ҒТАМР 28.23.01

А.Б. Абен[1] – *негізгі автор*, ©
Н.М. Жунисов[2], Ж.Б. Мырзатаев[3]

[1]Магистр, оқытушы, [2]PhD, аға оқытушы, [3]Магистрант

ORCİD *[1]https://orcid.org/0000-0001-8534-3288 [2]https://orcid.org/0000-0001-6531-9408*

[1,2,3]Қожа Ахмет Ясауи атындағы Халықаралық қазақ-түрік университеті, Түркістан қ., Қазақстан

@ *[1]arypzhan.aben@ayu.edu.kz*

https://doi.org/10.55956/NXZX2857

МАШИНАЛЫҚ ОҚЫТУ АЛГОРИТМДЕРІНІҢ КӨМЕГІМЕН СПАМ ХАБАРЛАМАЛАРДЫ АНЫҚТАУ

Аңдатпа. Бұл зерттеуде мәтіндік хабарламаларды автоматты түрде классификациялау арқылы спамды анықтау әдістері талданды. Зерттеудің мақсаты – машиналық оқыту модельдерінің тиімділігін бағалау және оларды спам фильтрлеу жүйелеріне қолдану мүмкіндіктерін зерттеу болды. Алдын ала деректерді өңдеу кезеңдерінде хабарламаларды төменгі регистрге ауыстыру, токенизация, арнайы символдар мен тыныс белгілерін жою, жиі кездесетін сөздерді алып тастау және сөздерді түбірге келтіру қолданылды. Осы қадамдар хабарламалардан маңызды ақпаратты бөліп көрсетуге және классификацияның сапасын арттыруға көмектесті.

Зерттеу барысында Наивті Байес, Логистикалық регрессия, SVM, Шешім ағашы, Рэндом Форест және Градиентті Бустинг модельдері сынақтан өткізілді. Нәтижелер көрсеткендей, SVM моделі барлық метрикалар бойынша ең жоғары көрсеткіштерге ие болды, сонымен қатар Рэндом Форест пен Логистикалық регрессия модельдері де жоғары тиімділігін көрсетті. Бұл модельдер нақтылық, дәлдік, еске түсіру және F1-өлшемі бойынша 95%-дан жоғары көрсеткіштерге қол жеткізді.

Зерттеу нәтижелері машиналық оқыту әдістерінің мәтіндік деректерді классификациялаудағы тиімділігін дәлелдейді және оларды нақты қолданбалы жүйелерге, мысалы, электрондық пошта немесе SMS арқылы келетін хабарламаларды сүзу жүйелеріне енгізу мүмкіндіктерін көрсетеді. Алдағы уақытта гиперпараметрлерді оңтайландыру және мәтіндік деректерді өңдеудің жетілдірілген әдістерін қолдану арқылы модельдердің өнімділігін арттыру жоспарланып отыр.

Тірек сөздер: машиналық оқыту, спам фильтрлеу, текст классификациясы, Наивті Байес, логистикалық регрессия, SVM моделі, деректерді өңдеу, градиентті бустинг.

Абен, А.Б. Машиналық оқыту алгоритмдерінің көмегімен спам хабарламаларды анықтау [Мәтін] / А.Б. Абен, Н.М. Жунисов, Ж.Б. Мырзатаев //Механика және технологиялар / Ғылыми журнал. – 2024. – №4(86). – Б.440-450. https://doi.org/10.55956/NXZX2857

Кіріспе. Ақпараттық технологиялардың қарқынды дамуы мен интернеттің кеңінен қолданылуы адамдарға жаңа мүмкіндіктермен қатар киберқауіп-қатерлерді де әкелді. Соның ішінде электрондық пошта мен хабар алмасу жүйелеріндегі спам хабарламалар ерекше назар аударуды қажет ететін мәселеге айналып отыр. Спам – бұл қолданушыларға қажетсіз,

жарнамалық немесе зиянды сипаттағы мазмұнды электрондық хаттар немесе хабарламалар, және олар кейде фишингтік шабуылдар, зиянды бағдарламаларды тарату және жеке деректерді ұрлау мақсатында пайдаланылады [1]. Бұл проблема әсіресе үлкен көлемде ақпарат алмасатын компаниялар мен жеке тұлғалар үшін маңызды болып табылады.

Қазіргі уақытта спам хабарламаларды анықтау үшін қолданылатын классикалық әдістер олардың көлемі мен алуан түрлілігіне байланысты жеткіліксіз болып отыр. Қолданушылардың күн сайын алатын хабарламалар санының артуы осы мәселені автоматтандырылған тәсілдермен шешуді талап етеді. Бұл тұрғыда машиналық оқыту алгоритмдері тиімді шешім ұсынады. Машиналық оқыту негізіндегі әдістер деректердің үлгілерін талдау арқылы жаңа спам түрлерін тануға және бейімделуге мүмкіндік береді, осылайша спамды анықтау сапасын жақсартады.

