



Kauno technologijos universitetas

Informatikos fakultetas

Žalingo turinio atpažinimo metodas

Baigiamasis magistro projektas

Informacijos ir informacinių technologijų sauga (6211BX008)

Mantas Pečiulis

Projekto autorius

Prof. dr. Agnius Liutkevičius

Vadovas

Prof. Jevgenijus Toldinas

Recenzentas

Kaunas, 2022



Kauno technologijos universitetas

Informatikos fakultetas

Mantas Pečiulis

Žalingo turinio atpažinimo metodas

Akademinio sąžiningumo deklaracija

Patvirtinu, kad mano, Manto Pečiulio, baigiamasis projektas tema „Žalingo turinio atpažinimo metodas“ yra parašytas visiškai savarankiškai ir visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

(vardą ir pavardę įrašyti ranka)

(parašas)

Mantas Pečiulis. Žalingo turinio atpažinimo metodas. Magistro baigiamasis projektas / vadovas prof. dr. Agnius Liutkevičius; Kauno technologijos universitetas, Informatikos fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Informatikos inžinerija (Informatikos mokslai).

Reikšminiai žodžiai: žalingo turinio aptikimas, teksto klasifikacija, žalingas turinys, prižiūrimas mokymasis.

Kaunas, 2022. 58 p.

Santrauka

Įvairios technologijos leidžia pasiekti informaciją ar ją dalintis. Dauguma nepagalvoja, kad veikla internete gali būti ne tik kūrybinga, edukacinė, bet ir destruktivi. Žalingo pobūdžio puslapiai tampa vis lengviau pasiekiami, nes bando maskuoti, klaidinti blokavimą ir jį apeiti, o visa tai paveikia saugų naršymą internete. Dauguma tinklapių filtravimo sistemų pagrįstos juodaisiais sąrašais ar raktinių žodžių mechanizmais, neatsižvelgiant į atidaryto puslapio turinį, kas ne visada užtikrina turinio identifikavimą puslapyje nustatant ar jis – žalingas ar kitokio pobūdžio.

Magistriniame darbe realizuotas prototipas, kuris nustato įvesto puslapio kategoriją. Sudarytas duomenų rinkinys pagal sudarytą metodologiją. Parinktas mašininio mokymo algoritmas ir palygintas su kitais algoritmais naudojant sudarytą duomenų rinkinį. Duomenų rinkinys sudarytas iš atrinktų puslapių turinio, kuris buvo normalizuotas ir išskirtos savybės naudojant vektorizavimo algoritmą. Atliktas apmokyto algoritmo tyrimas ir palygintas su kitais algoritmais.

Mantas Pečiulis. Harmful Content Detection Method. Master's thesis in information and information system security / Supervisor prof. dr. Agnius Liutkevičius; Kaunas University of Technology.

Study field and area (study field group): Informatics Engineering (Computing).

Keywords: harmful content detection, classification of text, harmful content, supervised learning. Kaunas, 2022. 58 p.

Summary

Different technologies make it possible to access or share information. What most people do not realize is that online activities can be creative and educational and destructive. Harmful sites are becoming increasingly easy to access as they try to mask or mislead and circumvent blocking, all of which affect safe web browsing. Most web filtering systems are based on blacklists or keyword mechanisms that do not consider the content of the opened page, which does not always ensure that the content on the page can be identified as harmful or otherwise.

In the master thesis, a prototype has been implemented that determines the category of the entered page. A dataset is built according to the methodology developed. A machine learning algorithm was selected and compared with other algorithms using the constructed dataset. The dataset consists of the content of the selected pages, which has been normalized and features extracted using a vectorization algorithm. The trained algorithm was tested and compared with other algorithms.

Turinys

Paveikslų sąrašas	7
Santrumpų ir terminų sąrašas	8
Įvadas.....	9
1. Žalingo turinio internete analizė	10
1.1. Žalinga informacija internete	10
1.2. Žalingo turinio internete vaikams analizė	11
1.3. Apsaugos nuo žalingo turinio metodai ir priemonės	11
1.4. Tėvų kontrolės sistemos ir būdai apsaugoti vaikus internete naudojant filtrus.....	12
1.4.1. Naršyklės filtrai	12
1.4.2. Paieškos sistemų filtrai	13
1.4.3. Interneto paslaugų tiekėjo ribojimai.....	13
1.5. Esamų tėvų kontrolės sprendimų analizė	14
1.5.1. Net Nanny.....	14
1.5.2. Qustodio	14
1.5.3. SafeDNS	15
1.6. Esamų tėvų kontrolės sprendimų analizės išvados.....	15
1.7. Internetinių tinklapių filtravimas naudojant domenų vardų sistemą	16
1.7.1. Domenų vardų sistemos mechanizmas.....	16
1.7.2. Domenų vardų sistemos filtravimo galimos architektūros analizė.....	16
1.8. Turinio filtravimas naudojant teksto klasifikacijos metodus.....	18
1.8.1. Naivus Bajesas metodas	19
1.8.2. K – artimiausių kaimynų metodas	19
1.8.3. Atraminių vektorių metodas	20
1.8.4. Neuroninio tinklo metodas	20
1.8.5. Klasifikacijos metodų panaudojimas filtruojant turinį	21
1.9. Teksto klasifikacijos metodų panaudojimas tyrimuose.....	24
1.10. Teksto klasifikacijos metodų išvada.....	24
1.11. Teksto klasifikacijos TF-IDF ir skaičiavimo vektorizavimas požymių išskyrimui	24
1.12. Netinkamo turinio filtravimo iššūkiai	26
1.12.1. Filtrų apėjimas	26
1.13. Išvada.....	26
2. Žalingo turinio atpažinimo metodas	28
2.1. Žalingo turinio atpažinimo metodo koncepcija	28
2.1.1. Žalingo turinio metodo reikalingumo pagrindimas	29
2.2. Žalingo turinio metodo reikalavimų specifikacija.....	29
2.2.1. Funkciniai reikalavimai	32
2.2.2. Nefunkciniai reikalavimai	33
2.3. Žalingo turinio aptikimas puslapiuose.....	33
2.3.1. Netinkamų žodžių atpažinimas tinklapiuose	33
2.3.2. Tinklapių klasifikavimas pagal sudarytą duomenų modelį	33
2.3.3. Žalingo turinio metodo apmokymų duomenų rinkinys.....	35
2.3.4. Žalingo turinio metodo duomenų rinkinio sudarymo metodas	35

3. Žalingo turinio atpažinimo metodo realizacija.....	38
3.1. Žalingo turinio atpažinimo metodo priemonės.....	38
3.2. Išankstinis duomenų apdorojimas prototipe	38
3.2.1. Dažniausiai naudojamų žodžių pašalinimas ir teksto normalizavimas	40
3.3. Žalingo turinio duomenų rinkinio apdorojimas prototipe	42
3.3.1. Duomenų rinkinio mažinimas	44
3.3.2. Žalingo turinio aptikimo prototipo diegimas.....	44
3.3.3. Žalingo turinio atpažinimo prototipo realizacijos išvados	45
3.4. Žalingo turinio atpažinimo tyrimas	45
3.4.1. Žalingo turinio atpažinimo metodo klasifikacijos žingsniai ir numatytieji parametrai.....	45
3.4.2. Sudarytos n-gramos ir b-gramos	47
3.4.3. Algoritmų įvertinimai.....	47
3.4.4. Išanalizuotų algoritmų palyginimas su surinktu duomenų rinkiniu	48
3.5. Tyrimo išvados	54
3.6. Išvalgos ateičiai	54
Išvados ir rezultatai.....	55
Literatūros sąrašas	56

Paveikslų sąrašas

1 pav.	„Bing“ saugios paieškos nustatymai.	13
2 pav.	„Google“ saugios paieškos nustatymai.	13
3 pav.	Domenų vardų sistemos mechanizmas [20].	16
4 pav.	Domenų vardų sistemos architektūra [21].	17
5 pav.	k – artimiausio kaimyno klasifikavimo pavyzdys [24].	19
6 pav.	SVM metodo veikimo pavyzdys [23].	20
7 pav.	Neuroninio tinklo mokymo pavyzdys [25].	21
8 pav.	„Fuzzy ontologijos“ siūlomas architektūros sprendimas [27].	22
9 pav.	Žalingo turinio atpažinimo metodo koncepcija.	28
10 pav.	Žalingo turinio atpažinimo internete sistemos prototipo panaudos atvejų diagrama.	29
11 pav.	Tinklapio turinio identifikavimo procesas.	34
12 pav.	Tinklapio HTML vaizdas.	34
13 pav.	„Google“ raktinių žodžių planavimo įrankis.	35
14 pav.	Duomenų rinkinio pasiskirstymas pagal kategorijas.	36
15 pav.	Veiksmų seka iš ankstinio duomenų paruošimo	37
16 pav.	Išankstinis tinklapio turinio apdorojimo procesas.	38
17 pav.	Duomenų rinkinio apdorojimas naudojant TF-IDF	43
18 pav.	Žalingo turinio aptikimo prototipo diegimo diagrama.	45
19 pav.	Žalingo turinio aptikimo metodo klasifikacijos apmokymo žingsniai.	46
20 pav.	Klasifikatoriaus kryžminė validacija [37].	47
21 pav.	Atraminio vektoriaus algoritmo išmaišymo matrica.	50
22 pav.	Atraminio vektoriaus ROC kreivė.	50
23 pav.	Multinominio Naivaus Bajeso išmaišymo matrica.	51
24 pav.	Multinominio Naivaus Bajeso išmaišymo matrica.	51
25 pav.	Gausinio Naivaus Bajeso išmaišymo matrica.	52
26 pav.	Gausinio Naivaus Bajeso ROC kreivė.	53
27 pav.	Atsitiktinio medžio išmaišymo matrica.	53
28 pav.	Atsitiktinio medžio ROC kreivė.	54

Santrumpų ir terminų sąrašas

Dažniausiai pasikartojantys žodžiai (angl. stop words) – žodžiai neturintys įtakos ar reikšmės tekste.

DNS (angl. Domain name system) – domenų vardų sistema.

IPT (angl. Internet Service Provider) – interneto paslaugų tiekėjas, juridinis asmuo teikiantis informacijos perdavimo viešo naudojimo kompiuterių tinklais arba šių tinklų paslaugas.

Lemavimas – žodžio pavertimas į bendrinę formą.

KNN (angl. k-nearest neighbour) – k- artimiausias kaimynas.

NLTK – Python natūralios kalbos biblioteka.

N – grama – sudarytas žodžių junginys iš apibrėžto N žodžių.

ODP (angl. Open Directory Project) – aktyvios direktorijos projektas.

Ontologija – dalykinės srities formali specifikacija.

SVM (angl. Support Vector Machine) – atraminių vektorių metodas.

TF-IDF – terminų dažnis, atvirkštinis dokumentų dažnis.

URL – svetainės (puslapio) adresas.

Vektorizavimas – tam tikros savybės pavertimas skaitine reikšme.

Įvadas

Darbo problematika ir aktualumas

Tobulėjant technologijoms įvairūs įrenginiai ir žaidimų konsolės prijungiamos prie namų tinklo, kurios leidžia namų vartotojams naudojant technologijas žaisti žaidimus, pasiekti informaciją ar ją dalintis, susisiekti su artimaisiais ar bendraamžiais. Veikla internete gali būti kūrybinga, edukacinė, bet kartais destruktvyvi. Kaip pavyzdys destruktvyvios veiklos internete galėtų būti azartiniai žaidimai, prekyba narkotikais, seksualinis priekabiavimas ir rasizmas. Naudojimas interneto blogais tikslais arba nežinant kas yra daroma gali privesti prie blogos kultūros, psichologinių traumų. Taip pat internetas turi ir tamsiąją pusę, kuri užpildyta negatyvios informacijos ir netinkamo turinio, dezinformacijos, kuri gali suformuoti netinkamai vaikų ir suaugusių žmonių pasaulėžiūrą. Žmogaus protas kartais nespėja prisitaikyti prie visų naujų technologijų, todėl šiuo atveju vaikams yra būtina tėvų priežiūra internete.

Darbo tikslas ir uždaviniai

Šio darbo tikslas sukurti žalingo turinio atpažinimo metodą. Atpažinimo metodas leistų identifikuoti ir nustatyti atidaryto puslapio kategoriją pagal jame esantį turinį. Norint įgyvendinti šį tikslą reikia išspręsti šiuos uždavinius:

- Išanalizuoti esamas apsaugas nuo žalingo turinio internete problemas.
- Pasiūlyti veiksmingą žalingo turinio atpažinimo metodą.
- Suprojektuoti ir įgyvendinti pasiūlyto apsaugos metodo prototipo programinę realizaciją.
- Eksperimentiškai įvertinti siūlomo apsaugos metodo prototipinės realizacijos efektyvumą ir palyginti su kitais metodais.

Darbo rezultatai ir jų svarba

Darbe siūlomas metodas žalingo turinio aptikimui internete. Šio metodo veikimo principas – atidarytą puslapio turinį išanalizuoja pagal jame esantį tekstą ir nustato kategoriją. Veikimas paremtas mašininio mokymosi algoritmais skirtais teksto klasifikavimui. Siūlomas metodas išsiskiria, kad – pritaikytas lietuvių k. ir naudojamas tik žodžių pavertimas bendrine forma, bet nėra žodžiai pakeičiami jų šaknimis. Darbo metu sudaryta duomenų rinkinio atrinkimo metodologija, kurios metu atrinktas azartinių, pornografinių, normalių puslapių rinkinys. Realizuotas prototipas pagal siūlomą metodą, kuris pagal puslapio nuorodą nustato tinklapio kategoriją – azartinis, pornografinis, normalus. Eksperimento metu ištirti 4 mašininio mokymo algoritmai, pagal jų veikimo tikslumą išrinktas geriausių rezultatų davęs algoritmas.

Darbo struktūra

Dokumentas susideda iš šešių skyrių. Pirmajame skyriuje – analizuojamas žalingas turinys ir jo poveikis, esami sprendimai rinkoje, kurie skirti turinio blokavimui puslapiuose, teksto klasifikavimo algoritmai. Antrajame skyriuje aprašomas apsaugos nuo žalingo turinio internete metodo koncepcija ir prototipas. Trečiajame skyriuje – pateikiama informacija apie tai, kaip buvo realizuotas prototipas, sudaryti duomenų rinkiniai mašininio mokymuisi, pateikiamas tyrimas bei jo išvados. Ketvirtajame skyriuje aprašoma gautas rezultatas ir išvados. Penktas skyrius skirtas literatūros sąrašui.

1. Žalingo turinio internete analizė

Pagrindinė vaikų apsaugos nuo žalingo turinio internete metodo esmė – pasirinktą tėvų turinį internete filtruoti ir uždrausti pasiekiamumą vaikų įrenginiuose. Šiame skyriuje analizuojama kas yra žalingas turinys internete vaikams ir kas turėtų būti filtruojama, ir draudžiama taip pat analizuojami tėvų kontrolės sistemos būdai apsaugoti vaikus internete.

1.1. Žalinga informacija internete

Žalingu laikomas toks turinys, kuris gali įžeisti asmens vertybes ir jausmus reiškiant politinę nuomonę, religinius įsitikinimus ar pažiūras rasistiniais klausimais. Kas laikoma kenksminga priklauso nuo kultūrinių skirtumų. Šalys gali apibrėžti ribas kas yra leistina ir ne, tačiau būtina atsižvelgti į tarptautines iniciatyvas, kad skirtingos etikos taisyklės nepažeistų žmogaus laisvės į saviraišką [1].

Šiandien daugelis Europos valstybių reaguoja į tam tikro turinio prieinamumą ir platinimą internete, o didžiausią rūpestį kelia seksualinio pobūdžio turinys įskaitant vaikų pornografiją, rasistinį turinį, religinį požiūrį, neapykantos kalbą ir propagandos prieinamumas. Pagal Europos teisės reglamentus nelegalus turinys skirstomas į keturis tipus [2]:

- Medžiaga apie seksualinę prievartą prieš vaikus;
- Rasistinė ir ksenofobinė neapykantos kalba;
- Teroristinis turinys;
- Intelektualinės nuosavybės teisės pažeidžiantis turinys;

ES nėra pilnai apibrėžusi neteisėto turinio internete. Tas pats turinio tipas valstybės narėse gali būti apibrėžtas skirtingai. Atskiroms valstybėms sprendžiant kas yra legalu ir neteisėta – atsižvelgiama į kultūrinius, moralinius, religinius ir istorinius aspektus. Pavyzdžiui Lietuvos Respublikoje yra priimtas Visuomenės informavimo įstatymas [3], kuriame nurodyta draudžiama ir neskelbtina informacija. Šiuo atveju žalingas turinys internete gali būti įvardijimas, kaip neskelbtina informacija, kuri susideda iš:

- Informacijos susijusios su fizinio, psichinio smurto atvaizdavimu;
- Erotinio pobūdžio informacijos, kai rodomas lytinis aktas ar jo imitacija, lyties organai;
- Informacija kurioje propaguojami žalingi įpročiai ar narkotinės medžiagos;
- Informacija, kuria bandoma tyčiotis, niekinti, skatinti neapykanta ar kurstyti tam tikros žmonių grupės diskriminaciją, kuriai gali priklausyti asmenys dėl lyties, seksualinės orientacijos ar kitokių pažiūrų.
- Kiti įstatymai ribojantys neskelbtiną informaciją;

Europos konsulato tyrimo duomenimis, kuris vykdytas 2020 m. rugsėjo mėnesį [4]. Iš 2 863 apklaustų dalyvių, kuriuos sudarė 66% ES piliečių ir 8% ne ES piliečiai iš jų įmonės ir organizacijos (7.4%), verslo asociacijos (6%) ir neorganizacinio pobūdžio įmonės (5.6%). Taip pat vieši asmenys (2.2%), mokymo įstaigos (1.2%) ir kiti (0.9%). Didžioji dauguma dalyvių įskaitant visas kategorijas nurodė, kad internete susidūrė su žalingu ir nelegaliu turiniu. Atkreipė dėmesį, kad „Covid-19“ pandemijos metu juntamas didesnis tokio turinio paplitimas. Konkrečiau 46% susidūrė su nelegaliomis prekėmis, o 67% teigia, kad susidūrė su neteisėtu turiniu ir išreiškė nepasitenkinimą mechanizmų funkcionalumu.

