, Antonio Navarro-García, Francisco Javier Rondán-Cataluña, A naive Bayes strategy for classifying customer satisfaction: A study based on online reviews of hospitality services, Journal of Business Research, Volume 101, 2019, P 499-506, ISSN

0148-2963.

16. Abdulkareem, N. M., Abdulazeez, A. M., Zeebaree, D. Q., & Hasan, D. A., "COVID-19 World Vaccination Progress Using Machine Learning Classification Algorithms," Qubahan Academic Journal, Vol.1, No.2, pp.100-105, 2021.

17. Al Dujaili, M. J., Ebrahimi-Moghadam, A., & Fatlawi, A., "Speech emotion recognition based on SVM and KNN classifications fusion," International Journal of Electrical and Computer Engineering, Vol.11, No.2, pp.1259, 2021.

18. Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W., "Applications of support vector machine (SVM) learning in cancer genomics," Cancer Genomics-Proteomics, Vol.15, No.1, pp.41-51, 2018.

19. X. Luo, Efficient English text classification using selected Machine Learning Techniques, Alex. Eng. J., 60 (3) (2021), pp. 3401-3409, 10.1016/j.aej.2021.02.009

20. Naeem MZ, Rustam F, Mehmood A, Mui-Zzud-Din, Ashraf I, Choi GS. Classification of movie reviews using term frequency-inverse document frequency and optimized machine learning algorithms. PeerJ Comput Sci. 2022 Mar 15;8:e914. doi: 10.7717/peerj-cs.914. PMID: 35494818; PMCID: PMC9044332.

ӘОЖ 004.056

МОБИЛЬДІ ҚОСЫМШАЛАРДЫҢ ҚАУІПСІЗДІГІН ҚОРҒАУ

Сагинтаев Абат Әлімжанұлы
sagintayev.abat@gmail.com

Л.Н.Гумилев атындағы ЕҰУ, «Ақпараттық қауіпсіздік» кафедрасының
1-ші курс магистранты, Астана, Қазақстан
Ғылыми жетекшісі – PhD, оқытушы Жеткенбай Лена

Андатпа. Қазіргі таңда мобильді қосымшалардың қауіпсіздігі өте маңызды болып табылады. Мобильді құрылғылардағы қосымшалардың қауіпсіздігі пайдаланушы деректерін әртүрлі шабуылдардан, хакерлерден, зиянды программалардан және т.б. манипуляциялардан немесе ұрланудан қорғау процесі. Осы зерттеу жұмысы барысында мобильді қосымшаларды қорғауда қолданылатын тиімді әдістер мен ақпаратты шифрлау алгоритмдері қарастырылды және қазіргі таңдағы танымал алгоритмдерге шолу жасалды.

Кілттік сөздер: ақпараттық қауіпсіздік, шифр, шифрлау алгоритмдері, ашық кілт, жабық кілт.

1. Кіріспе

Ақпараттық қауіпсіздікті қорғау өте маңызды және өзекті мәселеге айналған. Қазіргі таңда смартфондар адамдардың ең көп қолданатын және қажеттіліктерін өтейтін таптырмас құрал болып табылады. Смартфондардың танымал, әрі кең қолданысқа ие болуына әсер еткен әрине мобильді қосымшалар. Бұрын мобильді қосымшалар ойын түрінде дамыған болса, қазіргі уақытта бизнесте мықты маркетингтік құралға айналған. Жалпылай айтқанда, мобильді қосымшалардың қауіпсіздігі дегеніміз мобильді құрылғыларға арналған қосымшаларды зиянды программалардан, әртүрлі хакерлік шабуылдардан, сондай-ақ басқа да қылмыстық манипуляциялар сияқты цифрлық алаяқтықтан кешенді қорғау шарасы, әрі тиімді құралы. Мобильді қосымшалардың құрылымынан бөлек қауіпсіздігін қамтамасыз ету қажеттілігі туындады. Қауіпсіздікті қамтамасыз ету үшін яғни қауіптердің алдын алу немесе болдырмау мақсатында, сондай-ақ деректерді сыртқы шабуылдардан, ішкі шабуылдардан және де ақаулардан қорғау үшін әртүрлі тиімді әдістер мен ақпаратты шифрлау алгоритмдері қолданылады және сөзсіз қолдану қажет.

Негізгі түсініктер

Ақпараттық қауіпсіздік – бұл қарапайым мәтінді немесе хабарламаны ақпаратты оқуға арналған кілті барлардан басқа ешкім оқи алмайтындай етіп қауіпсіздікті қамтамасыз ету [1].