Осы мақалада машиналық оқыту алгоритмдерін қолданып спам хабарламаларды анықтау тәсілдері талқыланады. Зерттеу барысында әртүрлі алгоритмдердің тиімділігі салыстырылып, олардың нақты қолдану салаларында артықшылықтары мен кемшіліктері қарастырылады [2]. Мақаланың мақсаты – заманауи машиналық оқыту әдістерін қолдану арқылы спам хабарламаларды анықтаудың тиімді жолдарын ұсыну және бұл саладағы жаңа мүмкіндіктерді ашу.

Зерттеудің маңыздылығы спам хабарламалармен күресудің тиімділігін арттырумен ғана шектелмейді, сонымен қатар ақпараттық қауіпсіздік деңгейін жоғарылатып, қолданушылардың жеке деректерін қорғауға бағытталған алдын алу шараларын күшейтуге де ықпал етеді [3].

Спам хабарламаларды анықтау мәселесі ақпараттық қауіпсіздік саласындағы өзекті зерттеу тақырыптарының бірі болып табылады. Бұл бағытта машиналық оқыту әдістерін қолдану тиімділігі кеңінен зерттелген және көптеген авторлар әртүрлі әдістерді қолдану арқылы спам хабарламаларды анықтау тәсілдерін қарастырған.

Y. Kontsewaya және басқалар [4] өз зерттеулерінде машиналық оқыту әдістерінің спам хабарламаларды анықтаудағы тиімділігін бағалаған. Олар әртүрлі алгоритмдердің нәтижелілігін салыстыра отырып, олардың дәлдігі мен өнімділігін зерттеді. Зерттеу нәтижелері көрсеткендей, әртүрлі машиналық оқыту модельдері әртүрлі деректер жиынтығына байланысты түрлі нәтижелер көрсете алады, бұл алгоритмдерді нақты қолдану салаларына бейімдеудің қажеттілігін айқындайды. S.D. Gupta және басқалар [5] SMS спамды анықтауда машиналық оқыту әдістерін қолдану арқылы деректерді өңдеу мен классификациялаудың тиімділігін зерттеді. Олардың жұмысы SMS хабарламаларындағы мәтінді талдау арқылы спам хабарламаларды анықтау мәселесін шешуге бағытталған. Олар әртүрлі классификаторларды, соның ішінде Наивті Байес, логистикалық регрессия және шешім ағаштарын қолдана отырып, олардың тиімділігін салыстырмалы түрде бағалады. M.R. Julis және S. Alagesan [6] мәтіндерді өңдеу арқылы машиналық оқыту әдістерінің көмегімен SMS спамды анықтау әдістерін ұсынды. Бұл зерттеу мәтіндік деректерді тиімді өңдеу және спамды анықтаудың нақтылығы мен жылдамдығын арттыру жолдарын ұсынды. Зерттеу нәтижелері көрсеткендей, мәтінді талдаудың сапасы спам хабарламаларды анықтаудың дәлдігіне елеулі әсер етеді. Teja Nallamothu мен Shais Khan [7] өздерінің жұмысында спамды анықтауға арналған машиналық оқыту алгоритмдерінің тиімділігін зерттеген. Олар машиналық оқыту әдістерін әртүрлі спам түрлеріне бейімдеуге болатындығын көрсетіп, әдістердің өнімділігін жақсарту үшін деректерді

алдын ала өңдеудің маңыздылығын атап өтті. Y. Guo және басқалар [8] екі бағытты трансформерлерді және машиналық оқыту классификаторларын қолдана отырып спамды анықтау әдістерін зерттеді. Зерттеу нәтижелері көрсеткендей, трансформерлерді қолдану спамды жоғары дәлдікпен анықтауға мүмкіндік берді, бұл машиналық оқыту модельдерінің дамуындағы маңызды қадам болып табылады.

S. Gibson және басқалар [9] био-инспирленген метаэвристикалық алгоритмдер арқылы оптимизацияланған машиналық оқыту әдістерін қолдану арқылы спамды анықтау тәсілдерін зерттеді. Бұл зерттеу спамды анықтауда машиналық оқыту алгоритмдерін тиімдірек ету үшін оптимизациялық әдістерді қолданудың артықшылықтарын көрсетті. M. Abdullahi және басқалар [10] кескіндер негізіндегі спам хаттарды анықтауға арналған машиналық оқыту әдістерін шолу жасаған. Олар кескіндердегі спамды анықтауда қолданылатын әдістерді және олардың тиімділігін талдап, әртүрлі әдістердің кемшіліктері мен артықшылықтарын көрсеткен. S. Gadde және басқалар [11] машиналық және терең оқыту әдістерін қолданып SMS спамды анықтау мәселесін шешу жолдарын ұсынды. Бұл зерттеу әртүрлі машиналық оқыту және терең оқыту әдістерінің тиімділігін салыстырып, олардың әртүрлі деректер жиынтықтарында қалай жұмыс істейтінін зерттеді. Q. Yaseen [12] терең оқыту әдістерін қолдана отырып, спам хаттарды анықтауға бағытталған зерттеу жүргізді. Оның жұмысы терең оқыту әдістерінің күрделі деректерді өңдеуге және спамды жоғары дәлдікпен анықтауға қабілетті екендігін көрсетті.