1.2. Žalingo turinio internete vaikams analizė

Vaiko apsauga internete – sudėtingas ir keblus rūpestis reikalaujantis daug pastangų. Norint užtikrinti saugią aplinką reikia nuodugniai ištirti ir suprasti apie internetinės rizikos rūšis, ir galimus pavojus su kuriais gali susidurti jaunieji interneto naudotojai. Internetas pilnas įvairiausio turinio ir veiklos kuri yra linksma ir mokanti vaikus. Tačiau per trumpą laiko tarpą internetas tapo svarbesnė globalios ekonomikos dalis. UKCCIS Evidence Group apžvalga [5] interneto naudotojų kiekis smarkiai padidėjo, įskaitant ir nepilnamečius vaikus pastaraisiais metais. Milijonai vaikų ir paauglių keičiasi laiškais, trumpomis žinutėmis, dalyvauja pokalbių grupėse, naršo internete ir atlieka daugybę kitų veiklų. Didėjantis vaikų ir paauglių aktyvumas internete priverčia susimąstyti apie jų saugumą internete. Vaikų interneto naudojimas labai smarkiai keičiasi priklausomai nuo visuomenės, amžiaus, socialinio statuso ir lokacijos bei įrenginių, kuriuos naudoja. UKCCIS Evidence Group apžvalga [5] išskiria pagrindines rizikas vaikams internete, tai atsispindi pateiktoje 1 lentelėje.

1 lentelė. Interneto pavojai vaikams [5]

	Komercinė rizika	Agresyvi rizika	Seksualinė rizika	Rizikos reikšmės
Turinio peržiūrėjimas internete (vaikas, kaip gavėjas)	Reklama, Šlamštas, Asmeninė informacija	Smurtas / neapykanta	Nepageidaujamas seksualinis turinys	Šališkumas, rasizmas, klaidinantis
Kontaktas internete (vaikas, kaip dalyvis)	Sekimas, Atskleidimas asmeninės informacijos	Internetinės patyčios, priekabiavimas	Susitikimai su nepažystamais žmonėmis iš interneto	Savęs žalojimas
Elgsena internete (vaikas, kaip auka)	Nelegalaus turinio parsisiuntimas, Azartiniai žaidimai, Finansiniai apgaudinėjimai	Priekabiavimas, Internetinės patyčios	Netinkamo turinio kūrimas ir viešinimas internete	Teikiama klaidinanti informacija

Bendra problema yra ta, kad skaitmeniniame pasaulyje tarp vaikų, paauglių ir suaugusių kai pradama dalintis asmenine informacija, nuotraukomis interneto platybėse [6]. Vaikų problema nėra tik tapatybės pavogimo, bet taip pat didžiulė galimybė nukentėti nuo komercinių paslaugų ar sukčių internete. Paslaugos gali apimti vogtų daiktų pirkimą ar pardavimą be jų žinios, prieiga prie azartinių žaidimų internete, smurto, ginklų prekybos, sukčiavimo tinklapių, tinklapių, kuriuose yra mokama, kaip pasigaminti narkotinių medžiagų ar bombų, susidūrimą su žeminančiomis žinutėmis ar vedančiomis į priekabiavimą ar psichologinį smurtą, kuris galiausiai sumažina asmens savivertę ir kelia didelę grėsmę psichologinei savijautai [7].

1.3. Apsaugos nuo žalingo turinio metodai ir priemonės

Apsaugojimui tėvai dažniausiai naudoja tėvų kontrolės sistemas. Tėvų kontrolė iš esmės yra funkcija pridėta norint stebėti ir kontroliuoti naudojamų įrenginių esančių namų tinkle informaciją lengvai perskaitomu formatu. Tėvų kontrolės sistemos gali būti labai skirtingos, priklausomai nuo to kokiu tikslu yra naudojamos. Gali būti skirstomos į keturias kategorijas [8]: turinio filtravimas, naudojimo kontrolė, kompiuterio valdymo įrankiai ir stebėjimo priemonės.

Turinio filtravimas: tradicinis metodas tėvų kontrolei. Ši kontrolė leidžia tėvams kurti taisykles įrenginiams, kad būtų draudžiamos tam tikros programos ar prieiga prie interneto, išfiltruojamas netinkamas turinys.

Naudojimo kontrolė: šiai kategorijai priskiriamas įrenginių limitavimas, kokius tinklapius gali vartotojas lankyti, ką gali siųsti iš interneto ar įrašinėti į kompiuterį. Šiems dalykams yra pareikalaujamos administratoriaus teisės.

Kompiuterio valdymo įrankiai: tai leidžia tėvams reguliuoti kiek jų vaikai praleidžia savo laisvalaikio internete, kiek gali skirti laiko naršymui, kada gali naršyti.

Stebėjimas: ši kategorija dažniausiai būdinga brangiuose tėvų kontrolės sistemose, kurios leidžia tėvams peržiūrėti įrašų žurnalus ką ir kada jų vaikai veikė internete. Dažniausiai naudojami pačio kompiuterio ar įrenginio resursai, siunčiant informaciją ar išsaugant ir persiunčiant ją į serverį.

1.4. Tėvų kontrolės sistemos ir būdai apsaugoti vaikus internete naudojant filtrus

Filtrai gali būti įdiegti įvairiausiais būdais: naudojant programinę įrangą kompiuteriuose, telefonuose taip pat pasitelkiant tinklo infrastruktūrą. Pavyzdys: įgaliojimo (angl. proxy) serveriai, domenų vardų sistemų (angl. domain name system) serveriai ar ugniasienės ribojančios interneto pasiekiamumą. Nei vienas sprendimas pilnai neįgyvendina pilnavertiškos apsaugos, todėl yra būtina naudoti kelias technologijas vienu metu [9]. Kaip pavyzdys galėtų būti kai įmonės viduje dažniausiai yra blokuojami socialiniai tinklai norint sutelkti darbuotojų dėmesį į darbą, kad nebūtų švaistomas laikas nereikšmingiems dalykams. Tėvų atveju norima užtikrinti, kad vaikai nepasiektų netinkamo ar draudžiamo turinio, kuris galėtų sužaloti jų psichologinę būklę.

1.4.1. Naršyklės filtrai

Naršyklės tinklapių filtrai dažniausiai yra plėtiniai, programos ar priedai, kuriuos galima įdiegti ir naudoti tam tikrose interneto naršyklėse. Tokio tipo filtrai yra paplitę ir plačiai naudojami ir patogūs vartotojams. Naršyklės tokios, kaip Chrome, Firefox, Internet Explorer suteikia tėvų kontrolės funkcionalumą [10]. Saugumo funkcijos pateikiamos 2 lentelėje.

2 lentelė. Saugumo funkcijos naršyklėse [10]

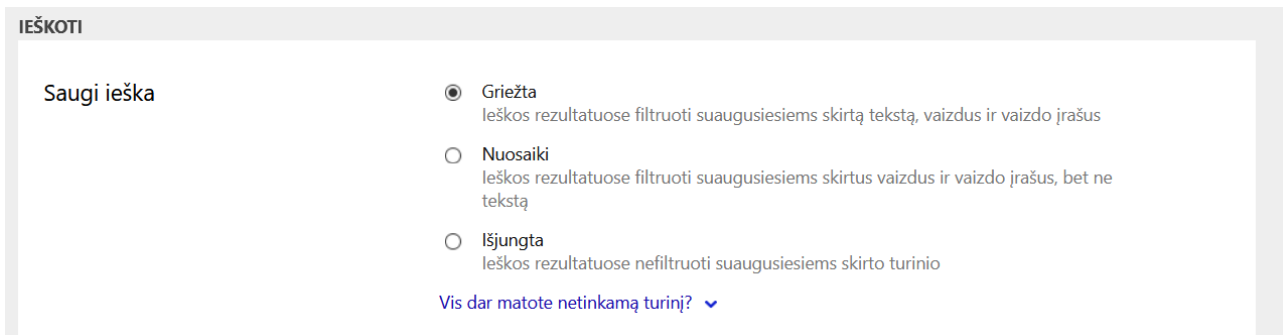
Naršyklė	Saugumo funkcijos
Chrome	Slaptažodžių išsaugojimo išjungimas, kas leidžia apsaugoti vaikų tinklapių apsilankymą, kuriuose lankosi tėvai; Pagal turinio nustatymus, galima panaikinti nuotraukas, sausainiukus. Tam tikri priedai gali būti išjungiami.
Firefox	Išsaugotų slaptažodžių išjungimas; Firefox naršyklė suderinama su „Windows“ operacinės nustatymais, kurie leidžia sustabdyti nepageidaujamų failų atsisiuntimą, svetainių filtravimą.
Internet Explorer	Blokuojami tam tikri tinklapiai pridodant į juodąjį sąrašą; Leidžiami turinio nustatymai, kurie apima nuotraukas ir aktyvų turinį, kuris leidžiamas kompiuteryje.

Netinkamo turinio filtravimas, kuris yra apibrėžiamas naršyklės plėtiniuose gali sumažinti paieškos rezultatų kiekį. Šis filtravimas yra svarbus nepilnamečių ir paauglių tėvams, kad nereikėtų prižiūrėti pastoviai vaikus ką jie veikia internete [10].

1.4.2. Paieškos sistemų filtrai

Paieškos sistemų filtrai išfiltruoja netinkamą turinį iš paieškos rezultatų, kuriuos vartotojas mato ieškant informacijos. Kai kurios paieškos sistemos turi šią funkciją integruotą į vartotojo sąsają „Google“ ir „Bing“ turi saugios paieškos nustatymus. Bing paieškos sistemos atveju galimos trys parinktys: išjungta, nuosaiki, griežta 1pav. pateikiamos galimos parinktys.

- Griežta – filtruojamas suaugusiems skirtas tekstas, vaizdo įrašai ir nuotraukos.
- Nuosaiki – filtruojamas tik suaugusiems skirti vaizdo įrašai ir nuotraukos.



IEŠKOTI

Saugi ieška

☒ **Griežta**
Ieškos rezultatuose filtruoti suaugusiems skirtą tekstą, vaizdus ir vaizdo įrašus

☐ **Nuosaiki**
Ieškos rezultatuose filtruoti suaugusiems skirtus vaizdus ir vaizdo įrašus, bet ne tekstą

☐ **Išjungta**
Ieškos rezultatuose nefiltruoti suaugusiems skirto turinio

[Vis dar matote netinkamą turinį?](#) ▼

1 pav. „Bing“ saugios paieškos nustatymai.

„Google“ paieškos sistemos atveju galimas tik įjungimas saugios paieškos (žiūrėti 2 pav.).

Saugios paieškos filtrai

Įjungus saugią paiešką slepiamas nepadorus turinys, pavyzdžiui, pornografija. Saugios paieškos nuostatas gali parinkti įrenginio arba tinklo administratorius. Jei negalite išjungti funkcijos, kreipkitės į tinklo ar įrenginio administratorių.

☒ Įjungti Saugią paiešką [Sužinokite daugiau](#)

2 pav. „Google“ saugios paieškos nustatymai.

Tačiau paieškos sistemų nustatymai turi didžiausią trūkumą, kad vartotojas be administratoriaus teisių gali išjungti arba įjungti filtrus. Taip pat šie filtrai gali būti lengvai apeinami jeigu vartotojas žino tinklapio adresą ir jį gali paprasčiausiai pasiekti, jam tai būtų nedraudžiama, nes turinys filtruojamas tik paieškos rezultatuose.

1.4.3. Interneto paslaugų tiekėjo ribojimai

Interneto paslaugų tiekėjai turi galimybę blokuoti pasirinktus tinklapius visiems tinklo įrenginiams naudojančioms tiekėjo paslaugas. Dažniausiai aptinkamas šalyse, kur žiniatinklyje turinys yra cenzūruojamas. Kaip pavyzdys: Nyderlanduose autorių teises ginanti asociacija „Brein“ kartu su interneto tiekėju „Ziggo“ dėl nelegalaus ir draudžiamo turinio blokavimo įvedimo pradėjo diskusijas [11]. Pamažu šis filtravimo tipas skinasi kelią į viso pasaulio šalis. Pavyzdžiui „Virgin Media“ naudoja tinklapių filtravimą, siekiant užblokuoti piratavimo tinklapį „Pirate Bay“ nuo vartotojų galimybės pasiekti nelegalų turinį [12]. Lietuvoje yra naudojama praktika susijusi su interneto svetainėmis, kurios pažeidžia vartotojų teises, azartinių lošimų įstatymo nustatyta tvarka yra taikomas neteisėtos informacijos blokavimas interneto paslaugų teikėjų lygmenyje [13].

1.5. Esamų tėvų kontrolės sprendimų analizė

Prieš pradėdant nagrinėti rinkoje esančius tėvų kontrolės sprendimus, kurie leistų tėvams pasirūpinti ar kontroliuoti vaikų saugumą internete, reikėtų išsikelti kriterijus, kuriais bus galima vadovautis analizuojant esamas sistemas rinkoje. Išskirti kriterijai padės nustatyti sistemų trūkumus ir privalumus.

Atliekant tėvų kontrolės sprendimų analizę, atsižvelgsime į šiuos kriterijus:

- Tinklapių filtravimas – pagrindinis tėvų kontrolės sistemų kriterijus – turinio filtravimas, galimybė blokuoti tinklapius, kuriuose galimai yra smurtas ar netinkamas turinys vaikams. Prieigos planavimas irgi įtraukiamas prie šio kriterijaus ar galima tėvams reguliuoti laiką prieigai prie interneto vaikams.
- Įrenginių limitavimas – keliuose įrenginiuose galima įdiegti sistemą ar nėra limituojami įdiegimai.
- Įrenginių suderinamumas – kokiuose įrenginiuose naudojant operacinę sistemą gali veikti tėvų kontrolės sistema.
- Pažangios funkcijos – kokios papildomos funkcijos yra tėvų kontrolės sistemose, išskyrus tinklapių filtravimo, kurios padėtų tėvams užtikrinti vaikų saugumą internete.
- Platforma – kokiam įrenginyje turi būti įdiegta sistema.

Nagrinėjami sprendimai pasirinkti tie, kurie galimai galėtų užtikrinti saugų internetą vaikams ir dalinai suteikti tėvams kontrolę peržiūrint ką jų vaikai naudoja, ar veikia internete. Atlikus analizę esamų sprendimų tėvų kontrolei vaikams padėtų suprasti jų trūkumus, bei privalumus ir leistų išsirinkti metodologiją, kuri galėtų būti tobulinama arba lyginama su naujai sukurtu metodu. Pasirinktos trys esančios rinkoje sistemos. „Qustodio“ ir „SafeDNS“ pasirinkimą lėmė ryšių reguliavimo tarnybos aprobuoto turinio filtravimo priemonių rekomendacija [14], o „Net Nanny“ reitingavimas internete, kaip vienos iš geresnių tėvų kontrolės sistemų [15].

1.5.1. Net Nanny

Viena iš sistemų tėvų kontrolės – „Net Nanny“ [16] skirta apsaugoti vaikus nuo netinkamo turinio internete, internetinių patyčių ir kitų grėsmių kylančių internete. Ši sistema tėvams leidžia filtruoti turinį, stebėti ir blokuoti nesaugų turinį leidžiant vaikams pasiekti internetą. Skirtingai nuo kitų tėvų kontrolės sistemų „Net Nanny“ naudoja dinامينius filtrus ir analizuoja kiekvieną tinklapį ar jis yra tinkamas vaikui pagal tėvų sukurtas taisykles [17]. Vieno įrenginio mokestis 32 Eur per metus, už kiekvieną papildomą įrenginį taikomas papildomas mokestis 16 Eur. „Net Nanny“ gali būti įdiegta Windows, macOS, Android, Ios ir Fire OS (Kindle) įrenginiuose. Mobiliuose įrenginiuose galima įdiegti dvi programas, vaikų įrenginiuose (stebėjimui), o kitą tėvų įrenginyje konfigūravimui ir taisyklių kūrimui, taip pat tėvams suteikiama ir internetinio tinklapio prieiga konfigūravimui.

1.5.2. Qustodio

Tėvų kontrolės įrankis, kuris veikia tik viename Jūsų šeimos kompiuteryje yra nenaudingas šiuolaikiniame pasaulyje. Tėvų kontrolės programa „Qustodio“ pasižymi beveik visomis funkcijomis, kurių gali prireikti tėvams rūpinantis vaikų saugumu [18]. Sistema turi žiniatinklio turinio filtravimą, programų blokavimą ir išsamų veiklos žurnalo pateikimą. Visą veiklos stebėjimą bei konfigūravimą gali atlikti naudojant prietaisų skydelį mobiliuose telefonuose, tai reiškia, kad

tėvams suteikiama visur ir visada galimybė stebėti vaikų veiklą internete. Ši sistema yra brangi, tačiau turi galimybę, kad vienam įrenginiui gali būti netaikomas joks mokestis su limituotomis funkcijomis. Nemokama versija leidžiama peržiūrėti tik 7 dienų įrenginio veiklą internete ir suteikiamas tik pagrindinis filtravimas tinklapių. Mokama versija – pagal įrenginių skaičių: 5 įrenginiai, 10 įrenginių, 15 įrenginių. Atitinkamas ir mokestis 45 Eur, 79 Eur, 113 Eur per metus. Ši sistema gali veikti Windows, macOS, Android ir iOS operacinėse sistemose. Taip pat reikėtų atkreipti dėmesį, kad nors ir turi platų platformų suderinamumą, tačiau programos funkcijos specifiskai priklauso nuo pasirinkto įrenginio operacinės sistemos.

1.5.3. SafeDNS

Atskiros programos įdiegimas tėvų kontrolės produkto šeimos kompiuteryje nėra labai efektyvus, kai vaikai gali pereiti prie naršymo internete išmaniajame telefone ar planšetiniame kompiuteryje, ar kokiame nors kitame įrenginyje. Už sąlyginai mažą kainą 16 Eur per metus „SafeDNS“ tinkamai išfiltruos, bet kokiame įrenginyje, kuriame bus įdiegta programa netinkamą turinį, tačiau nesugebės užblokuoti apgaulingų ir kenkėjiškų puslapių [19]. Nemokamos versijos ši programa neturi, suteikiamas tik bandomasis 15 dienų laikotarpis. Šio serviso diegimas yra galimas lokaliai kompiuteryje arba tinklo lygmenyje namų maršrutizatoriuje. Didžiausias trūkumas šios sistemos, kad laiko planavimas (prieigos prie interneto) apamai apriboja pilnai interneto prieigą, kas nėra labai gerai. Dar vienas iš trūkumų tas, kad sistemoje yra rodomos ir visos DNS užklausos, todėl tarp tinklapių statistikos ir papuola serverių adresai, kurie vartotojui nesuteikia pridėtinės vertės [19].

1.6. Esamų tėvų kontrolės sprendimų analizės išvados

Kaip matome 2 lentelėje išanalizuotų tėvų kontrolės sistemų geriausia yra „Qustodio“, nes suteikiamas didžiausias tėvų kontrolės funkcionalumas ir galimybės kontroliuoti vaikus. Įrenginių limitavimas lyginant su „Net Nanny“ yra priimtinesnis žmonėms turintiems daugiau įrenginių šeimos ūkyje, kuriais naudojasi vaikai.

3 lentelė. Tėvų kontrolės sistemų lyginamoji lentelė

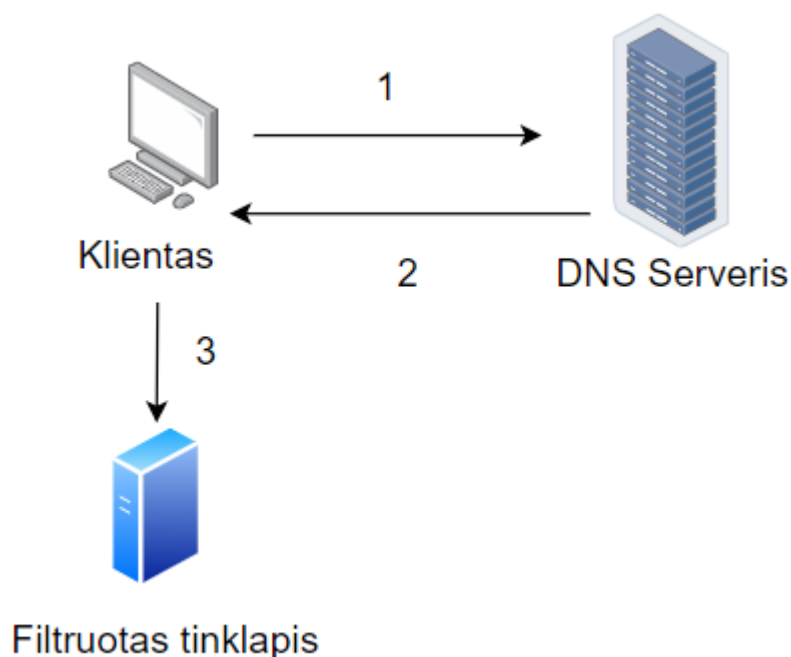
Kriterijus	Net Nanny	Qustodio	SafeDNS
Tinklapių filtravimas	Yra	Yra	Dalinai
Įrenginių limitavimas	1 įrenginys (už papildomą mokestį)	5 / 10 / 15 įrenginių	Nėra
Įrenginių suderinamumas	Windows, Mac OS x, Android, iOS, Kindle Fire	Windows, Mac OS x, Android, iOS, Kindle Fire	Windows, Mac OS x, Android, iOS, Linux OS
Pažangios funkcijos	Aplikacijų blokavimas, ekrano laiko sekimas	Aplikacijų blokavimas, žinučių ir skambučių sekimas, kelių platformų palaikymas	Lengvas tinklapių filtravimas, profilių kūrimas ir apribojimas pagal laiko planavimą
Platforma	Lokaliai įrenginiuose	Lokaliai įrenginiuose	Lokaliai įrenginiuose ir maršrutizatoriuose

Reikėtų pastebėti, kad „Qustodio“ tėvų kontrolės sistema yra viena iš rekomenduojamų ryšių reguliavimo tarnybos [14].

1.7. Internetinių tinklapių filtravimas naudojant domenų vardų sistemą

1.7.1. Domenų vardų sistemos mechanizmas

Domenų vardų sistema yra paskirstyta duomenų bazių sistema skirta saugoti pagrindinio kompiuterio ir domeno vardo informaciją tinkle. DNS funkcijos identifikuoja kiekvieną kompiuterį, kaip mazgą kompiuterių tinkle naudojant interneto protokolo (IP) adresą [20]. Domenų vardų sistemos mechanizmas pavaizduotas 3 pav.



3 pav. Domenų vardų sistemos mechanizmas [20].

Mechanizmo paaiškinimas:

1. Domenų vardo sistemos sprendėjas / klientas siunčia užklausą į DNS serverį.
2. Vardų serveris apdoroja užklausą, patikrina vartotojo užklausą apie prašomą pasiekti domeną, jei interneto adresas nėra įtrauktas į sąrašą, klientui grąžinamas atsakymas su leidimu pasiekti adresą.
3. Priklausomai nuo grąžinto atsakymo iš domenų vardų serverio – pasiekiamas arba ne.

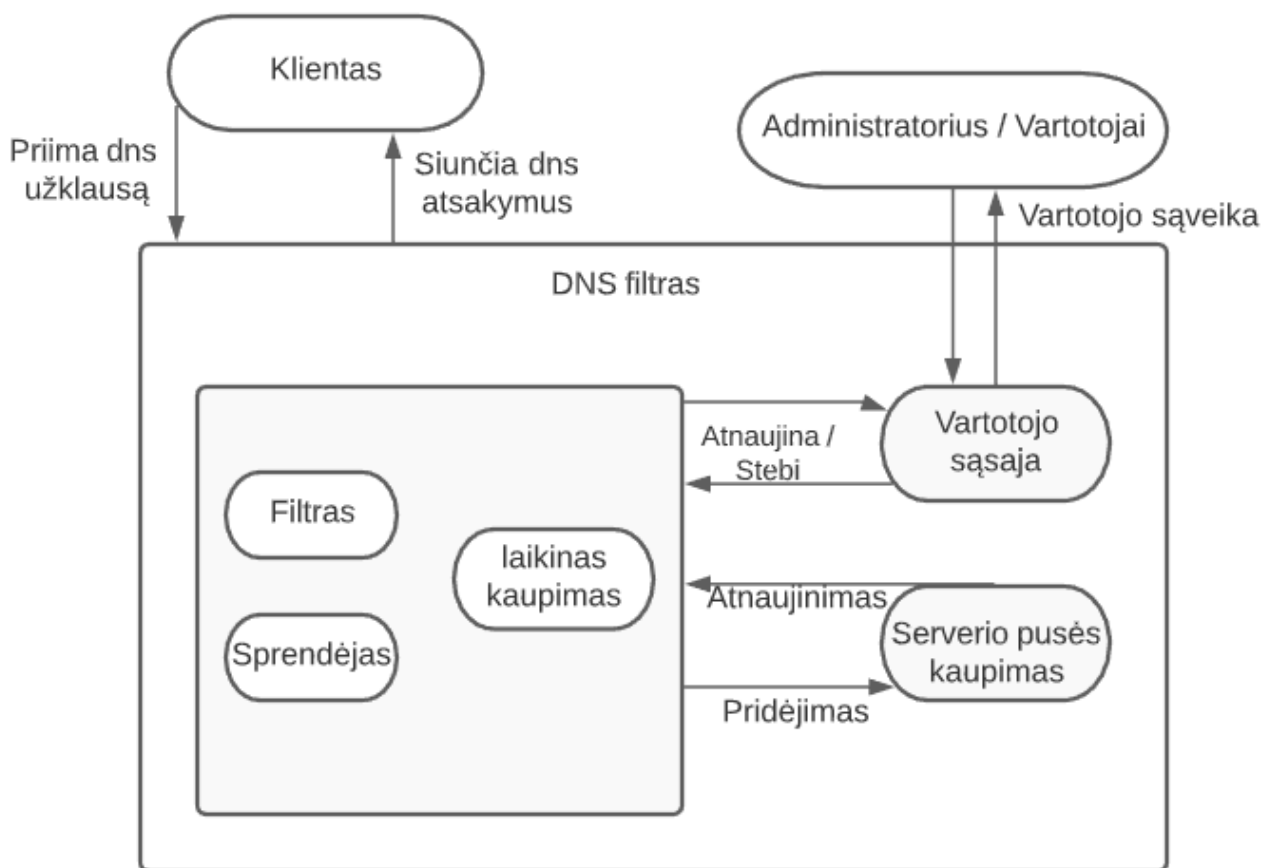
1.7.2. Domenų vardų sistemos filtravimo galimos architektūros analizė

Sistemos architektūra buvo sukurta, kaip alternatyvus sprendimas debesies pagrindu veikiančios DNS žiniatinklio filtravimo sistemos [21]. Filtruoja svetaines domenų vardų sistemos lygyje ir teikia visą žurnalo informaciją vartotojo sąsajoje. Filtras sukurtas naudojant ir pritaikant šias funkcijas:

- Domeno vardo sprendimas – rekursyviai sprendžiami domenų vardai saugyklose ir grąžinamas serverio atsakymas.

- Juodasis sąrašas – blokuojamos DNS užklausos ir įtraukiamos į juodąjį sąrašą.
- Įtraukimas į baltąjį sąrašą – užtikrina, kad tam tikro domeno kategorija arba visas domenas nebūtų įtraukiamas į juodąjį sąrašą.
- Tik baltasis sąrašas – leidžiamos tik tos DNS užklausos, kurios įtrauktos į baltąjį sąrašą, o visą kitą yra blokuojama.
- Laiko blokavimas – leidžia nurodyti, kuriomis savaitės dienomis ir kokių laiku būtų taikomi prieigos draudimai.
- Domeno pakaitinis simbolis – leidžia suderinti domeną su pakaitiniu simboliu paieškoje vardų.
- Konfigūruojamas blokavimo puslapis – tikrinamas blokuojamas puslapis žiniatinklio serveryje, kuris suteikia galimybę peradresavimui.

DNS filtravimo sistemos architektūrą sudaro trys pagrindinės posistemės [21]: sistemos vartotojai ar administratorius ir pagrindinis kompiuteris (klientas) iš kurio yra gaunamos DNS užklausos. Šios architektūros pavyzdys pavaizduotas 4 pav.



4 pav. Domenų vardų sistemos architektūra [21].

Klientai arba įrenginiai tokie, kaip nešiojamieji kompiuteriai, mobilieji įrenginiai, kuriuose gali būti naudojamas DNS skiriamasis įrankis, kad būtų susiejami domenų vardų adresai į interneto adresą.

Sistemos administratorius arba vartotojas: administruojantis sistemą naudojant vartotojo sąsajos posistemę. Sistemos administravimas apima nurodant serverio parametrus, apibrėžiant prieigos taškus ir stebint DNS serverio filtravimo posistemę.

Vartotojo sąsajos posistemė: žiniatinklio pagrindu sukurtas – tinklapis, kuriame administratorius gali konfigūruoti prieigos sąrašus ir peržiūrėti DNS užklausų žurnalus. Visi šie nustatymai yra saugomi serverio dalies posistemėje. Pakeitimai nėra automatiškai atliekami DNS serverio filtravimo sistemoje. Pradžioje pakeitimai atliekami vartotojo sąsajoje, kurie išsaugomi laikinoje saugykloje ir tik po to perduodama į serverio dalį.

Serverio dalies posistemė: nuolatinė saugykla, kurioje talpinami visi nustatymai ir sąrašai. Tai yra bendras portalas tarp vartotojo sąsajos ir DNS sprendėjo ir filtravimo posistemės.

DNS sprendėjo ir filtravimo posistemė: praplečia įprastą DNS sprendėją pridėdant filtravimo galimybes. Tai suteikia galimybę naudoti prieigos sąrašus iš serverio pusės. Ši posistemė atsakinga už prieigos sąrašų sprendimą ir tikrinimą kiekvienai gautai DNS užklausiai. Šie visi žurnalai patikros saugomi serverio dalyje.

Šios sistemos testavimo rezultatai pateikiami 4 lentelėje. Autorius [21] lygina su kitomis panašiomis sistemomis, kurios filtruoja žiniatinklio turinį.

4 lentelė. Domenų vardu patikrinimo ir blokavimo testavimo rezultatai

DNS serveris	Išspręsti domenai (iš 40 testavimo domenų)	Blokuoti domenai (iš 40 testuotų domenų)
DNS filtravimo sistema	40	25
OpenDNS	40	Nėra pagal nutylėjimą filtrų
Dyn	40	8
DNSAvantage	40	2
YandexDNS	40	2

Reikėtų atkreipti dėmesį į tai, kad ši sistema išfiltravo daugiausiai tinklapių dėl funkcionalumo ir galėjimo kurti taisykles lokaliai, kai kitos lyginamos sistemos nesuteikia šios galimybės, ir yra naudojami viešieji juodieji sąrašai.

1.8. Turinio filtravimas naudojant teksto klasifikacijos metodus

Žiniatinklio filtravimas naudojant teksto klasifikaciją yra plačiai naudojamas šiomis dienomis [22]. Šis metodas naudojamas norint nustatyti puslapius, kurie turėtų būti blokuojami. Automatinio teksto klasifikavimo užduotis - organizuoti dokumentus pagal nustatytas klases ar grupes naudojant mašininio mokymosi algoritmus. Paprastai tariant tai yra vienas iš svarbesnių būdų grupuoti milžinišką informacijos kiekį, kuris egzistuoja nestruktūruotu tekstiniu formatu.

Yra daugybė algoritmų skirtų teksto klasifikavimui: Naivus Bjesas, K – artimiausių kaimynų (angl. K-Nearest Neighbours, K-NN), atraminių vektorių (angl. Support Vector Machine, SVM) ir neuroninių tinklų.

Nagrinėjant klasifikavimo metodus išsikelsime tokius kriterijus:

- Tikslumas – kaip efektyviai naudojant metodą gali būti filtruojamas turinys;
- Klasifikavimas – ar klasifikavimui naudojamos klasės;
- Mokymosi mechanizmo sudėtingumas – metodas gali neturėti mokymosi mechanizmo.

1.8.1. Naivus Bajesas metodas

Naivus Bajesas – tikimybinis klasifikatorius naudojantis skirtingas variacijų grupes. Klasifikatoriaus principas – sekami, kokie požymiai yra būdingi tam tikrai klasei [23]. Skaičiavimuose yra remiamasi Bajeso taisykle (1 lygtis). Klasė parenkama ta, kuri tikėtina, kad bus teisinga (2 lygtis).

$$P(A) \cdot P(B|A) = P(B) \cdot P(A|B) \quad (1)$$

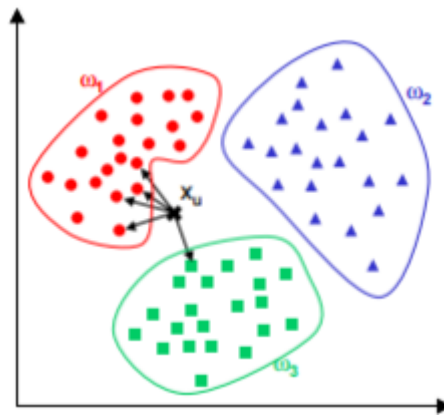
$$C_{best} = \arg \max_{C \in c} P(C = c) \cdot \prod_{i=1}^n P(x_i|C = c) \quad (2)$$

Išskiriamos skirtingos Naivaus Bajeso versijos priklausomai kokia pasiskirstymo funkcija yra naudojama [23]:

- Gausinis Naivus Bajesas (angl. Gaussian Naïve Bayes) – laikoma, kad požymio vertė pasiskirsčiusi normaliuoju skirstiniu;
- Multinominis Naivus Bajesas (angl. Multinomial Naïve Bayes) – laikoma, kad požymiai žymi dažnius, o požymiuose neturėtų būti neigiamų reikšmių.
- Bernulio Naivus Bajesas (angl. Bernouli Naïve Bayes) – naudojami binariniai požymiai.

1.8.2. K – artimiausių kaimynų metodas

K – artimiausio kaimyno taisyklė yra labai intuityvus metodas, kuris leidžia klasifikuoti nežymėtus mėginius panaudojant jų panašumus su kitais esančiais mėginiais mokymosi rinkinyje [24]. Ši taisyklė reikalauja trijų dalykų: sveiko skaičiaus k, žymėtų mėginio rinkinio ir metrikos priemonių matuojant artumą. Pateikiamas pavyzdys 5 pav. kaip atrodo klasifikavimas naudojant šį metodą.

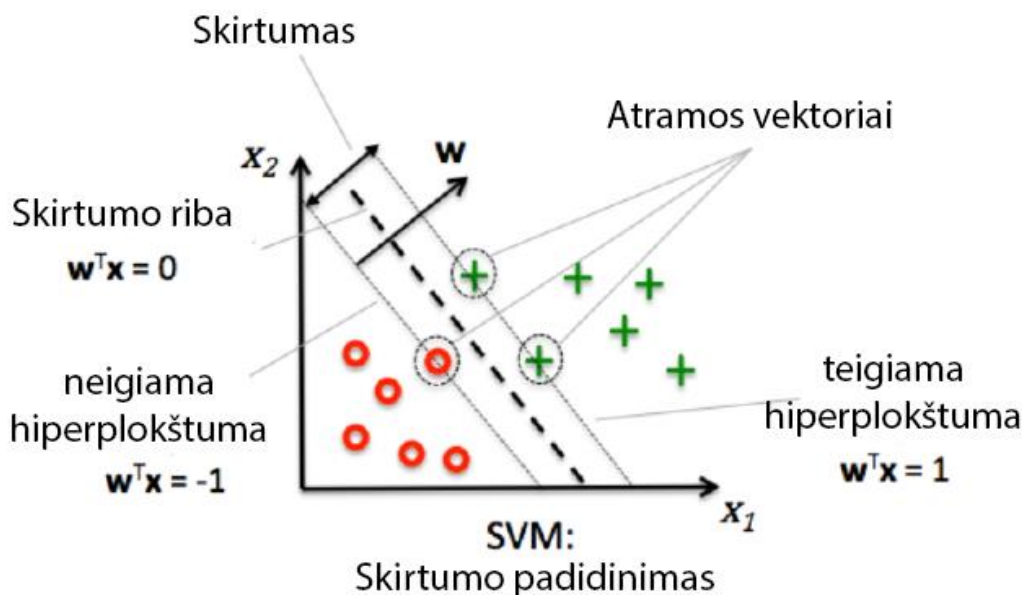


5 pav. k – artimiausio kaimyno klasifikavimo pavyzdys [24].

Iš 5 pav. matome, kad yra trys klasės ir pagrindinis tikslas rasti klasės žymą nežinomam mėginiui X_u . Šitame atvejyje naudojamas Euklidinis atstumas ir $k = 5$. Iš penkių artimiausių kaimynų, keturi priklauso ω_1 ir vienas priklauso ω_3 , todėl X_u priskiriama prie ω_1 . Privalumai šio metodo yra tokie, kad jis yra paprastas, nereikalaujantis sudėtingų parametų modelio sukūrimui [24]. Neturi jokio mokymosi mechanizmo, todėl metodas dar dažnai vadinamas „tingiuoju“.

1.8.3. Atraminų vektorių metodas

Atraminų vektorių metodo veikimo principas – siekimas maksimizuoti atstumą tarp klases skiriančios hiperplokštumos ir arčiausiai šalia jos esančių pavyzdžių, vadinamų atraminiais vektoriais [23]. Šio metodo iliustracija pavaizduota 6 pav. SMV metodą galima naudoti su skirtingais branduoliais, pavyzdžiui tiesiniu (angl. linear) ir Gausiniu branduoliu. Galimybė naudoti skirtingus branduolius leidžia SVM metodą pritaikyti įvairiems uždaviniams spręsti.