Осы зерттеулерден көрініп тұрғандай, спам хабарламаларды анықтауда машиналық оқыту мен терең оқыту әдістері үлкен мүмкіндіктерге ие. Әрбір әдіс өзінің ерекшеліктері мен артықшылықтарын көрсете отырып, нақты қолдану саласына байланысты түрлі нәтижелер береді. Сондықтан спамды анықтаудың ең тиімді жолын таңдау үшін деректердің түрі, мөлшері және қолдану мақсаты ескерілуі қажет.

Зерттеу шарттары мен әдістері. Бұл зерттеуде SMS Spam Collection Dataset деректер жиынтығы пайдаланылды. Аталған деректер жиынтығы SMS хабарламаларындағы спамды анықтау бойынша зерттеулер жүргізу мақсатында жиналған және тегтелген 5,574 хабарламадан тұрады. Әрбір хабарлама екі бағаннан тұрады: біріншісі (v1) хабарламаның меткасын (спам немесе заңды), ал екіншісі (v2) хабарламаның мәтінін көрсетеді. Хабарламалар жинағы спам (403 хабарлама) және заңды (ham) хабарламалар (5,171 хабарлама) болып бөлінген.

Деректер көзі. SMS Spam Collection Dataset бірнеше дереккөзден жиналған:

– Grumbletext веб-сайтынан 425 спам хабарлама қолмен алынды. Бұл Ұлыбританияның ұялы телефон пайдаланушылары SMS спамға қатысты шағымдарын жариялайтын форумы. Спам хабарламаларды анықтау өте қиын және уақытты қажет ететін тапсырма болды, өйткені жүздеген веб-беттерді мұқият қарап шығу керек болды [13];

– NUS SMS Corpus (NSC) мәліметтер жинағынан 3,375 заңды SMS кездейсоқ таңдалды. Бұл деректер жинағы шамамен 10,000 заңды хабарламалардан тұрады және Сингапур Ұлттық университетінің информатика кафедрасында зерттеу жүргізу үшін жиналған;

– Caroline Tag's PhD диссертациясынан 450 заңды SMS хабарлама қосылды;

– SMS Spam Corpus v.0.1 Big құрамына 1,002 заңды және 322 спам хабарлама кірді.

Алдын ала өңдеу. Зерттеу барысында хабарламаларды тиімді талдау үшін алдын ала өңдеу кезеңі жүргізілді:

1. Төменгі регистрге ауыстыру: Барлық мәтін төменгі регистрге келтірілді, бұл сөздерді бірегей түрде талдауды жеңілдетті [14];

2. Таңбаларды жою: Қаріптік таңбалар, нүктелер, тыныс белгілері және басқа арнайы таңбалар жойылды;

3. Сөздерді түбірге келтіру (stemming) және қосымшаларды жою (lemmatization) әдістері қолданылды. Бұл сөздердің түбірлерін анықтауға және сөз нұсқаларының санын азайтуға мүмкіндік берді.

Машиналық оқыту әдістері. Бұл зерттеуде әртүрлі машиналық оқыту алгоритмдері қолданылды және олардың спам хабарламаларды анықтаудағы тиімділігі бағаланды. Қолданылған негізгі алгоритмдер мыналар:

1. Наивті Байес классификаторы (Naive Bayes): Бұл әдіс шартты тәуелсіздік принципіне негізделген ықтималдықтық әдіс. Наивті Байес классификаторы мәтіндік деректерді өңдеуде және спам хабарламаларды анықтауда кеңінен қолданылады;

2. Логистикалық регрессия (Logistic Regression): Бинарлы классификация үшін қолданылатын әдіс. Логистикалық регрессия спам және заңды хабарламалар арасындағы айырмашылықты анықтауға мүмкіндік береді [15];

3. Қолдау векторлық машиналары (Support Vector Machines, SVM): Бұл әдіс деректерді сызықтық және сызықтық емес бөлетін гипержазықтықты құруға бағытталған. SVM әдісі күрделі деректер құрылымдарымен тиімді жұмыс істейді;

4. Шешім ағаштары (Decision Trees): Деректерді олардың қасиеттеріне байланысты бөлетін иерархиялық құрылым. Бұл әдіс шешім қабылдау ережелерін түсіндіруге жеңіл болғандықтан, спамды анықтауда қолданылады;

5. Рэндом Форест (Random Forest): Бұл бірнеше шешім ағаштарын құру арқылы нәтижені жақсартатын ансамбльдік әдіс. Әр түрлі ағаштардың шешімдерін біріктіру қателіктерді азайтуға және жіктеудің дәлдігін арттыруға мүмкіндік береді;

6. Градиентті Бустинг (Gradient Boosting): Бұл әдіс ағаш құрылымды классификаторларды қайталап қолдану арқылы қателіктерді біртіндеп азайтуды мақсат етеді. Градиентті бустинг күрделі спам үлгілерін тиімді анықтауға мүмкіндік береді [16].

Алгоритмдерді бағалау. Әрбір машиналық оқыту алгоритмі төмендегі бағалау метрикалары бойынша бағаланды:

– Нақтылық (Accuracy): Спам және заңды хабарламаларды дұрыс анықтау пайызы;

– Дәлдік (Precision): Дұрыс анықталған спам хабарламалардың жалпы анықталған спам хабарламаларға қатынасы;

– Еске түсіру (Recall): Дұрыс анықталған спам хабарламалардың жалпы шынайы спам хабарламаларға қатынасы;

– F1-өлшемі (F1-score): Дәлдік пен еске түсіру көрсеткіштерінің орташа гармониясы.

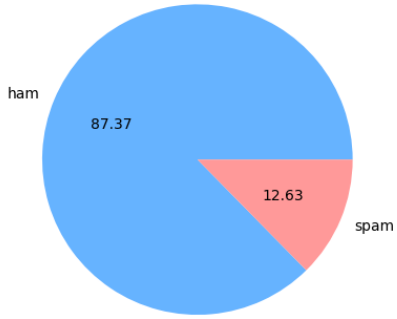
Зерттеу нәтижелеріне сәйкес, әрбір алгоритмнің нәтижелілігі деректердің сипатына және алдын ала өңдеу әдістеріне байланысты өзгерді. Алынған нәтижелер алдағы уақытта спам хабарламаларды анықтаудың тиімді

әдістерін таңдауға және машиналық оқыту модельдерін жақсартуға мүмкіндік береді [17].

Зерттеу нәтижелері және оларды талқылау. Бұл зерттеу спам хабарламаларды анықтау мақсатында әртүрлі мәтіндік деректерді талдау және оларды өңдеу әдістерін қолдануға бағытталды. Осы бөлімде қолданылған әдістердің нәтижелері мен олардан алынған қорытындылар сипатталады.

1. Деректерді зерттеуге шолу.

Деректер жиыны 5169 мәтіндік хабарламадан тұрады, олардың ішінде 4516-сы «ham» (яғни, спам емес) және 653-і спам ретінде таңбаланған. Бұл деректердің бастапқы зерттеулері арқылы спам және «ham» хабарламалардың ерекшеліктерін анықтау мақсатында әртүрлі визуализация әдістері қолданылды. Нәтижелер көрсеткендей, спам хабарламалар жалпы хабарламалар санының шамамен 12,6%-ын құрайды, ал қалған 87,4% «ham» санатына жатады.



Сурет 1. Ham және спам хабарламалардың үлестірілімі

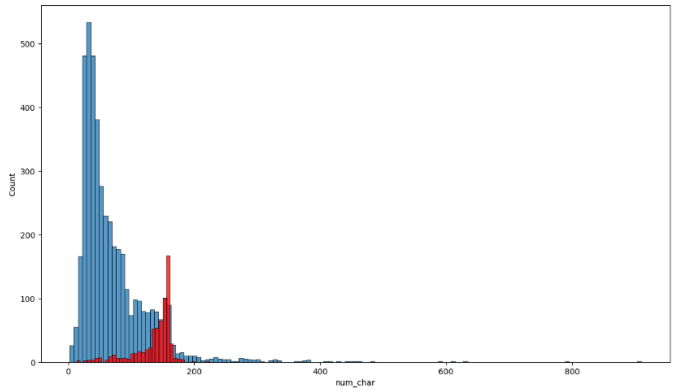
Бұл суретте «ham» және спам хабарламалардың жалпы саны көрсетілген, мұнда «ham» хабарламалардың басым екендігі көрінеді. Бұл ақпарат модельді құру барысында теңсіздік мәселесін шешуді талап етеді, себебі деректердің теңгерімсіздігі алгоритмдердің нәтижелеріне айтарлықтай әсер етуі мүмкін.

2. Деректерді визуализациялау.

Хабарламалардың ұзындығы мен олардың сипаты бойынша талдау жүргізілді. Бұл талдауда «ham» және спам хабарламалардың ұзындығы бойынша ерекшеліктері айқындалды. 2-суретте хабарламалардағы символдар санының «ham» және спам хабарламалар үшін таралуы көрсетілген.

Спам хабарламалар көбіне коммерциялық немесе жарнамалық сипатқа ие болғандықтан, олар «ham» хабарламаларға қарағанда орташа ұзын болады. Бұл факт спам хабарламалардың мәтіндерінде жиі кездесетін түрлі ұсыныстар, акциялар, және жарнамалық мазмұнның болуының нәтижесі ретінде түсіндіріледі.