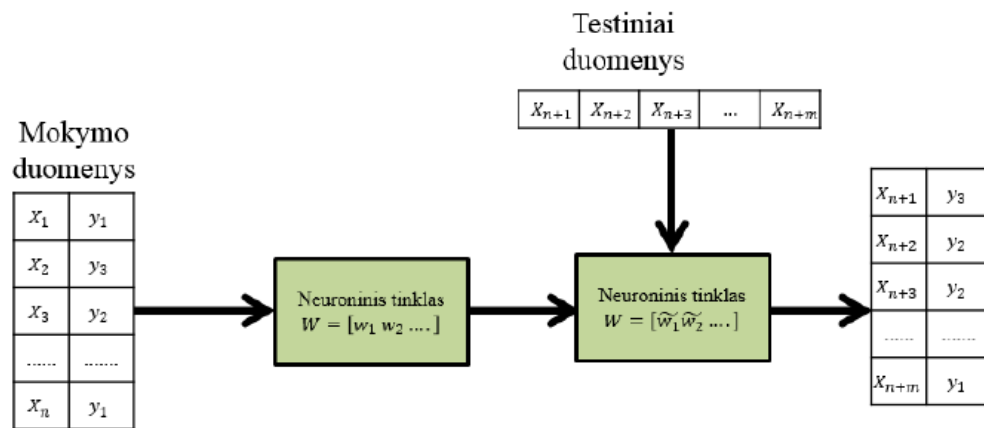
6 pav. SVM metodo veikimo pavyzdys [23].

Hiperplokštuma kuri brūkšniuotos linijos pagalba yra atskiriama į dvi kategorijas ir apibrėžiama naudojant formulę $(w \cdot x) + b = 0$; kur w yra svorio vektorius, o b apibrėžia hiperplokštumos poslinkį nuo pradinio taško [10].

1.8.4. Neuroninio tinklo metodas

Neuroninių tinklų mokymui yra naudojami duomenų rinkiniai (masyvai). Dirbtiniu neuronu – vadinama tokia funkcija, kuri turi baigtinį įvesties reikšmių skaičių ir išvesties reikšmė yra viena. Tarpusavyje sujungtus neuronus galime vadinti neuroniniu tinklu [25]. Neuronų kiekis priklauso nuo sprendžiamo uždavinio sudėtingumo. Klasifikavimo uždavinyje, kai norima atpažinti objektą reikia priskirti duomenų rinkinius kuriai nors pasirinktai klasei (neuroninio tinklo mokymo pvz. pateikiamas 6 pav.). Tarkim objektas yra aprašomas vektoriumi: $X = [x_1, x_2, x_3 \dots x_n]$

Aibė $Y = [y_1, y_2, y_3 \dots y_m]$ aprašo visas galimas klases y_i , kurioms priklauso objektas X . Mokymo fazėje imami mokymo duomenys su priskirtomis klasėmis – (X_i, y_j) .



7 pav. Neuroninio tinklo mokymo pavyzdys [25].

Optimizavimo algoritmas išrenka neuroninio tinklo koeficientus $\omega_1, \omega_2, \dots, \omega_k$, kad klasifikavimo funkcijoje K būtų atvaizduojami mokymo fazės objektai X, kuriuos paverstų į klasę y, kuriai turėtų priklausyti [25].

1.8.5. Klasifikacijos metodų panaudojimas filtruojant turinį

Tyrėjas Ptaszynski pirminiame savo darbe atliko tyrimą, kuriame sukūrė duomenų rinkinį susijusį su interneto patyčiomis, naudojant vulgarius žodžius [26]. Žodžių klasifikavimui panaudojo atraminių vektorių metodą. Optimizavus sistemą sugebėta nustatyti internetines patyčias 88.2% tikslumu. Tačiau didinant duomenis pradėjo mažėti rezultatai, kas autorių privertė daryti išvadą, kad SVM metodas nėra idealus žodžių klasifikavimui nustatinėjant patyčias internete.

Mašininio mokymosi algoritmais pagrįstos sistemos, klasifikuojant tinklapių nuorodas ir papildomai naudojant aktyvios direktorijos projektus (angl. open directory project, ODP), kuriuose talpinami juodieji sąrašai [27]. Kiekviena nuoroda yra lyginama su esančiomis juodajame sąrašė. Pagrindinis tokios sistemos trūkumas yra toks, kad tinklapio nuoroda ne visada atitinka esantį turinį, todėl dažnai tinklapiai būna neišfiltruojami, todėl papildomai – naudojama tinklapių klasifikacija atraminių vektorių mašininio mokymosi algoritmas turinio filtravimui. 5 lentelėje pateikiami atrinkti skirtingi naudojami metodai su įvairiais tikslais ir jų pasiektais filtravimo tikslumais.

5 lentelė. Skirtingų klasifikacijos metodų palyginimas [27]

Tikslas	Naudojamas metodas	Duomenų rinkinys	Tikslumas
Tinklapių nuorodų klasifikacija	N – Gram	4,167	94%
Kenkėjiškų svetainių aptikimas pagal nuorodą	Naivus Bajesas ir SVM	15000	90%
Pornografinio turinio filtravimas pagal nuorodą	K – artimiausių kaimynų	Nil	96%
Filtravimas naudojant hiperteksto klasifikaciją	K – artimiausių kaimynų ir SVM	1,400	87%

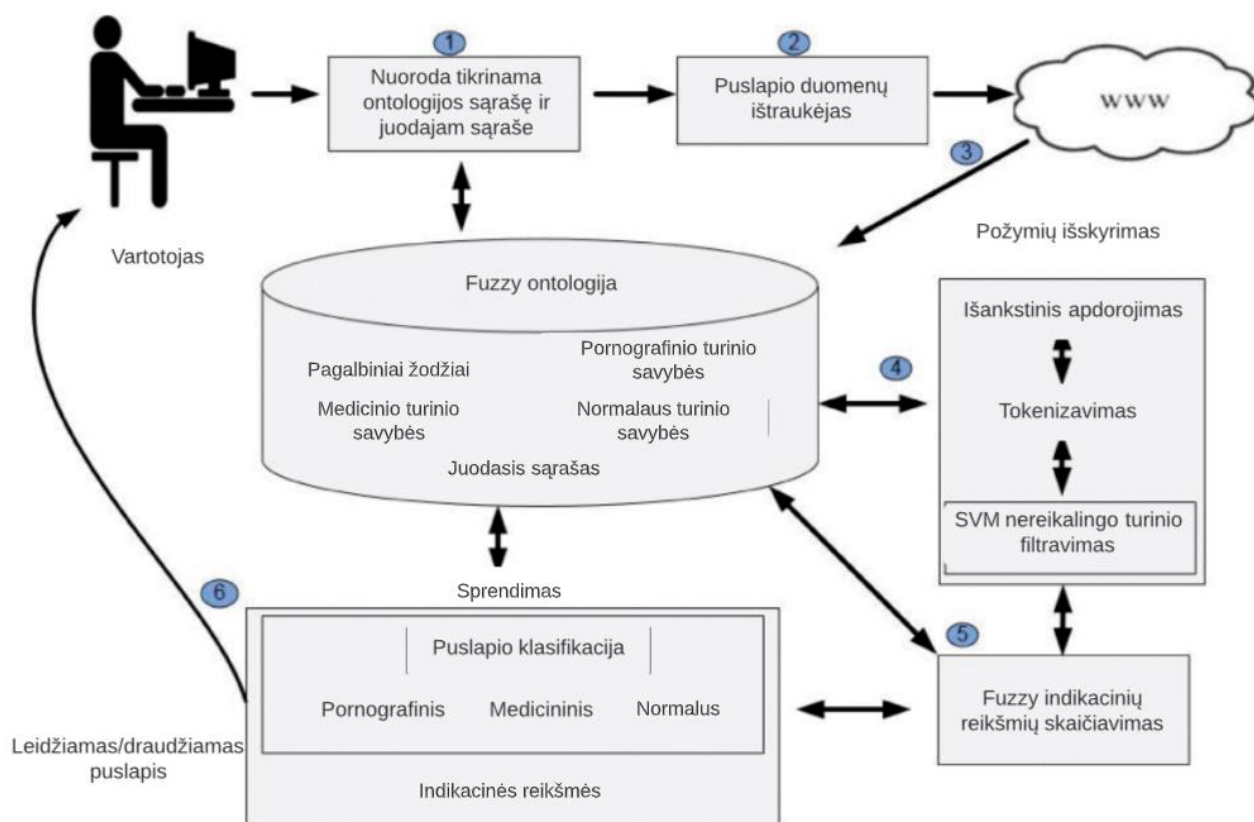
Kaip matome iš lentelės naudojant SVM ir k – artimiausių kaimynų algoritmą su 1 400 tinklapiais buvo pasiektas 87% filtravimo tikslumas [27]. Geriausias rezultatas buvo pasiektas naudojant k – artimiausių kaimynų metodą filtruojant pornografinio turinio tinklapius. Suaugusių turinio filtravimui naudojama keletas klasifikacijos metodų: nuotraukų filtravimas ir žalingų žodžių filtras. Nuotraukų filtravimas naudoja statistinį modelį žmogaus odos aptikimui pasitelkiant neuroninio

tinklo klasifikavimo metodą taip pat pateiktoje lentelėje atspindi tyrimai, kuriuose ribotai ištrauktas turinys iš pasirinkto puslapio, todėl klaidingai būna aptiktas pornografinio pobūdžio turinys [27].

SVM metodu pagrįstas turinio filtravimas vyksta analizuojant atidarytą tinklapį HTML formatu [27]. Tinklapiai turi dažniausiai įvairių skriptų, stiliaus failų, tam tikrų žymėjimų (angl. tags), pavadinimo, meta duomenų (angl. metadata) informacijos. Autoriaus pasiūlyta sistema ištraukia informaciją ir naudoja morfologinę analizę siekiant nustatyti žodžių formas ir leksiką.

Kaip pavyzdys tinklapio žymoje <title>(angl. Becouse every tumour is unique | Generic Testing for Breast Rates of) </title> susidaro iš žodžių testavimas ir reitingai (angl. testing, rates). Šie išvardinti žodžiai turi pagrindines žodžių formas test, rate analizavimui naudojamos būtent šios formos reikšmės. Atrinkti žodžiai nebūtinai gali turėti neigiamas reikšmes, bet jie gali žymėti, kaip nenaudingą turinį, kuriame naudojami šie žodžiai. Tokiu būdu autorius [27] klasifikuoja pornografinio turinio tinklapius naudojant „Fuzzy“ ontologiją ir SVM turinio filtravimą. „Fuzzy“ ontologija naudojama dalintis informacija, kurią galėtų suprasti ir žmogus, ir mašina. Informacijos dalinimasis vyksta specifiniu domenu nepertvarkant informacijos.

Ontologija yra semantinio tinklo pagrindas. Joje saugoma informacija, kuri svarbi programų sąveikai. Turi keturis pagrindinius elementus: klases, sąvokas, atvejus (angl. instances) ir santykius [27]. Egzistuoja trys tipai tinklapių ontologijos (angl. web ontology language, OWL): OWL-Lite, OWL-Descriptive ir OWL-Full. Šiuo atveju autoriaus darbe naudojamas OWL sukuria duomenų rinkinius ir nustatymus turinio klasifikavimui.



8 pav. „Fuzzy ontologijos“ siūlomas architektūros sprendimas [27].

„Fuzzy ontologijos“ sistemos pateiktoje architektūroje 8 pav. matome, kad norint įvertinti tinklapį gaunama iš ontologijos ištrauktų raktažodžių reikšmės taip pat vyksta tikrinimas ir juodajame sąraše, ir tuomet yra nusprendžiama kuriai kategorijai puslapis būtų priskiriamas. Pagal aprašytą metodologiją autorius sudaro kiekvienai kategorijai žodžių sąrašą priskirdamas jiems reikšmes, kaip pavyzdys jeigu puslapis yra medicininis, tai žodis krūtinė gauna reikšmę 0.21 arba 0.91 tokiu būdu sudaromas rinkinys pagal kurį po to – identifikuojama puslapio kategorija. Pagal jo atliktą tyrimą jam pavyko pasiekti 98% tikslumą naudojant ontologiją ir atraminių vektorių metodą. Atrinktų puslapių sąrašas nėra pateikiamas, kuris naudojamas pateikti tyrimo rezultatus.

Kitas sprendimas automatinio pornografinio ir azartinių lošimų svetainių aptikimo [35] architektūriniu požiūriu panašus į pateiktą 8 pav., tačiau jame nėra naudojamas ontologijos sprendimas. Teksto ypatybių išskyrimui naudojamas doc2vec modulis. Esminė doc2vec idėja – mokytis fiksuoto ilgio požymių reprezentacijų iš kintamo ilgio teksto pvz. sakinių, pastraipų ar dokumentų remiantis neuroniniais tinklais. Todėl šiame sprendime lyginant [27] autoriaus sprendimą nebelieka žmogiškosios klaidos sudarinėjant žodžių rinkinį su priskirtomis skaitinėmis reikšmėmis. Puslapio išgautas tekstas yra vektorizuojamas doc2vec pagalba ir paverčiamas į skaitinę reikšmę. Šiame tyrime buvo atliktos šios išankstinio apdorojimo veiksmai:

- HTML žymų pašalinimas.
- Skyrybos ženklų ir specialių simbolių pašalinimas (pvz. =, \$ ir kt).
- Teksto segmentavimui naudojamas NLPIR-ICTCLAS2016 pritaikytas Kinų kalbai.
- Šalinami nereikšmingi žodžiai arba dažniausiai pasikartojantys, kurie kartojasi tekste (angl. stop words) ir neturi reikšmės klasifikavimo rezultatams.

Eksperimentinį duomenų rinkinį sudarė 2722 puslapiai. Naudojant tokio dydžio duomenį rinkinį ir doc2vec modulį naudojant atraminių vektorių metodą buvo pasiektas 91,94% ir 89.58% (pornografinio ir azartinių) puslapių aptikimo tikslumas.

Sekančiame tyrime - neigiamo turinio svetainėse nustatymas naudojant mašininio mokymosi metodą [36] tyrėjas panaudoja teksto pavertimui į skaitinę reikšmę TF-IDF metodą, kuris aprašytas 1.11 poskyryje. Analizuojant puslapio turinį HTML žymos pašalinamos, skaičiai ir simboliai, tačiau tekstui nėra pritaikomas normalizavimas ar kokie nors kitokie apdorojimo procesai. Geriausias šiame metode pasiektas 95,38 tikslumas naudojant atraminių vektorių mašininio mokymosi metodą.

6 lentelė. Veiksmų palyginimas su siūlomu vaikų apsaugos nuo žalingo turinio interneto metodo

Kriterijus	„Fuzzy“ ontologijos sprendimas [27]	Sprendimas [35]	Sprendimas [36]	Siūlomas žalingo turinio apsaugos metodas
HTML žymų pašalinimas	Taip	Taip	Taip	Taip
Simbolių ir skaičių pašalinimas	Taip	Taip	Taip	Taip
Nereikšmingų žodžių šalinimas	Taip	Taip	Nėra	Taip
Lemavimas	Taip	Nėra	Nėra	Taip
Metodas teksto vertimui į skaitines reikšmes	„Fuzzy“ skaičiavimai	Doc2vec	TF-IDF	TF-IDF

Kriterijus	„Fuzzy“ ontologijos sprendimas [27]	Sprendimas [35]	Sprendimas [36]	Siūlomas žalingo turinio apsaugos metodas
Kalbos pritaikymas	Anglų	Kinų	Anglų	Lietuvių
Juodieji sąrašai	Yra	Yra	Nėra	Nėra
Pasiektas rezultatas	98%	91,91%	95,38%	-

6 lentelėje pateikiamas esamų puslapių kvalifikacijos veiksmų sekų palyginimas su siūlomu žalingo turinio metodu. Siūlomą metodą lyginant su kelių tyrėjų darbais bus naudojamas teksto vektorizavimas naudojant TF-IDF biblioteką, o ne doc2vec ar „Fuzzy“ metodologijos žodžių prilyginimą į skaitines vertes. Kitas svarbesnis nebuvimas, tai – juodojo sąrašo. Tinklapio domeno įtraukimas į juodąjį sąrašą nereiškia, kad visame puslapyje yra žalingas turinys.

1.9. Teksto klasifikacijos metodų panaudojimas tyrimuose

Svarbiausias teksto klasifikavimo etapas – pasirinkti klasifikatorių. Pagal atliktą tyrimą [32] atraminių vektorių ir k – artimiausių kaimynų algoritmai – plačiausiai naudojami mašininio mokymosi tekstų klasifikavimo srityje. Iš 132 tirtų darbų 88 tyrėjai naudojo šiuos metodus. Taip pat apžvelgiami teksto klasifikavimo būdai, kuriuose daugiausia dėmesio skiriama prižiūrimiems mašininio mokymosi metodams, tokiems, kaip Naivus Bajesas, atraminių vektorių ir k – artimiausių kaimynų [33]. Eksperimentų rezultatai parodė, kad dažnu atveju atraminių vektorių metodo taikymas leido pasiekti ir išgauti didesnę tikslumą.

1.10. Teksto klasifikacijos metodų išvada

Apibendrinant teksto klasifikacijos metodų atliktą analizę 7 lentelėje galime matyti, kad metodai yra identiški ir viskas priklauso nuo pasirinkto duomenų rinkinių, kurie naudojami klasifikacijų kūrimui.

7 lentelė. Klasifikacijos metodų palyginimas

Kriterijus	Naivus Bajesas	Atraminių vektorių	K – artimiausių kaimynų	Neuroninių tinklų
Tikslumas	Tikslus	Tikslus	Tikslus	Tikslus
Klasifikavimas	Pagal klases	Pagal klases	Pagal klases	Pagal klases
Mokymosi mechanizmas	Turi	Turi	Neturi	Turi

Reikėtų atkreipti dėmesį į 1.8.5 poskyrio 5 lentelę, kurioje matosi metodų palyginimas naudojant turinio filtravimui. K – artimiausių kaimynų metodas ir paprasčiausias, ir tiksliausiai išfiltruoja turinį, tačiau naudojant abu metodus Naivus Bajesas ir SVM pasiekiamas 90% tikslumas.

1.11. Teksto klasifikacijos TF-IDF ir skaičiavimo vektorizavimas požymių išskyrimui

Mašininio mokymosi algoritmai išnagrinėti ankstesniame poskyryje reikalauja įvesties skaitinių požymių, todėl tekstas turėtų būti konvertuojamas. Pagrindinis principas teksto pavertimas į žodžių pasikartojimų vektorių. Straipsnyje [30] nagrinėjami du metodai:

- Skaičiavimo vektorizavimas (angl. count vectorization)

- Terminų dažnis atvirkštinis dokumentų dažnis (angl. TF-IDF)

Skaičiavimo vektorizavimas skaičiuoja, kiek kartų žodis pasitaiko dokumente. Tai lemia dažniausią žodžių iškraipymą. Šio metodo trūkumas yra tas, kad reto žodžio svoris – mažas [30]. Tam tikrais atvejais retas žodis gali turėti didelę semantinę reikšmę.

TF-IDF – panašus į skaičiavimo vektorizavimą, tačiau jame įvedamas svorio koeficientas. Jį sudaro dvi dalys: terminų dažnis (angl. Term Frequency - TF) ir atvirkštinis dokumentų dažnis (angl. IDF). TF atsižvelgia į bendrą žodžio paplitimą dokumente. Skirtingų straipsnių dydis skiriasi, gali būti, kad tam tikras žodis pasitaiko dažniau, nes straipsnis yra ilgesnis. Norint tai išspręsti, bendras pasikartojimas dalijamas iš bendro žodžių skaičiaus straipsnyje. Pavyzdys straipsnis a1 priklauso rinkiniui A, kuriame yra 10 straipsnių iš viso sudaro 1000 žodžių. Žodis „trumpas“ pasitaiko 100 kartų, pateikta 3 formulė.

$$TF(trumpas, a1) = \frac{\text{Bendras pasikartojimas}}{\text{Bendras žodžių skaičius}} = \frac{100}{1000} = 0,1 \quad (3)$$

TF visus raktažodžius vertina vienodai, neatsižvelgiant į jų reikšmę ar semantinę svarbą. Norint pabrėžti semantinę žodžio svarbą, naudojamas IDF. Jis atsižvelgia į dažnumą tarp straipsnių. Didesnė reikšmė priskiriama terminams, kurie dažnai pasitaiko dokumente ir yra retesni tarp visų dokumentų. Pateiktame pavyzdyje žodis „trumpas“ pasitaiko 5 iš 10 straipsnių priklausančių A. Žodžio „trumpas“ tarp 10 straipsnių galime apskaičiuoti naudojant 4 formulę.

$$TF(trumpas, A) = \log_e\left(\frac{\text{Straipsniai}}{\text{Straipsniai, kur kartoja žodis}}\right) = \log_e\left(\frac{10}{5}\right) = 0,3 \quad (4)$$

TF-IDF žodžio reikšmę galima apskaičiuoti, kaip TF ir IDF verčių sandaugą (5 formulė).

$$TF - IDF(trumpas, A) = TF(trumpas, a1) * IDF(trumpas, A) = 0,03 \quad (5)$$

Reikšmės didėja proporcingai žodžio atsiradimo dokumente skaičiui ir yra kompensuojamos dokumentų, kuriuose – šis žodis, skaičiumi, o tai leidžia subalansuoti, kad kai kurie žodžiai apskirtai atsiranda dažniau. Teigiama, kad TF-IDF – pranašesnis požymių išskyrimo metodais, tai patikrino ir tyrėjai [30] klasifikuojant melagingų naujienų tekstus ir jų rezultatas taip pat parodė, kad šis metodas davė geresnį rezultatą negu skaičiavimo vektorizavimas.

Teksto klasifikacijos algoritmų apklausoje [31] TF-IDF metodo privalumai išskirti tokie, kaip:

- Lengvas apskaičiavimas,
- Lengva apskaičiuoti panašumą tarp kelių dokumentų,
- Pagrindinė metrika, leidžia išgauti labiausiai aprašomuosius terminus dokumente,
- Bendriniai žodžiai neturi įtakos rezultatams dėl IDF pvz.: esu, yra ir t.t.

Šio metodo trūkumai:

- Nefiksuoja padėties tekste (sintaksinis),
- Neatspindi teksto prasmės (semantinės).

Atsižvelgiant į tai, kad buvo klasifikuota 13470 straipsnių iš bandomosios aibės, bendras TF-IDF ir skaičiavimo vektorizavimo tikslumas – aukštas [31]. Naudojant skaičiaus vektorizavimą, logistinė regresija pasiekia 98,91% tikslumą, o atraminių vektorių – 98,87%. Taikant TF-IDF gaunami geresni rezultatai: atraminio vektoriaus padidėja iki 99,18%, o logistinės regresijos iki 99,13%. Tyrėjo rezultatai patvirtina prielaidą, kad taikant TF-IDF gaunami geresni rezultatai nei naudojant skaičiavimo vektorizavimą.

1.12. Netinkamo turinio filtravimo iššūkiai

Internetas siūlo gyvenimą be ribų, kai visi gali ieškoti informacijos nepriklausomai nuo vartotojo amžiaus. Mums labai svarbu yra suvokti, ar gaunama informacija internete yra tiksli, nepadirbta. Pagrindiniai klausimai filtravimo procese yra tikslumas pateikiant vartotojui informaciją. Internetas gali pateikti įvairius rezultatus, kurie gali turėti blogą poveikį mūsų visuomenei, ypač jaunai kartai, kai paieškos rezultatuose parodomas netinkamas arba iškreiptas turinys [29]. Nepaprasta interneto įvairovės medžiaga, kurios didžioji dalis kuriama arba generuojama dinamiškai socialiniuose tinkluose, svetainėse sukuria praktinius iššūkius turinio ir tinklapių filtravimui.

1.12.1. Filtrų apėjimas

Egzistuoja ir filtrų apėjimas ne iš vartotojo pusės (kliento), bet iš tinklapių ar paslaugų kūrėjų. Pavyzdžiui turinio kūrėjas gali panaudoti statinius tinklapius su veidrodžiu be interaktyvių elementų [29]. Vartotojas bandydamas pasiekti tokio tipo tinklapį kuriame egzistuoja filtras būtų nukreipiamas prie tinklapio veidrodžio. Kitas būdas piktavalių yra pateikiami keli tinklapių adresai, kur egzistuoja blokuotas turinys tame pačiame interneto adrese. Šis būdas neveikia naudojant interneto adresų filtravimą, bet veiksmingas prieš domenų vardų sistemų apsinuodijimą ar įgaliotuosius serverius. Dar vienas būdas, kad turinio kūrėjas gali pastoviai keisti ir atnaujinti interneto adresus. Dinaminis adresų pakeitimas gali sukelti problemų filtruojant tinklapius naudojant juoduosius sąrašus ar net paketų filtravimą.

1.13. Išvada

Atlikus analizę išsiaiškinta kas yra žalinga informacija internete pagal ES teisinius aktus bei LR įstatymus. Paanalizuotos galimos grėsmės vaikams internete, kokią žalingą įtaką gali daryti netinkamas turinys nepilnamečiams. Netinkamo turinio filtravimas bandant šalinti visas grėsmes vaikams internete, naudojant vieną metodą būtų sudėtingas arba neužtikrintų grėsmių pašalinimo.

Atlikus esamų tėvų kontrolės sprendimų analizę pastebėta, kad dabartinių sistemų didžiausias trūkumas yra įrenginių limitavimas ir išskylantis pavojus, kad atsijungus nuo namų tinklo t. y. perjungus internetą turinys nebebus filtruojamas, jeigu naudojamas DNS serveris maršrutizatoriuje.

Analizuojant galimus sistemų būdus vienas iš jų būtų domenų vardų sistemos mechanizmas. Pateiktas architektūros pavyzdys leistų užtikrinti tinkamą pradinį turinio filtravimą visuose įrenginiuose esančiuose namų tinkluose arba papildomai įdiegiant įrenginius. Tačiau reikėtų atkreipti dėmesį, kad tokio tipo filtravimas naudoja juoduosius sąrašus, kurie yra aprašomi pačio vartotojo, todėl tinklapiai ne visada būtų tinkamai išfiltruojami, o ir ne visas turinys tinklapyje būna pavojingas nepilnamečiams.

Pasirinkta išanalizuoti teksto klasifikavimo metodus, kurie leistų filtruoti tinklapiuose esantį turinį, naudojant mašininį mokymąsi. Tikslas pagal duomenų rinkinius sugrupuoti tinklapius į kategorijas,

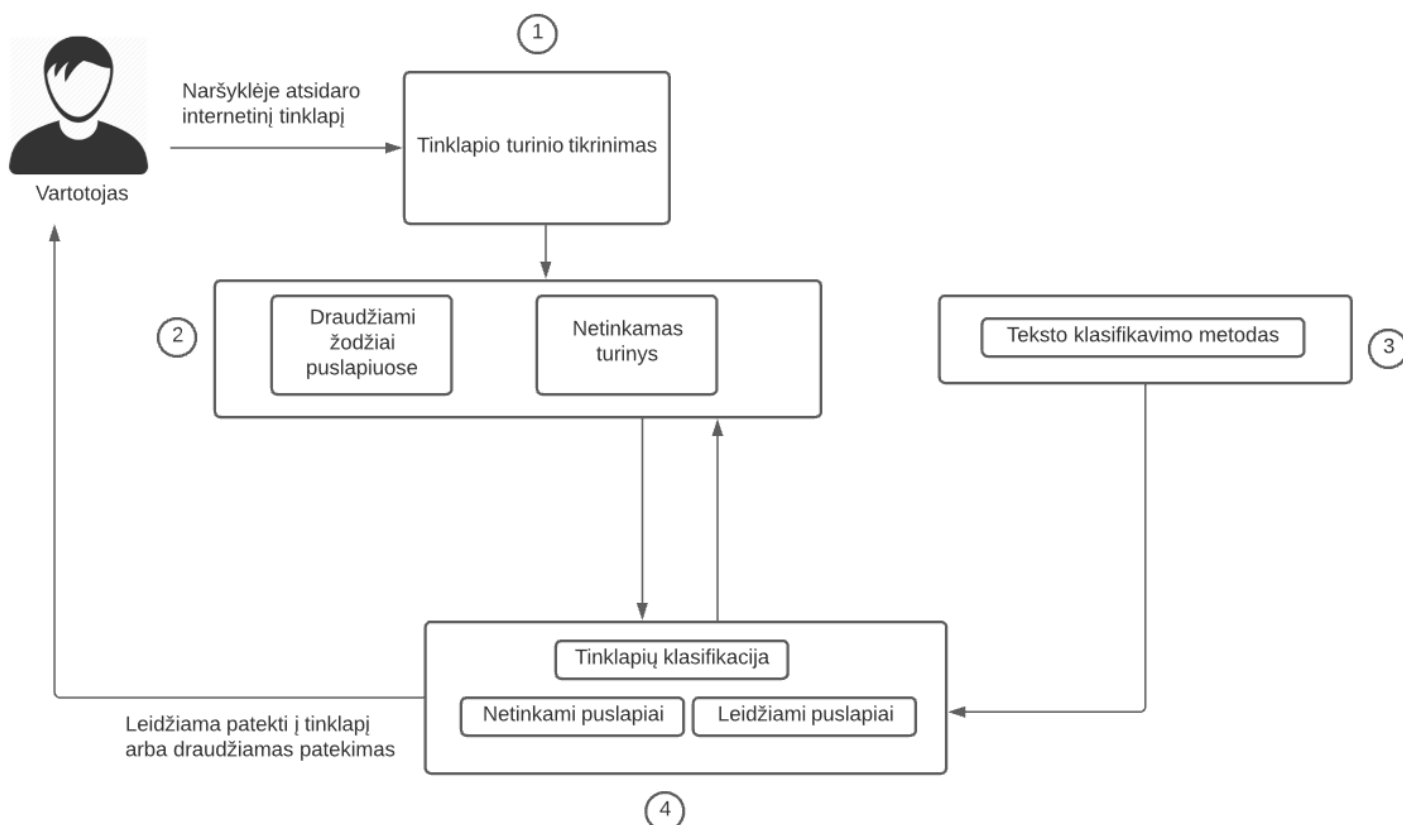
kurie būtų filtruojami pagal atidaryto puslapio turinį. Išanalizuoti trys esami klasifikacijos metodai, kurie palyginti su siūlomu žalingo turinio atpažinimo metodu. Teksto pavertimui į skaitinę reikšmę pasirinkta naudoti TF-IDF vektorizavimo metodą. Tyrimais ir bandymais yra įrodyta [26, 27], kad dažniausiai vieno mechanizmo neužtenka tinkamai užtikrinti, kad bus išfiltruojamas turinys.

2. Žalingo turinio atpažinimo metodas

Šiame skyriuje aprašomas metodo projektas, kuris apima kuriamo metodo viziją (konceptiją), sistemos reikalavimų analizę (panaudojimo atvejų diagramas, funkcinius ir nefunkcinius reikalavimus).

2.1. Žalingo turinio atpažinimo metodo koncepcija

Pagrindinis šio tyrimo tikslas yra pasiūlyti metodą, kuris automatiškai aptiktų ir nustatytų suaugusiųjų turinį – pornografinį ir azartinį. Netinkamų tinklapių aptikimas pagrįstas internetinių tinklapių turinio analizės pagrindu. Šiame skyriuje siūlomo metodo koncepcija pateikiama 9 pav.



9 pav. Žalingo turinio atpažinimo metodo koncepcija.

Netinkamo turinio tinklapiai bus įtraukiami į draudžiamų puslapių sąrašą. Jeigu tinklapis jau yra draudžiamų puslapių sąrašė tada nereikia tinklapio turinio klasifikuoti ir iškart tinklapis vartotojui yra užblokuojamas. Jeigu šiame sąrašė nėra tinklapio, turinys klasifikuojamas naudojant teksto klasifikavimo atraminių vektorių metodą. Šiame etape visas tinklapio turinys – apdorojamas dalimis (angl. scrape) ir analizuojami jame esantys žodžių junginiai. Pagal paruoštą duomenų rinkinį – netinkamų žodžių ir paveikslukų tinklapis yra įtraukiamas į draudžiamų puslapių sąrašą. Jeigu tinklapyje aptinka netinkamą žodį, kaip pavyzdys lažybos, statymai tinklapis įtraukiamas į netinkamų puslapių kategoriją. Tokiu principu tinklapis klasifikuojamas į netinkamą arba leidžiamą puslapį.

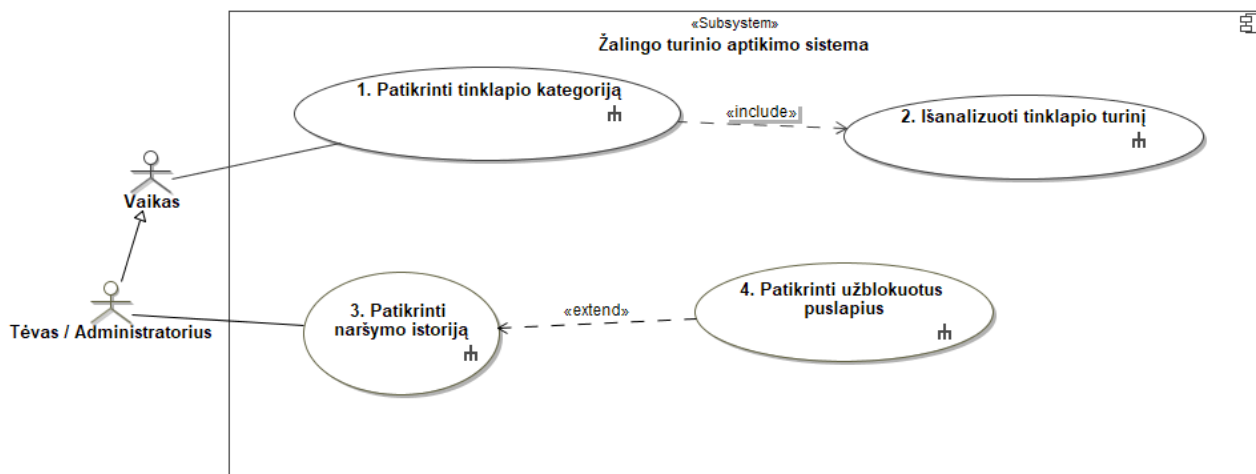
2.1.1. Žalingo turinio metodo reikalingumo pagrindimas

Vienas iš paprastesnių apsaugos metodų nuo žalingo turinio internete – juodieji sąrašai. Sudaromas puslapių sąrašas, kurie būtų užblokuojami ir draudžiami vaiko įrenginyje. Šiuolaikiniame pasaulyje, kai puslapių kiekvieną dieną atsiranda naujų arba piktavaliai naudoja veidrodinius puslapius (nukreipiamuosius), juodieji sąrašai ar statiniai filtrai ne visus puslapius ar turinį identifikuoja, kaip pavojingus. Taip pat daugėjant puslapių kiekiui turėtų būti atnaujinami, pildomi juodieji sąrašai.

Atliktoje analizėje esamų sprendimų naudojant mašininio mokymosi algoritmus tekstų klasifikacijai 1.8.5 poskyryje išskirtos savybės, kuo pasižymi sprendimai. Pagrindinis skirtumas tarp analizuotų ir siūlomo sprendimo – dviejų kalbų suderinamumas. Tai leidžia apsisaugoti, kai vaikas savo įrenginyje ieško puslapių ne tik lietuvių k., bet ir anglų k. Kadangi teksto klasifikacijos geresniems rezultatams pasiekti – naudojamas teksto normalizavimas, kalbos skirtumai ir žodžiai lemia, teksto apdorojimo ypatybės tokias, kaip žodžių lematizavimą ir dažniausiai pasikartojančių žodžių sąrašą. „Fuzzy“ ontologija [27] autoriui leido pasiekti didžiausią rezultatą, tačiau nėra žinoma kokius parametrus ir duomenų rinkinį naudojo klasifikuojant tekstus pagal kategorijas. Kitas skirtumas didžioji dalis analizuotų esamų sprendimų vis dar naudoja juoduosius sąrašus, kadangi siūlomas metodas analizuoja atidarinęjamą puslapį iš karto, nebėra tikslo sistemoje papildomai saugoti juodąjį sąrašą.

2.2. Žalingo turinio metodo reikalavimų specifikacija

Šiame poskyryje aprašoma vaikų apsaugos nuo žalingo turinio internete realizuojamo prototipo projektas, kuris apima sistemos reikalavimų analizę.



10 pav. Žalingo turinio atpažinimo internete sistemos prototipo panaudos atvejų diagrama.