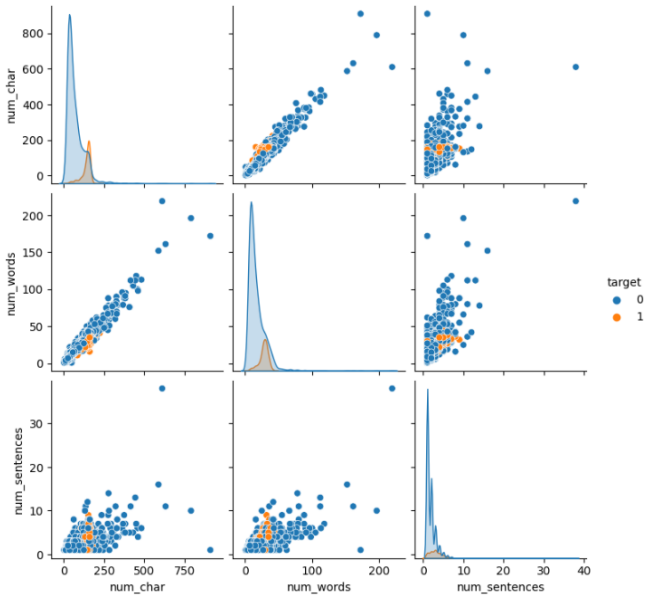
Бұл диаграммада «ham» хабарламалардың орташа ұзындығы шамамен 50-70 символ болса, спам хабарламаларда бұл көрсеткіш 100 және одан да көп символды құрайды. Осы ұзындықтың үлкен болуы спам хабарламалардың жиі жарнамалық сипатта болуына байланысты.



Сурет 2. Хабарламалардың ұзындығы бойынша үлестірілім

3. Жұп диаграмма анализі.

Жұп диаграммалар (pair plot) көмегімен хабарламалардағы негізгі көрсеткіштер арасындағы байланысты анықтау мақсатында талдау жүргізілді (3-сурет). Бұл талдау хабарламаның ұзындығы, сөздер саны, сөйлем саны сияқты көрсеткіштердің спам және «ham» хабарламалар үшін әртүрлі болатынын көрсетеді.



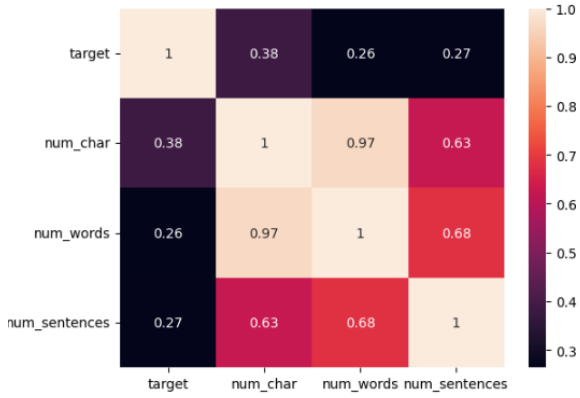
Сурет 3. Жұп диаграммалары арқылы хабарламалар арасындағы байланыстар

Бұл диаграммалар спам және «ham» хабарламалар арасындағы сипаттамалық ерекшеліктерді анықтауға көмектесті. Спам хабарламаларда көбінесе сөйлем саны көп, әрі сөздер саны да жоғары. Бұл ақпарат модель құру барысында ерекшелік белгілерін (features) таңдау кезінде ескерілуі керек.

4. Корреляция матрицасы.

4-суретте әртүрлі көрсеткіштер арасындағы корреляциялық қатынасты анықтайтын корреляция матрицасы көрсетілген. Мақсаты – спам

хабарламаларға тән сипаттамалар арасындағы өзара байланыс деңгейін анықтау.



Сурет 4. Корреляция матрицасы

Бұл матрицада хабарламаның ұзындығы мен сөздер санының арасында оң корреляция бар екенін байқауға болады. Басқаша айтқанда, хабарламада неғұрлым көп сөз болса, оның ұзындығы да соғұрлым жоғары болады. Дегенмен, басқа көрсеткіштер арасындағы корреляция салыстырмалы түрде әлсіз болып шықты, бұл олардың модельдеудегі әсерін шектеулі ететінін көрсетеді.

5. Деректерді алдын ала өңдеу.

Хабарламаларды тиімді талдау және өңдеу үшін мәтіндік деректерді алдын ала өңдеу қадамдары жүзеге асырылды. Төменде қолданылған өңдеу әдістері келтірілген:

1. Барлық мәтінді кіші әріптерге ауыстыру – бұл қадам мәтіндер арасындағы сәйкестікті қамтамасыз ету үшін қолданылды;
2. Токенизация – мәтінді жеке сөздерге бөлу арқылы деректердің құрылымын оңтайландырады;
3. Арнайы символдарды жою – мағынаға қатысы жоқ символдарды алып тастау мәтіннің сапасын арттыруға көмектеседі;
4. Стоп сөздер мен тыныс белгілерді алып тастау – бұл қадам маңызды емес сөздер мен тыныс белгілерін сүзу арқылы модельдің тиімділігін арттырады;
5. Сөздерді түбірге келтіру – сөздерді түбір формасына келтіру арқылы бірдей мағыналы сөздерді стандарттау, яғни бір сөздің әртүрлі формаларын бір мағынаға біріктіру.

Бұл қадамдардың барлығы мәтіндерді ықшамдап, маңызды ерекшеліктерді сақтауға бағытталған.

Бірнеше мысал келтірілгенде, хабарламалардың түпнұсқалық түрі мен алдын ала өңделген нұсқаларының арасында айқын айырмашылықтар байқалады. Мысалы, бастапқыда:

Бастапқы мәтін: «Congrats! Nokia 3650 video camera phone is your Call 09066382422 Calls cost 150ppm Ave call 3mins vary from mobiles 16+ Close 300603 post BCM4284 Ldn WC1N3XX».

Өңделген мәтін: «congrat nokia 3650 video camera phone call 09066382422 call cost 150ppm ave call 3min vari mobil close 300603 post bcm4284 ldn wc1n3xx».

Бұл мысал көрсеткендей, арнайы символдар мен жиі қолданылатын сөздер жойылып, мәтін ықшамдалған және мағыналық тұрғыдан маңызды бөліктер сақталған.

6. Модельдерді оқыту және бағалау.

Өңделген деректер негізінде түрлі машиналық оқыту модельдері құрастырылып, олардың тиімділігі өлшенді (1-кесте). Модельдердің жұмысын бағалау үшін стандартты көрсеткіштер – дәлдік (accuracy), сәйкестік (precision), шақыру (recall), және F1 көрсеткіші (F1-score) пайдаланылды.

Кесте 1

Әр түрлі машиналық оқыту модельдерінің классификация тиімділігінің көрсеткіштері

Модель	Нақтылық (%)	Дәлдік	Еске түсіру	F1-өлшем
Наивті Байес	97,27	0,98	0,89	0,93
Логистикалық регрессия	98,01	0,96	0,95	0,95
SVM	98,76	0,97	0,97	0,97
Шешім ағашы	95,35	0,93	0,93	0,93
Рэндом Форест	97,91	0,96	0,96	0,96
Градиентті Бустинг	97,38	0,94	0,94	0,94

Бұл модельдердің нәтижелері көрсеткендей, спам анықтау үшін Random Forest алгоритмі жоғары дәлдікке ие, бірақ әр модельдің жеке артықшылықтары мен кемшіліктері бар екенін де ескерген жөн.

Қорытынды. Бұл зерттеудің негізгі мақсаты – әр түрлі машиналық оқыту модельдерін қолдану арқылы мәтіндік хабарламалардың спам екенін анықтауға бағытталған тиімді әдістерді зерттеу. Жүргізілген тәжірибелер мен талдаулардың нәтижесінде алынған қорытындылар, модельдердің ерекшеліктері мен оларды оңтайландыру жолдарын көрсетуге мүмкіндік береді.

Деректер жиынында 5169 мәтіндік хабарлама қамтылды, олардың ішінде тек 653 хабарлама ғана спам ретінде анықталды. Алғашқы талдаулардан хабарламалардың ұзындығы, сөздер саны және сөйлем саны сияқты сипаттамалар бойынша ерекшеліктер байқалды. Бұл ерекшеліктер спам және «ham» хабарламаларды ажыратуда маңызды рөл атқаратын факторлар болып шықты. Осылайша, хабарламалардың ұзындығы мен сөздер саны бойынша байқалған айырмашылықтар классификация үшін маңыздылығын көрсетті.

Деректерді алдын ала өңдеу қадамдары, соның ішінде мәтінді төменгі регистрге ауыстыру, токенизация, арнайы символдар мен жиі қолданылатын сөздерді жою, сондай-ақ сөздерді түбірге келтіру, хабарламаларды ықшамдап, мағынасына қарай маңызды бөліктерді ғана қалдыруға мүмкіндік берді. Бұл қадамдар модельдерді оқыту барысында қолданылатын ерекшеліктердің сапасын арттыруға және модельдердің жалпы өнімділігін жақсартуға ықпал етті. Мысалы, бастапқыда мәтіндер әртүрлі тыныс белгілері мен формаларда болса, өңдеуден кейін олар стандартты пішімге келтірілді, бұл өз кезегінде модельдердің оқу процесін оңтайландырды.

Жүргізілген тәжірибелер барысында бірнеше танымал машиналық оқыту модельдері қолданылды, олардың ішінде Наивті Байес, Логистикалық регрессия, SVM, Шешім ағашы, Рэндом Форест және Градиентті Бустинг

сияқты алгоритмдер бар. 1-кестеде берілген нәтижелер бұл модельдердің әрқайсысының тиімділігін салыстыруға мүмкіндік берді.