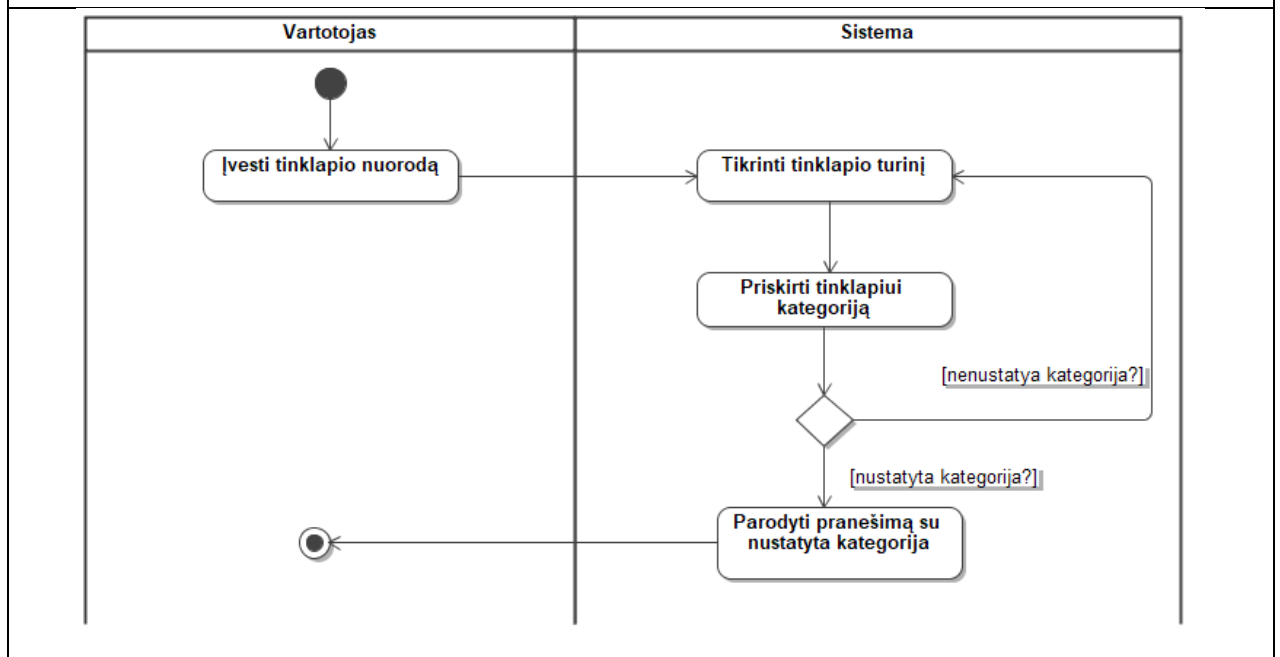
Vaikas – pagrindinis sistemos aktorius, kuris turi prieigą prie apsaugos nuo žalingo turinio sistemos per interneto naršyklę. Sistemos naudotojas įvesdamas tinklapio nuorodą gali patikrinti kokiai kategorijai įvestas tinklapis yra priskiriamas, ar jis neturėtų būti blokuojamas, leidžiama peržiūrėti užblokuotų tinklapių istoriją, kurie buvo įvesti anksčiau tikrinimui. Tėvams (administratoriams) leidžiama peržiūrėti esamus filtrus (duomenų rinkinius).

Žemiau lentelėse pateikiami detalūs kiekvieno panaudos atvejo aprašymai.

8 lentelė. Patikrinti tinklapio kategoriją panaudos atvejo aprašymas

Panaudojimo atvejis	1	Patikrinti tinklapio kategoriją
Aprašymas	Funkcija leidžianti vartotojui nustatyti tinklapio kategoriją pateikiant tinklapio nuorodą. Nustatoma ar turinys esantis puslapyje – žalingas (azartinis, pornografinis).	
Aktoriai	Vartotojas	
Prieš sąlyga	Vartotojas turi būti atsidaręs sistemos sąsają.	
Po sąlyga	Sistemoje atlikus tinklapio turinio analizę – pateikiama kategorija, kuriai tikrinamas internetinis puslapis priklauso.	
Pagrindiniai žingsniai	Įvedama puslapio nuoroda arba pateikiama nuoroda, kaip kintamasis funkcijoje ir paleidžiama programa.	

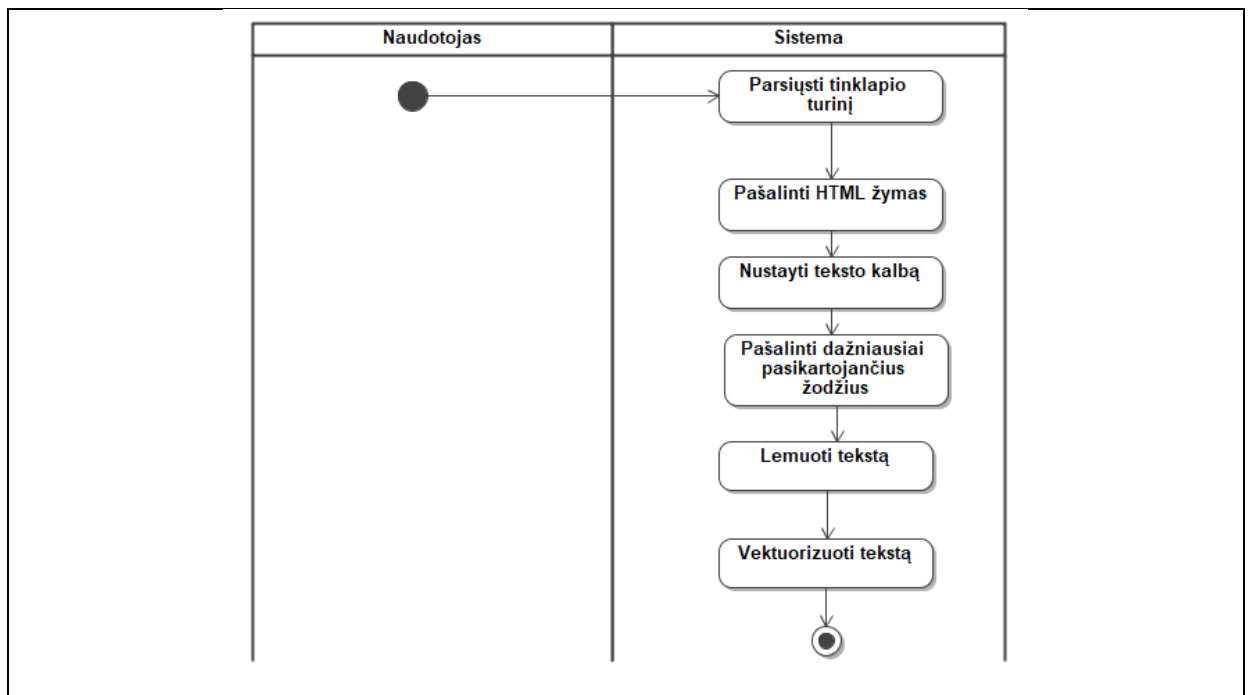
Veiklos diagrama



9 lentelė. Patikrinti tinklapio kategoriją panaudos atvejo aprašymas

Panaudojimo atvejis	2	Išanalizuoti tinklapio turinį
Aprašymas	Funkcija papildanti PA-1. Šis antrinis procesas leidžia sistemai pasirinktam puslapiui atlikti turinio analizę – išrinkti iš puslapio tekstą ir jį apdoruoti paruošiant mašininio mokymosi algoritmams.	
Aktoriai	Vartotojas	
Prieš sąlyga	Vartotojas turi būti atsidaręs sistemos sąsają ir pasirinkęs tikrinti tinklapio kategoriją.	
Po sąlyga	Sistema iš pasirinkto puslapio iškrapšto tekstą ir jį apdoroja mašininio mokymosi algoritmui.	
Pagrindiniai žingsniai	Įvedama puslapio nuoroda arba pateikiama nuoroda, kaip kintamasis funkcijoje ir paleidžiama programa.	

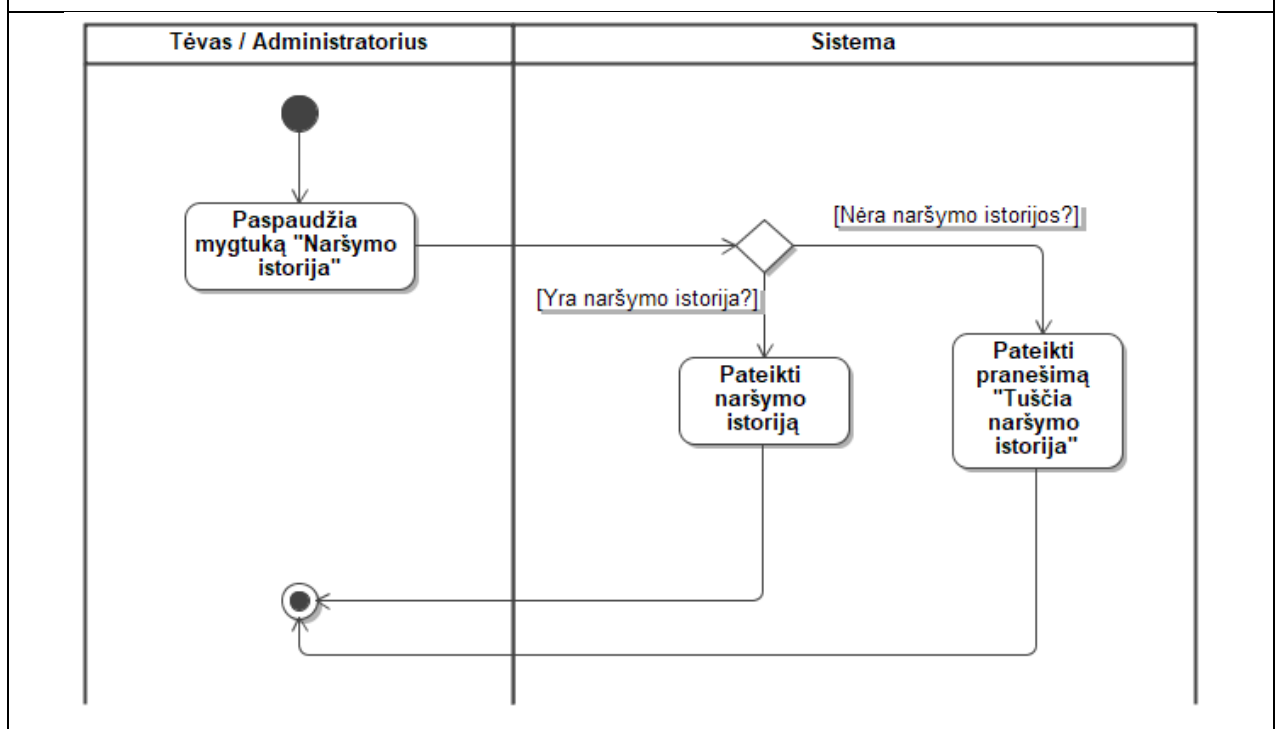
Veiklos diagrama



10 lentelė. Patikrinti naršymo istoriją panaudos atvejo aprašymas

Panaudojimo atvejis	3	Patikrinti naršymo istoriją
Aprašymas	Funkcija leidžianti tėvams arba sistemos administratoriui patikrinti vaiko naršymo istoriją, kokius puslapius vaikas atsidarė savo įrenginyje.	
Aktoriai	Tėvas, administratorius	
Prieš sąlyga	Tėvas turi būti atsidaręs vaikų apsaugos nuo žalingo turinio internete sąsają.	
Po sąlyga	Žalingo turinio apsaugos sistema pateikia tėvui, vaiko įrenginio puslapių naršymo istoriją, kuriuos savo įrenginyje vaikas bandė atsidaryti.	
Pagrindiniai žingsniai	Žalingo turinio apsaugos sistemos sąsajoje paspaudžiamas mygtukas „Naršymo istorija“.	

Veiklos diagrama



11 lentelė. Patikrinti užblokuotus puslapius panaudos atvejo aprašymas

Panaudojimo atvejis	4	Patikrinti užblokuotus puslapius
Aprašymas	Funkcija leidžianti tėvams arba sistemos administratoriui patikrinti užblokuotus puslapius, kuriuos vaikai bandė atsidaryti, bet vaikų apsaugos nuo žalingo turinio sistema identifikavo, kaip nesaugius.	
Aktoriai	Tėvas, administratorius	
Prieš sąlyga	Tėvas turi būti atsidaręs vaikų apsaugos nuo žalingo turinio internete sąsają ir paspaudęs mygtuką „Naršymo istorija“.	
Po sąlyga	Žalingo turinio apsaugos sistema pateikia tėvui užblokuotus puslapius, kuriuos vaikas bandė atsidaryti savo įrenginyje, tačiau žalingo turinio sistema juos identifikavo, kaip nesaugius ir vaiko įrenginyje tapo nepasiekiami.	
Pagrindiniai žingsniai	Įvedama puslapio nuoroda arba pateikiama nuoroda, kaip kintamasis funkcijoje ir paleidžiama programa.	

Veiklos diagrama

Tėvas / Administratorius

Pasirinkti filtrą "Užblokuoti puslapiai"

Sistema

[Užblokuotų puslapių nėra?]

Pateikti pranešimą "Užblokuotų puslapių nėra"

[Yra užblokuotų puslapių?]

Pateikti užblokuotus puslapius

2.2.1. Funkciniai reikalavimai

12 lentelė. Funkcinių reikalavimų lentelė

Nr.	Reikalavimas	Susiję panaudos atvejai	Sudėtingumas
FR-01	Vartotojas turi galėti nustatyti tinklapio kategoriją.	PA-01	Sudėtingas
FR-02	Tinklapio turinys turi būti iškrapšytas be HTML žymų.	PA-02, PA-01	Sudėtingas
FR-03	Tinklapio turinio tekstas turi sudaryti teksto žodžiams lemas.	PA-02, PA-01	Vidutinis
FR-04	Lemuotas tekstas turi būti vektorizuojamas naudojant TfIdfVectorizer biblioteką.	PA-02, PA-01	Vidutinis
FR-05	Tėvai arba administratorius gali patikrinti naršymo istoriją.	PA-03	Vidutinis
FR-05	Tėvai arba administratorius peržiūrėdami naršymo istoriją gali išfiltruoti puslapius, kurie buvo užblokuoti apsaugos nuo žalingo turinio sistemos.	PA-03, PA-04	Vidutinis

2.2.2. Nefunkciniai reikalavimai

13 lentelė. Nefunkcinių reikalavimų lentelė

Nr.	Reikalavimas	Reikalavimų grupė	Panaudos atvejis	Sudėtingumas
NR-01	Sudėtingesni programinės įrangos skaičiavimai turi rodyti kiek laiko užtruko atliktas veiksmas.	Programinei įrangai	-	Nesudėtingas
NR-02	Įrenginiai turi būti prijungti prie interneto.	Aparatūrai	-	Nesudėtingas
NR-03	Įvairiose įrenginiuose apsaugos nuo žalingo turinio sistema turi veikti sklandžiai.	Aparatūrai	-	Vidutinis

2.3. Žalingo turinio aptikimas puslapiuose

2.3.1. Netinkamų žodžių atpažinimas tinklapiuose

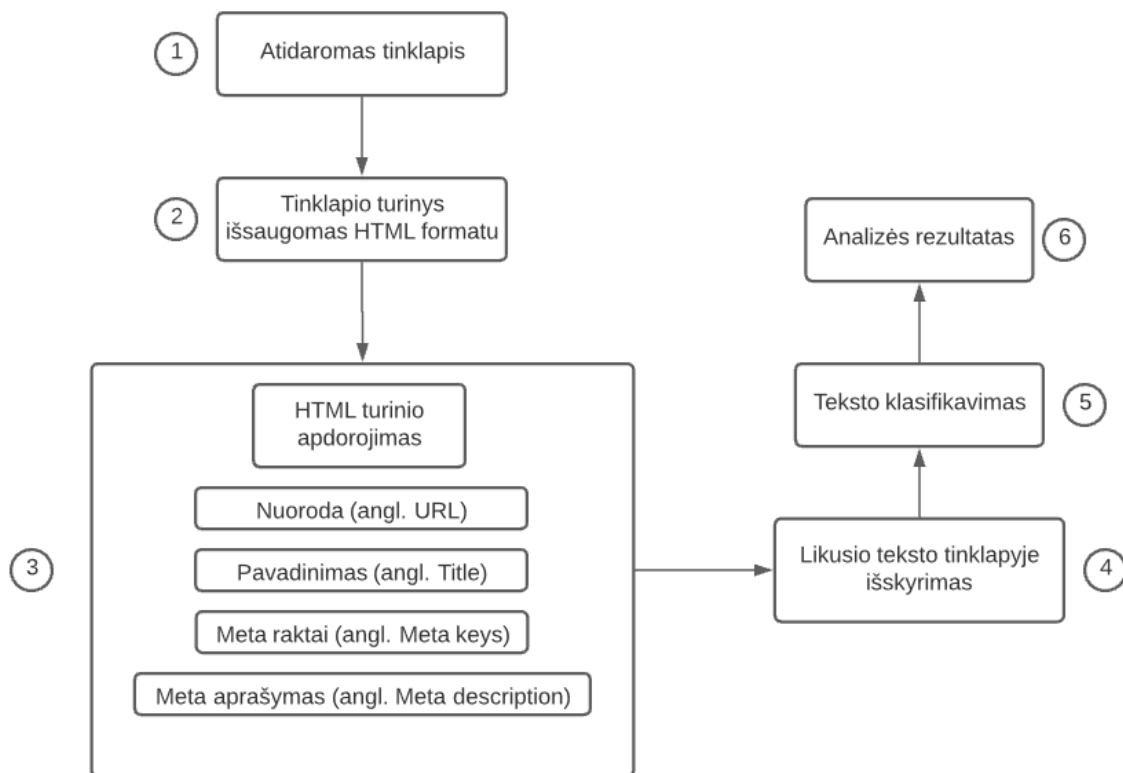
Tinklapių klasifikavimo metu esantis turinys – tekstui bus naudojamas žodžių kanonizavimas (angl. canonicalization) – transformacija, kurios metu žodžiai bus normalizuojami ir gali būti paversti į standartinę formą. Žodžių kanonizavimo būdai:

- Rašybos taisymas.
- Žodžio raidžių suvienodinimas.
- Šaknies nustatymas (angl. stemming), kurio metu pašalinama žodžio galūnė.
- Lemavimas (angl. lemmatization), kurio metu žodis yra paverčiamas į pagrindinę jo formą, pavyzdžiui bendratį, jei žodis buvo veiksmažodis arba jeigu buvo daiktavardis paverčiama į vienaskaitos vardininko linksį.

Pagrindinė paskirtis būtų tokia, kad analizuojant tekste žodžius būtų lengviau identifikuoti ir sutapatinti su atrinktais žodžiais ar jų junginiais kurie neturėtų būti naudojami.

2.3.2. Tinklapių klasifikavimas pagal sudarytą duomenų modelį

Vartotojui įvedus tinklapį į sistemą jis yra atidaromas ir jo statinis tinklapių turinys parsisiunčiamas HTML formatu. Toliau tinklapių apdorojimas vyksta šiame formate ir yra išskaidomos pagrindinės dalys tinklapių – pilna tinklapių nuoroda, pavadinimas ir visi kiti meta duomenys. Likusi tinklapių dalis analizuojami be žymų, tik kaip tekstas. Teksto analizavimas vyksta pagal aprašytą procesą 2.3.1 skyriuje. 11 pav. pavaizduotas visas supaprastintas procesas tinklapių apdorojimo.