SVM алгоритмі барлық көрсеткіштер бойынша жоғары нәтижелер көрсетіп, нақты спам және «ham» хабарламаларды анықтау үшін ең тиімді модель болып табылды. Бұл модельдің жоғары нақтылығы, дәлдігі, еске түсіру және F1-өлшемі оның тұрақты және тиімді жұмысының дәлелі. Сонымен қатар, Логистикалық регрессия мен Рэндом Форест модельдері де жоғары нәтиже көрсетіп, спам классификацияда сенімді балама ретінде пайдалануға болатындығын көрсетті.

Бұл зерттеу нәтижелері мәтіндік деректер негізінде спам хабарламаларды дәл анықтау үшін әртүрлі машиналық оқыту модельдерінің қолданылу мүмкіндіктерін көрсетеді. Әрбір модельдің артықшылықтары мен кемшіліктері бар, мысалы, SVM алгоритмі жоғары дәлдік көрсетсе де, есептеу күрделілігі жоғары болып келуі мүмкін. Шешім ағашы салыстырмалы түрде қарапайым және тез нәтиже береді, бірақ оның нақтылығы төмен болуы мүмкін.

Келешектегі зерттеулерде модельдердің өнімділігін одан әрі арттыру мақсатында гиперпараметрлерді оңтайландыру әдістерін қолдану, мәтіндік деректердің жаңа ерекшеліктерін қосу және ансамбльдік әдістерді (мысалы, бірнеше модельдерді біріктіру) қарастыру қажет болады. Сондай-ақ, деректердің теңгерімсіздігін шешу үшін SMOTE сияқты техникаларды пайдалану арқылы өнімділікті жақсарту мүмкіндігін зерттеу маңызды.

Бұл зерттеудің нәтижелері коммерциялық ұйымдар мен жеке қолданушыларға спам хабарламаларды тиімді түрде анықтап, қажетсіз ақпараттан қорғауға көмектеседі. Мұндай шешімдер клиенттердің уақытын үнемдеп, жұмыс тиімділігін арттырады, сондай-ақ алаяқтық әрекеттердің алдын алуға ықпал етеді. Осы зерттеуде қарастырылған модельдер мен әдістер нақты жүйелерге енгізіліп, электрондық пошта, әлеуметтік желілер немесе SMS арқылы таратылатын спам хабарламаларды сүзу үшін пайдаланылуы мүмкін.

Бұл зерттеу мәтіндік хабарламаларды спам ретінде тиімді түрде анықтаудың әртүрлі әдістерін зерттеуге бағытталған. Әртүрлі машиналық оқыту модельдерін қолдану арқылы алынған нәтижелер олардың әрқайсысының тиімділігі мен шектеулерін анықтауға мүмкіндік берді. Бұл зерттеу қолданылатын модельдерді жетілдіру және деректерді алдын ала өңдеу әдістерін жақсарту арқылы болашақта спам классификациясында жоғары дәлдікке қол жеткізуге көмектеседі.

Әдебиеттер тізімі

1. Kumar N. et al. Email spam detection using machine learning algorithms //2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA). – IEEE, 2020. – P. 108-113.
2. GuangJun L. et al. Spam detection approach for secure mobile message communication using machine learning algorithms //Security and Communication Networks. – 2020. – Vol. 2020. – No. 1. – P. 8873639.
3. Siddique Z.B. et al. Machine Learning-Based Detection of Spam Emails //Scientific Programming. – 2021. – Vol. 2021. – No. 1. – P. 6508784.
4. Kontsewaya Y., Antonov E., Artamonov A. Evaluating the effectiveness of machine learning methods for spam detection //Procedia Computer Science. – 2021. – Vol. 190. – P. 479-486.