11 pav. Tinklapių turinio identifikavimo procesas.

Analizės rezultate pateikiama vartotojui informacija apie tinklapį, priskirtą kategoriją tinklapiui, ar jo turinys nebūtų draudžiamas ir ar jis nebūtų užblokuotas.

```

<!DOCTYPE html>
<html lang="lt">
<head>
<meta charset="utf-8">
<meta http-equiv="X-UA-Compatible" content="IE=edge"><script type="text/javascript">(window.NREUM||(NREUM={})).loader_config={licenseKey:"287ceeab32",applicationID:"99361373"};window.NREUM||(NREUM={});
<meta name="theme-color" content="#000000">
<meta name="apple-mobile-web-app-status-bar-style" content="black">
<meta name="application-name" content="www.topsport.lt" />
<meta name="msapplication-TileColor" content="#000000" />
<link rel="shortcut icon" type="image/x-icon" href="https://cdn.topsport.lt/favicon.ico">
<link rel="preconnect" href="https://www.google.com" crossorigin>
<link rel="preconnect" href="https://www.google.lt" crossorigin>
<link rel="preconnect" href="https://www.google-analytics.com" crossorigin>
<link rel="preconnect" href="https://www.googleadmanager.com" crossorigin>
<link rel="preconnect" href="https://www.googleadservices.com" crossorigin>
<link rel="preconnect" href="https://googleads.g.doubleclick.net" crossorigin>
<link rel="preconnect" href="https://fonts.gstatic.com" crossorigin>
<link rel="preconnect" href="https://www.facebook.com" crossorigin>
<link rel="preconnect" href="https://connect.facebook.net" crossorigin>
<link rel="preconnect" href="https://myrtmark.net" crossorigin>
<link rel="preconnect" href="https://track.adform.net" crossorigin>
<link rel="preconnect" href="https://s2.adform.net" crossorigin>
<link rel="preconnect" href="https://dmp.adform.net" crossorigin>
<link rel="preconnect" href="https://vars.hotjar.com" crossorigin>
<link rel="preconnect" href="https://script.hotjar.com" crossorigin>
<title>Lažybos internetu kiekvieną dieną ❏ | TOPsport statymai</title>
<meta name="description" content="Lažybos sporo įvykiuose visą parą! ➔ Ateik ir pasiimk savo nemokamą statymą be rizikos, kurio pralaimėti negali! ➔ Patinka Lažybos internetu? Rinkis TOPsport statymus!</meta>
<link href="https://www.topsport.lt/laizyba" rel="canonical" />
<meta name="og:title" content="Lažybos internetu kiekvieną dieną ❏ | TOPsport statymai" />
<meta name="og:description" content="Lažybos sporo įvykiuose visą parą! ➔ Ateik ir pasiimk savo nemokamą statymą be rizikos, kurio pralaimėti negali! ➔ Patinka Lažybos internetu? Rinkis TOPsport statymus!</meta>
<meta name="og:image" content="/assets/app/static/images/logo/topsport.png" />
<meta name="og:type" content="website" />
<meta name="twitter:title" content="Lažybos internetu kiekvieną dieną ❏ | TOPsport statymai" />
<meta name="twitter:description" content="Lažybos sporo įvykiuose visą parą! ➔ Ateik ir pasiimk savo nemokamą statymą be rizikos, kurio pralaimėti negali! ➔ Patinka Lažybos internetu? Rinkis TOPsport statymus!</meta>
<meta name="twitter:image" content="/assets/app/static/images/logo/topsport.png" />
<meta name="twitter:card" content="summary" />
<meta name="twitter:site" content="@topsportbet" />
<meta content="width=device-width, initial-scale=1, maximum-scale=1, user-scalable=no" name="viewport">
<link rel="stylesheet" href="https://cdn.topsport.lt/assets/app/app.78a25fe8.css">
</head>
<body class="page-prelive page-prelive-landing" data-locale="lt">
<script>dataLayer = [{"loggedIn":false,"route":"application_prelive_default_betting","path":"/odds-overview"}];</script>

```

12 pav. Tinklapių HTML vaizdas.

Tinklapių turinio apdorojimas, kuris pateiktas 11 pav. 3 punkte vyksta tinklapių analizė HTML formatu, 12 pav. matome, kaip yra atidaromas tinklapis šiame formate. 3 punkte išskiriamos HTML žymos pavadinimas ir meta raktai, kurie aprašomi tokiu būdu: <title>Tinklapių pavadinimas </title> <meta name="description" content="Tinklapių aprašymas/>

<meta name="og:title" content="Tinklapio pavadinimas" />

<meta name="og:description" content="Tinklapio aprašymas"/>

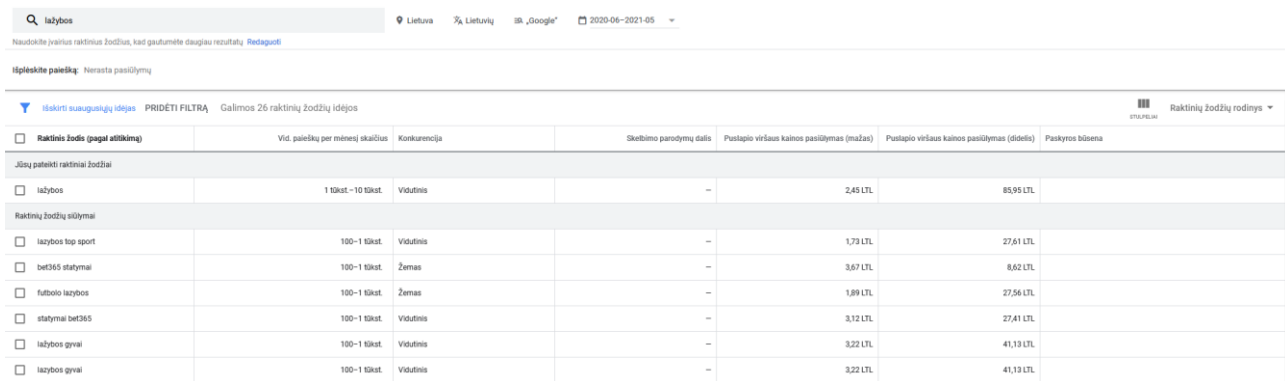
Šiose žymose atspindi tinklapio turinys, tačiau gali pasitaikyti ir taip, kad likęs žalingas turinys bus ne žymose, o likusiame tinklapio turinyje, todėl būtina išskirti visą <body></body> žymą, kaip tekstą be likusių žymų tokių, kaip <div id="header"></div> ar Nuoroda ir atlikti teksto analizę.

2.3.3. Žalingo turinio metodo apmokymų duomenų rinkinys

Ankstesniame skyriuje 1.11 nuspręsta, kad bus bandoma filtruoti azartinius ir pornografinius tinklapius, todėl tinklapių turinys bus skirstomas į tris kategorijas (klases):

- Normalus turinys,
- Azartinis turinys,
- Suaugusių turinys.

Pagal atliktą analizę suaugusių turinys bus priskiriamas toks, kuriame yra medžiagos apie seksualinę prievartą, erotinio pobūdžio informaciją. Pagal turinio pobūdį galime išskirti žodžius, kuriuos bandysime ieškoti tinklapio turinyje ir pagal tai nuspręsti ar tinklapis bus įtraukiamas į kategoriją suaugusių turinys. Pagrindiniai tokio pobūdžio turinio žodžiai: pornografija, rasizmas, seksas, nuogumas, diskriminacija. Azartinio turinio žodžių išskyrimui panaudotas „Google“ raktinių žodžių planavimo įrankis, kuris pateiktas 13 pav.. Žodžiui lažybos automatiškai parenkama virš 26 raktinių žodžių, tokių, kaip lažybos, lažybos top sport, futbolo lažybos. Įvedus žodį kazino pasiūloma 100 raktinių žodžių, o statymai 34 raktiniai žodžiai, kurie gali būti panaudoti identifikuojant azartinį turinį.



Raktinis žodis (pagal atlikimą)	Vid. paieškų per mėnesį skaičius	Konkurencija	Skelbimo parodymų dalis	Postapio viršaus kainos pasiūlymas (mažas)	Postapio viršaus kainos pasiūlymas (didelis)	Paskyros biudžetas
lažybos	1 tūkst.–10 tūkst.	Vidutinis	–	2,45 LTL	89,95 LTL	
lažybos top sport	100–1 tūkst.	Vidutinis	–	1,79 LTL	27,61 LTL	
bet365 statymai	100–1 tūkst.	Žemas	–	3,67 LTL	8,62 LTL	
futbolo lažybos	100–1 tūkst.	Žemas	–	1,89 LTL	27,56 LTL	
statymai bet365	100–1 tūkst.	Vidutinis	–	3,12 LTL	27,41 LTL	
lažybos gval	100–1 tūkst.	Vidutinis	–	3,22 LTL	41,13 LTL	
lažybos gval	100–1 tūkst.	Vidutinis	–	3,22 LTL	41,13 LTL	

13 pav. „Google“ raktinių žodžių planavimo įrankis.

Normaliam turiniui bus priskiriamas visas likęs turinys, kuris nepatenka į suaugusių ir azartinio pobūdžio turinį.

2.3.4. Žalingo turinio metodo duomenų rinkinio sudarymo metodas

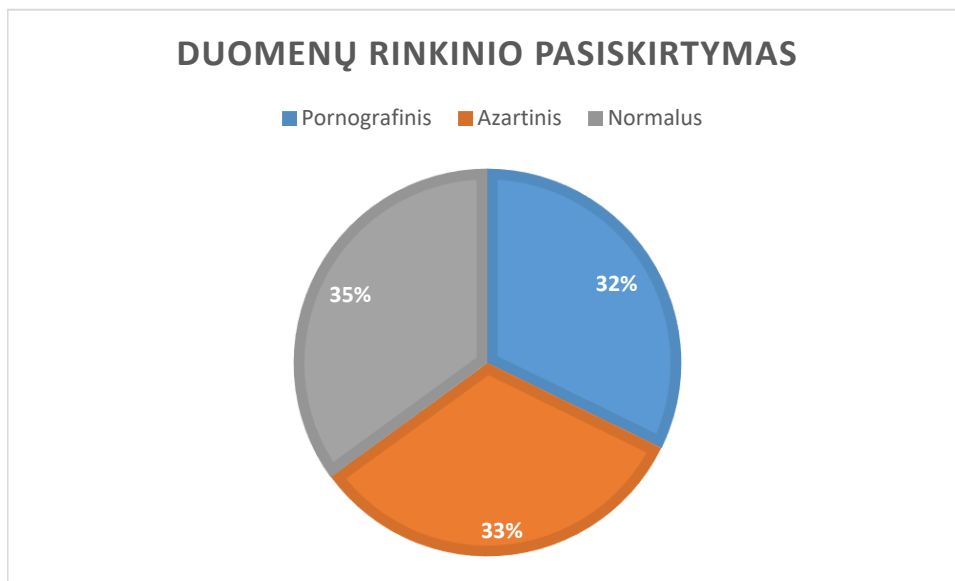
Duomenų rinkinį sudarys X nuorodų su 3 atributais. Pagrindiniai atributai, kurie bus naudojami mašininiam mokymuisi – tinklapio nuoroda (url) ir tinklapio pagrindinė kategorija (main_category).

- id
- url – atributas nurodo tinklapio adresą,

- `main_category` – atributas aprašo pagrindinę tinklapio kategoriją

Pasirenkamas raktažodis, kuris įvedamas į „Google“ raktinių žodžių planavimo įrankį, kuris pateiktas 12 pav. ir pagal jį parenkama papildomų raktinių žodžių pagal kuriuos galima atrinkti tinklapius pagal reikiamą kategoriją. Kiekvienas puslapis – atidaromas ir patikrinamas jo turinys (pobūdis) ir priskiriama atitinkama kategorija duomenų rinkinio faile.

Duomenų rinkinys sudarytas iš 400 puslapių. Sudaryto duomenų rinkinio pasiskirstymas pagal kategorijas pateikiamas 14 pav.

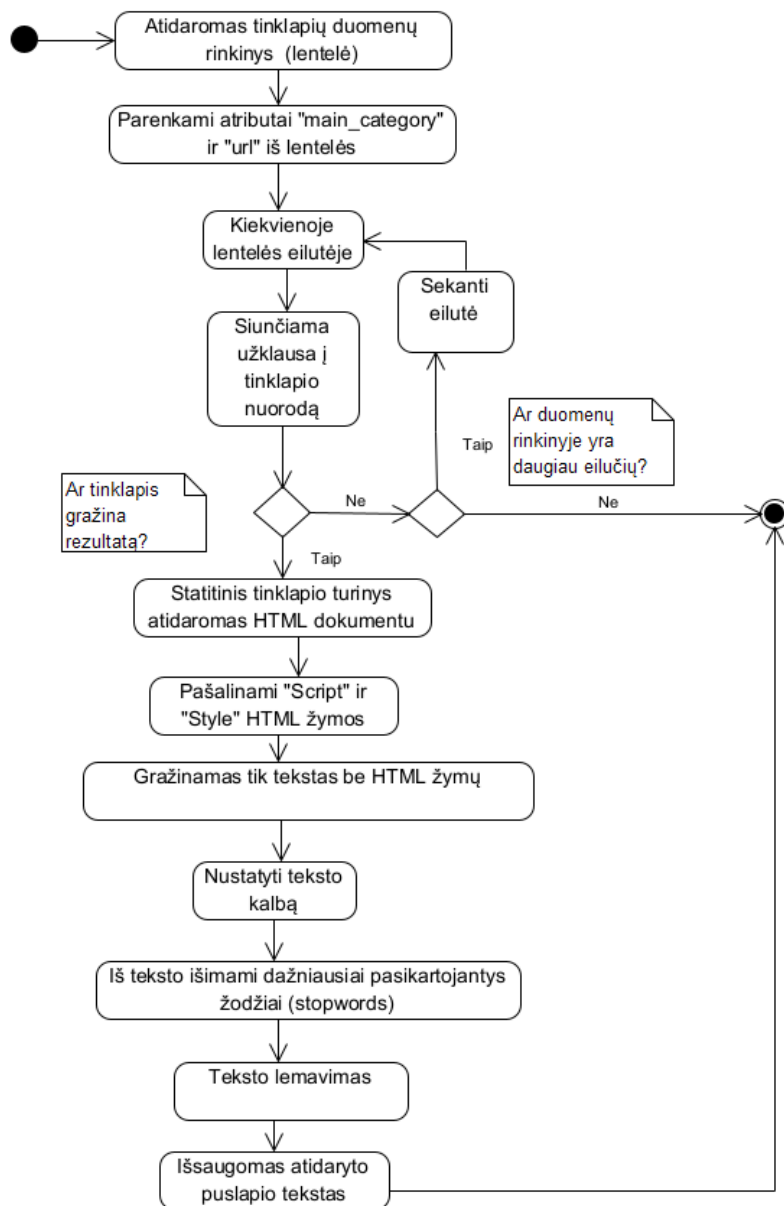


14 pav. Duomenų rinkinio pasiskirstymas pagal kategorijas.

Pirminių duomenų nepakanka norint nustatyti, kuriai kategorijai būtų priskirta tinklapio nuoroda, todėl reikalingas duomenų apdorojimas, siekiant sukurti mašininio mokymosi algoritmų vektorių ir etikečių rinkinį. Išankstinį duomenų apdorojimą sudaro 4 pagrindiniai veiksmas:

- Statitinio puslapio HTML vaizdo apdorojimas ir teksto išskyrimas,
- Teksto išskyrimas atskirais žodžiais,
- Dažniausiai naudojamų žodžių pašalinimas (angl. stop words) ir teksto normalizavimas.

Visas pilnas prototipo išankstinio duomenų paruošimo procesas pavaizduotas 15 pav.



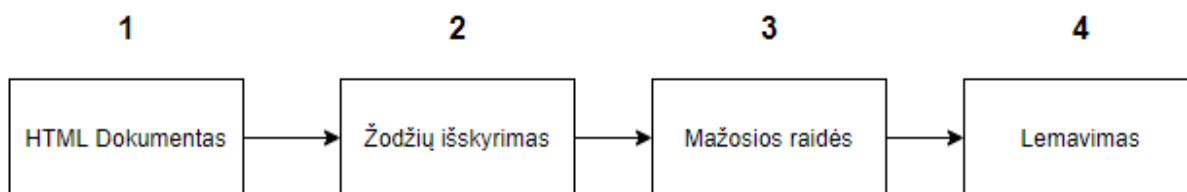
15 pav. Veiksmų seka iš ankstinio duomenų paruošimo

Pirmasis procesas – atidaryto puslapio svetainės analizė ir jos turinio atsisiuntimas. Kiekviena svetainė, sukurta naudojant HTML žymėjimo kalbą. Svetainės atsakyme grąžinamas HTML kodas, kuriame saugomas visas tekstas. Kadangi klasifikacijai tekstas nėra reikalingas, kad jis būtų HTML žymose, todėl reikia jį išsifiltruoti be žymų. Tam naudojama „Python“ biblioteka „beautifulsoup“. Pilną teksto analizę prototipe sudaro šie veiksmai:

- 1) Naudojant „beautifulsoup“ biblioteką tekstas yra ištraukiamas iš nurodyto tinklapio.
- 2) Išfiltruojamas CSS ir Javascript turinys. HTML dizainas būna pažymėtas <style> HTML žyma, o Javascript <script> žymomis, šiose žymose dažniausiai nebūna jokio reikalingo turinio ar naudingos informacijos, todėl šios žymos – pašalinamos panaudojant „beautifulsoup“ bibliotekos funkciją `decompose()`.
- 3) Kadangi rezultatas vis dar – su HTML žymomis, visos kitos likusios žymos pašalinamos su „beautifulsoup“ bibliotekos funkcija `get_text()`.

3. Žalingo turinio atpažinimo metodo realizacija

Pagrindinis prototipo tikslas – aptikti nurodyto tinklapio kategoriją. Supaprastintas išankstinis tinklapio turinio apdorojimas pavaizduotas 16 pav.



16 pav. Išankstinis tinklapio turinio apdorojimo procesas

Atsižvelgiant į tinklapio HTML dokumentą, išankstinio apdorojimo žingsnis išskiria žodžių lemavimo vektorius. Natūralios kalbos apdorojimo kontekste žodis lemavimas arba antraštė – kanoninė forma susijusių žodžių rinkinys. Išankstinio apdorojimo proceso metu nuskaitomas neapdorotas statinio tinklapio HTML failo tekstas, po kurio atliekami veiksmai, kurie leidžia panaudoti duomenų rinkinį mašininio mokymosi algoritmams.

Prototipo realizacijos tikslas išspręsti vaikams žalingo turinio tinklapių nuorodos klasifikavimo metodą, kuris leistų identifikuoti tinklapio kategoriją.