5. Gupta S.D., Saha S., Das S.K. SMS spam detection using machine learning //In Journal of Physics: Conference Series. – IOP Publishing, 2021. – Vol. 1797. – No. 1. – P. 012017.
6. Julis M. R., Alagesan S. Spam detection in SMS using machine learning through textmining //Internatıonal Journal Of Scıentıfic & Technology Research. – 2020. – Vol. 9. – No. 02.
7. Teja Nallamothu P., Shais Khan M. Machine learning for SPAM detection //Asian Journal of Advances in Research. – 2023. – Vol. 6. – No. 1. – P. 167-179.
8. Guo Y., Mustafaoglu Z., Koundal D. Spam detection using bidirectional transformers and machine learning classifier algorithms //journal of Computational and Cognitive Engineering. – 2023. – Vol. 2. – No. 1. – P. 5-9.
9. Gibson S. et al. Detecting spam email with machine learning optimized with bio-inspired metaheuristic algorithms //Ieee Access. – 2020. – Vol. 8. – P. 187914-187932.
10. Abdullahi M. et al. A Review on Machine Learning Techniques for Image Based Spam Emails Detection //2020 IEEE 2nd International Conference on Cyberspac (CYBER NIGERIA). – IEEE, 2021. – P. 59-65.
11. Gadde S., Lakshmanarao A., Satyanarayana S. SMS spam detection using machine learning and deep learning techniques //2021 7th international conference on advanced computing and communication systems (ICACCS). – IEEE, 2021. – Vol. 1. – P. 358-362.
12. Yaseen Q. et al. Spam email detection using deep learning techniques //Procedia Computer Science. – 2021. – Vol. 184. – P. 853-858.
13. Sethi M. et al. Spam email detection using machine learning and neural networks //Sentimental Analysis and Deep Learning: Proceedings of ICSADL 2021. – Springer Singapore, 2022. – P. 275-290.
14. Srinivasan S. et al. Spam emails detection based on distributed word embedding with deep learning //Machine intelligence and big data analytics for cybersecurity applications. – 2021. – P. 161-189.
15. Mashaleh A. S. et al. Detecting spam email with machine learning optimized with Harris Hawks optimizer (HHO) algorithm //Procedia Computer Science. – 2022. – Vol. 201. – P. 659-664.
16. Mashaleh A. S. et al. Detecting spam email with machine learning optimized with Harris Hawks optimizer (HHO) algorithm //Procedia Computer Science. – 2022. – Vol. 201. – P. 659-664.
17. Kaddoura S., Alfandi O., Dahmani N. A spam email detection mechanism for English language text emails using deep learning approach //2020 IEEE 29th international conference on enabling technologies: infrastructure for collaborative enterprises (WETICE). – IEEE, 2020. – P. 193-198.

Материал редакцияға 15.10.24 түсті.

А.Б. Абен[1], Н.М. Жунисов[1], Ж.Б. Мырзатаев[1]

[1]Международный казахско-турецкий университет им. Ходжи Ахмеда Ясави, г. Туркестан, Казахстан

ОБНАРУЖЕНИЕ СПАМ-СООБЩЕНИЙ С ИСПОЛЬЗОВАНИЕМ АЛГОРИТМОВ МАШИННОГО ОБУЧЕНИЯ

Аннотация. Данное исследование анализирует методы обнаружения спама с помощью автоматической классификации текстовых сообщений. Целью исследования было оценить эффективность моделей машинного обучения и исследовать их применимость в системах фильтрации спама. Этапы предварительной обработки включали преобразование сообщений в нижний регистр, токенизацию, удаление специальных символов и знаков препинания, исключение распространенных слов и стемминг. Эти шаги помогли выделить важную информацию из сообщений и улучшить качество классификации.

В ходе исследования были протестированы такие модели, как Наивный Байес, Логистическая регрессия, SVM, Дерево решений, Случайный лес и Градиентный бустинг. Результаты показали, что модель SVM достигла наивысших показателей по всем метрикам, в то время как модели Случайного леса и Логистической регрессии также продемонстрировали высокую эффективность. Эти модели достигли более 95% в точности, точности, полноте и F1-мере.

Результаты исследования демонстрируют эффективность методов машинного обучения в классификации текстовых данных и указывают на их потенциальное применение в реальных системах, таких как фильтрация сообщений, получаемых по электронной почте или SMS. В будущем планируется оптимизация гиперпараметров и применение усовершенствованных методов обработки текстовых данных для повышения производительности моделей.

Ключевые слова: машинное обучение, фильтрация спама, классификация текстов, Наивный Байес, логистическая регрессия, модель SVM, предварительная обработка данных, градиентный бустинг.

A.B. Aben[1], N.M. Zhunissov[1], Zh,B. Myrzatayev[1]

[1]Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan

DETECTION OF SPAM MESSAGES USING MACHINE LEARNING ALGORITHMS

Abstract. This study analyzes methods for detecting spam through the automatic classification of text messages. The aim of the research was to evaluate the effectiveness of machine learning models and to explore their applicability in spam filtering systems. Preprocessing steps included converting messages to lowercase, tokenization, removing special characters and punctuation, eliminating common words, and stemming. These steps helped to highlight important information from the messages and enhance the quality of classification.

During the study, models such as Naive Bayes, Logistic Regression, SVM, Decision Tree, Random Forest, and Gradient Boosting were tested. The results showed that the SVM model achieved the highest metrics across all measures, while the Random Forest and Logistic Regression models also demonstrated high effectiveness. These models attained over 95% in accuracy, precision, recall, and F1-score.

The findings of the research demonstrate the effectiveness of machine learning techniques in classifying textual data and indicate their potential application in real-world systems, such as filtering messages received via email or SMS. Future work aims to optimize hyperparameters and apply advanced methods for processing textual data to enhance the performance of the models.

Keywords: machine learning, spam filtering, text classification, Naive Bayes, logistic regression, SVM model, data preprocessing, gradient boosting.