3.1. Žalingo turinio atpažinimo metodo priemonės

- 1) Metodui reikalinga programinė įranga:
 - a) Python 3.9.5 ar naujesnė versija, palaikanti Scikit-learn, NLTK ir Spacy bibliotekas.
- 2) Sudarymui duomenų rinkinio reikalinga papildoma įranga:
 - a) Įrenginys leidžiantis išsaugoti atrinktus žalingus puslapius pagal aprašytą poskyryje 2.3.3 metodologiją.
 - b) Programinė įranga: Python 3.9.5, ar naujesnė versija palaikanti BeautifulSoup ir Urllparse bibliotekas.

3.2. Išankstinis duomenų apdorojimas prototipe

Duomenų rinkinyje bus 3 kategorijos, kurios buvo aprašytos 2.3.3 poskyryje. Duomenų rinkinio pavyzdys pavaizduotas 14 lentelėje, sudarytas naudojant ankstesniame poskyryje aprašytą būdą. Pateiktoje lentelėje atsispindi pagrindinės kategorijos.

14 lentelė. Atrinktų puslapių pavyzdys su kategorijomis

Id	Url	Main_category
1	https://www.topsport.lt/lazybos-gyvai-beta	Azartinis
2	https://www.betsafe.lt/lt/lazybos-gyvai	Azartinis
3	https://www.safesex.lt/pagrindinis/irasas/39_kaip-gerai-atlikti-oralini-vaikinui-.html?page_type=post	Pornografinis
4	https://seksualios.com/nuoga-zyma/nuoga-mergina	Pornografinis
5	https://www.statymai.com	Azartinis
6	http://www.technologijos.lt	Normalus
7	https://bigsport.lt	Normalus

8	http://www.putytes.lt/	Pornografinis
---	---	---------------

Išskrapštytas tekstas iš atidaryto puslapio panaikinus HTML žymas pateikiamas 15 lentelėje. Šiame etape dar nėra pašalinti dažniausiai pasikartojantys žodžiai ir tekstas nėra papildomai apdorotas mašininio mokymosi algoritmais.

15 lentelė. Išanalizuoto puslapio tekstas panaikinus HTML žymas

Id	1
Url	https://www.topsport.lt/lazybos-gyvai-beta
Main_category	Azartinis
Text	Lažybos gyvai + TV transliacijos <u>TOPsport statymai</u> TOPsport yra ne tik lažybų punktas - tai moderniausia ir didžiausia pramogų erdvė bei populiariausia lažybų internetu + TV platforma visoje Lietuvoje. <u>TOP SPORT LAŽYBOS GYVAI + TV - REZULTATAI REALIU LAIKU TV TIESIOGIAI TOP SPORT LAŽYBOS GYVAI + TV - KATEGORIJOS TOPsport</u> Pagalba <u>TOPsport</u> remia Šis puslapis reikalauja JavaScript palaikymo! Draudžiama lošti nuotolinius stalo lošimus ir lošimus a kategorijos automatais asmenims, kuriems nėra sukakę 21 metų. Draudžiama lošti asmenims, kuriems nėra sukakę 18 metų. Ispėjame: neatsakingas lošimas gali tapti priklausomybės nuo azartinių lošimų priežastimi. Prisijungti Nepavykta prisijungti? Spauskite čia! Registruotis English <u>Русский Sportas Gyvai Gyvai + Kazino Kazino gyvai Lažybų žaidimai Lenktynės E-sportas Punktai Blogas Sportas Gyvai Gyvai + Kazino Kazino gyvai Lažybų žaidimai Lenktynės E-sportas Punktai Blogas</u> Mobiliosios programėlės Tikrinti lažybų kortelę Statistika Krepšinio statymai priimami ir rezultatai fiksuojami įskaitant pratesimą(-us), jeigu nenurodyta kitaip. Išimtį sudaro įvykiai, kuriuose galima spėti baigtį ""0"" t.y. priimami statymai už lygiąsias. Lažinantis lažybose gyvai (<u>Live</u>) dėl FUTBOLO VARŽYBŲ STATISTIKOS (t.y. kiek kartų komandos prasižengs; kiek kartų komanda smūgiuos į vartus; geltonos kortelės; kiek bus kampinių varžybose ir t.t.) rezultatas fiksuojamas tik pagrindinio laiko metu. Statymai už geltonas korteles: Jei už ANTRĄJĄ GELTONĄ kortelę futbole parodoma RAUDONA kortelė, tai SPĖJANT GELTONŲ KORTELIŲ SKAIČIŲ ĮSKAITOMA TIK PIRMOJI parodyta GELTONA kortelė. Realaus laiko rezultatų lentelė matoma lažybų puslapyje naudotina tik kaip orientacinė. Kompanija neprisima atsakomybės dėl informacijos ar duomenų savaikiškumo. Supratau, neberodyti daugiau TOP SPORT LAŽYBOS GYVAI + TV tai daugiau negu gyvai atnaujinami rezultatai. Tai - viskas ko reikia fanui vienoje vietoje. Stebėkite rezultata, nuolat atnaujinamą statistiką, būkite įvykių sukuryje ir nepraleiskite svarbiausių detalių. <u>TopSport</u> gyvos TV transliacijos prieinamos kiekvienam, tiek jūsų kompiuteryje ar išmaniajame. Rinkitės ir sekite karščiausių rungtynių rezultatus iš viso pasaulio. Nepraleiskite varžybų bei čempionatų - čia rasite visą informaciją, transliacijas, tiek FIBA, FIFA ar LKL naujienas. Rinkitės iš pasaulio fanų pamėgtų sportų ir sekite rungtynių eigą gyvai - jūsų patogumui krepšinis <u>live</u>, futbolas <u>live</u>, tenisas, tinklinis, beisbolas, stalo tenisas, rankinis, salės futbolas, badmintonas, golfas <u>snukeris</u>, žolės riedulys, paplūdimio tinklinis, kriketas, regbis, net internetiniai sportai kaip „<u>League of Legends</u>“, „<u>Counter-Strike: GO</u>“, „<u>Hearthstone</u>“. Rinkitės rungtynių apžvalgą ir vienoje vietoje raskite viso pasaulio karščiausių žaidimų informaciją viename puslapyje. Stebėti rezultatus dar niekada nebuvo taip greitai ir paprasta. Svarbiausią informaciją apie lažybas rasite lažybų kalendoriuje, taigi niekuomet nepraleisite savo šanso. Sėkmingi azarto mėgėjai žino - fortūna lydi tik pasiruošusius ir suprantančius, kad sėkmė reikalauja žinių bei laiko. Rezultatų išklotinės bei statistika padės sukurti unikalią ir užtikrintą strategiją. Stebėkite rungtynes, sekite varžybų baigtį, domėkitės statistika ir pažinkite komandas bei žaidėjus. Vienos rungtynės gali būti lemtingos, tad nepraleiskite net mažiausių detalių visų svarbiausių sportų rezultatų suvestinėse ir statistikos lentelėse. Sukurkite patogią, personalizuotą lažybų lentelę bei patirtį sau - viską ko reikia rasite LAŽYBOS bei MANO LAŽYBOS kategorijose. Skambinkite +370 37 409409 Rašykite pagalba@topsport.lt Klausimai D.U.K Sporto lažybos <u>Lažybos gyvai Lenktynės Kazino žaidimai Lažybų žaidimai Tikrinti lažybų kortelę Įmokos/Išmokos Taisyklės Pranešimai Lažybų rezultatai Slapukų politika Privatumo politika Pateikti prašymą neleisti lošti Apie mus Karjera Lažybų punktai ir kazino Mobiliosios programėlės Blogas Licencijuojama Lietuvos Lošimų priežiūros tarnyba, licencijos Nr. 0046. Registruota S.Dariaus ir S.Girėno g. 31A, Kaunas, Lietuva, LT-46265. Tel.: +370 37 409409 Lošimų priežiūros tarnyba prie Lietuvos Respublikos finansų ministerij</u>
Id	2
Url	https://www.betsafe.lt/lt/lazybos-gyvai,Azartinis

Main_category	Azartinis
Text	Lažybos gyvai - spēk krepšinio, futbolo ir kitus rezultatus - <u>Betsafe</u>Betsafe lažybos gyvai su besikeičiančiais statymais. Tūkstančiai sporto mačų vykstančių čia ir dabar - kasdien. Žaibiškos lažybos ir tiesioginės rungtynių transliacijos. <u>betsafe</u>, lažybos, lažybos internetu, statymai, statymai gyvai, lažybos gyvai, lažybų koeficientai, lažybų pasiūla, sąlyginiai statymai, akcijos ir nuolaidos, kazino, internetinis kazino, pokeris, internetinis pokerisBūtinieji slapukai Reklamos slapukai Cash Out kodas Lažybos gyvai - spēk krepšinio, futbolo ir kitus rezultatus Statymų gyvai pasirinkimai Futbolo lažybų statymai gyvai Krepšinio lažybos gyvai >Kokiose sporto įvykiuose taip pat galima atlikti statymus gyvai? Pagalba gyvai Pagalba BETSAFE ŽAIDIMAI Apie Betsafe Atsisiųsk Betsafe programėlęŠie slapukai yra labai svarbūs, kad galėtum naudotis mūsų svetaine ir jos funkcijomis, pvz., pasiekti saugias svetainės vietas. Be šių slapukų, prašomų paslaugų nebus galima suteikti. Tam tikri slapukai užtikrina saugumą, kurį privaloma užtikrinti siekiant įvykdyti tam tikrus norminius reikalavimus ar saugumo reikalavimus, pavyzdžiui, dėl internetinių mokėjimų. Štai kodėl kai kurie našumo arba funkciniai slapukai yra įtraukti į šią kategoriją. Būtinieji slapukai Šie slapukai naudojami norint pateikti tavo interesus geriau atitinkančius skelbimus. Be to, jie yra naudojami norint apriboti reklamos parodymų skaičių ir įvertinti reklaminių kampanijų efektyvumą. Jie „prisimena“, kad apsilankei mūsų tinklalapyje, ir ši informacija gali būti pateikta kitoms organizacijoms, pvz., reklamdaviams. Šie slapukai padės mums pateikti tik susijusią reklamą pagal tavo pageidavimus. Reklamos slapukai Išsaugoti Mes naudojame slapukus, kad galėtume tau suteikti pačią geriausią patirtį ir, kad galėtume rodyti, tik tavo interesus atitinkančias reklamas vadovaujantis mūsų slapukų politika. Jų nustatymus gali pakeisti čia. Slapukų nustatymai Sutinku DRAUDŽIAMA LOŠTI NUOTOLINIUS STALO LOŠIMUS IR LOŠIMUS A KATEGORIJS AUTOMATAIS ASMENIMS, KURIEMS NĖRA SUKAKĘ 21 METAİ •

3.2.1. Dažniausiai naudojamų žodžių pašalinimas ir teksto normalizavimas

Išsaugotas puslapio tekstas turi daug žodžių, tačiau tai nereiškia, kad jie yra visi tinkami naudojant mašininio mokymosi modelį. Žodžiai, kurie neturi įtakos kategorijai – pašalinami iš esamo teksto. Tokių žodžių – apie 400 [34], kaip pavyzdys: aš, be, ir, anas, dar, kurie tame pačiame tekste gali pasikartoti po kelis kartus, tačiau tinklapio kategorijos visiškai neturėtų ir negalėtų lemti. Šio proceso esmė realizuojant prototipe, kad sudarytas tinklapio tekstas būtų prasmingas ir suteiktų kuo daugiau informacijos apie tam tikrą kategorijos puslapį.

Žodžių kanonizavimui bus naudojamas lemavimo būdas naudojant Spacy programavimo kalbos biblioteką. Kanonizavimas būtinas norint atrinkti tiksliau žodžių pasikartojimus, kitokiu būdu žodžiai turintys skirtingas galūnes ar priešdėlius, būtų identifikuojami, kaip skirtingi žodžiai. 16 lentelėje pateikiamas pavyzdys, kaip po šio etapo analizuojamo puslapio tekstas atrodo.

16 lentelė. Puslapio turinys po teksto normalizavimo ir dažniausiai naudojamų žodžių pašalinimo

Id	1
Url	https://www.topsport.lt/lazybos-gyvai-beta
Main_category	Azartinis

Cleaned_text	lažybos gyvas <u>tv transliacija</u> <u>topsport</u> lažybos punktas modernus didus pramoga erdvė populiarius lažybos internetu <u>tv platforma</u> visoje lietuvoje <u>top sport</u> lažybos gyvai <u>tv rezultatai</u> realiu laiku <u>tv tiesiogiai</u> <u>top sport</u> lažybos gyvai <u>tv kategorijos</u> <u>topsport</u> pagalba <u>topsport</u> remti puslapis reikalauti <u>javascript</u> palaikymas draudžiamas loštas nuotolinis stalas lošimas <u>lošimas</u> kategorija automatas asmuo nebūti sukakę metas draudžiamas loštas asmuo nebūti sukakę metas išpėti neatsakingas lošimas galėti taptas priklausomybė azartinis lošimas priežastis prisijungti nepavykta prisijungtas spausti čia registruotis <u>english</u> <u>русский</u> sportas gyvas <u>gyvas</u> kazino <u>kazino</u> gyvas lažybos žaidimas lenktynės sportas punktas blogas sportas gyvas <u>gyvas</u> kazino <u>kazino</u> gyvas lažybos žaidimas lenktynės sportas punktas blogas mobilus programėlės tikrinti lažybos kortelė statistika krepšinis statymas priimamas rezultatas fiksuojami įskaitant nenurodyta kitaip išimtis sudaryti įvykis galimas spėtas baigtis priimamas statymas lygus lažinantis lažybos gyvas <u>live</u> futbolo varžybų statistikos kartus komanda prasižengs kartus komanda smūgiuos vartai geltonas kortelė būti kampinis varžybose rezultatas fiksuojamas pagrindinis laikyti metas statymas geltonas kortelė antrąją geltoną kortelė futbolas parodomas raudona kortelis spėjant geltonų kortelių skaičių įskaitoma pirmoji parodytas geltona kortelis realus laikyti rezultatas lentelė matomas lažybos puslapis naudotina orientacinis kompanija neprisiima atsakomybė informacija duomuo savalaikiškumo suprasti neberodyti daug <u>top sport</u> lažybos gyvai <u>tv</u> daug gyvas atnaujinami rezultatas viskas ko reikti fanui vienoje vietoje stebėti rezultatas nuolat atnaujinamą statistikas būti įvykis sukurs nepraleiskite svarbus detalė <u>topsport</u> gyvas <u>tv transliacija</u> prieinamas tiekti kompiuteris išmanus rinktis sekti karštas rungtynės rezultatas viso pasaulis nepraleiskite varžybų čempionatas čia rasti visą informacija transliacija tiekti <u>fiba</u> <u>fifa</u> <u>lkl</u> naujiena rinktis pasaulis fanų pamėgtas sportas sekti rungtynės eiga gyvas patogumas krepšinis <u>live</u> futbolas <u>live</u> tenisas tinklinis beisbolas stalas tenisas rankinis salėti futbolas badmintonas golfas <u>snukeris</u> žolėti riedulys paplūdimys tinklinis kriketas regbis internetiniai sportai <u>league of legends</u> <u>counter strike</u> <u>go hearthstone</u> rinktis rungtynės apžvalga vienoje vietoje rasti viso pasaulis karštas žaidimas informacija vienas puslapis stebėti rezultatas niekada nebūti greitas paprastas svarbus informacija lažybos rasti lažybos kalendorius niekuomet nepraleisite šansas sėkmingas azartas mėgėjas žinti fortūna lydėti pasiruošęs suprantančius sėkmė reikalauti žinia laikyti rezultatas išsklotinė statistika padėti sukurtas unikalūs užtikrintas strategija stebėti rungtynės sekti varžybų baigtis domėtis statistika pažinti komanda žaidėja vienos rungtynės galėti būtas lemtingas nepraleiskite mažiausia detalė visų svarbus sportas rezultatas suvestinis statistika lentelė sukurti patogus personalizuotą lažybos lentelė patirtis viską ko reikti rasti lažybos <u>lažybos</u> kategorija skambinti rašyti klausimas sportas lažybos <u>lažybos</u> gyvas lenktynės kazino žaidimas lažybos žaidimas tikrinti lažybos kortelė įmoka išmoka taisyklė pranešimas lažybos rezultatas slapukas politika privatumo politika pateiktis prašymas neleistas loštas karjera lažybos punktas kazino mobilus programėlės blogas licencijuojama <u>lietuvos</u> lošimas priežiūra tarnyba licencija <u>nr</u> registruotas g 31a kautis <u>lietuva</u> tel lošimas priežiūra tarnyba <u>lietuvos</u> respublika finansai ministerija
Id	2
Url	https://www.betsafe.lt/lt/lažybos-gyvai,Azartinis
Main_category	Azartinis

Cleaned_text	lažybos gyvas spėti krepšinis futbolas rezultatas <u>betsafebetsafe</u> lažybos gyvas besikeičiančiais statymais tūkstantis sportas mačų vykstančių čia dabar kasdien žaibiškas lažybos tiesioginis rungtynės transliacija <u>betsafe</u> lažybos <u>lažybos</u> internetu statymas <u>statymas</u> gyvas lažybos gyvas lažybos koeficientas lažybos pasiūla sąlyginis statymas akcija nuolaida kazino internetinis kazino pokeris internetinis <u>pokeris</u> būtinieji slapukas reklama slapukas <u>cash out</u> kodas lažybos gyvas spėti krepšinis futbolas rezultatas statymas gyvas pasirinkimas futbolas lažybos statymas gyvas krepšinis lažybos gyvas sportas įvykis galimas atlikti statymas gyvas pagalba gyvas pagalba <u>betsafe</u> žaidimai <u>betsafe</u> atsisiųsk <u>betsafe</u> programėlės slapukas irti labas svarbus galėti naudotis svetainė funkcija <u>pvz</u> pasiektas saugias svetainė vietas slapukas prašomas paslaugus nebūti galimas suteiktas tikras slapukas užtikrinti saugumas privalomas užtikrintas siekiant įvykdytas tikras norminis reikalavimas saugumas reikalavimas pavyzdys internetinių mokėjimas našumas funkcinis slapukas irti įtrauktas kategorija būtinieji slapukas <u>slapukas</u> naudojamas norint pateiktis interesas geras atitinkantis skelbimas irti naudojamas norint apribotas reklama parodymas skaičius įvertintas reklaminis kampanija efektyvumas prisiminti apsilankei tinklalapyje informacija galėti būtas pateiktas organizacija <u>pvz</u> reklamaviams slapukas padėti pateiktis susijęs reklama pageidavimas reklama slapukas išsaugoti naudoti slapukus galėti suteiktas geras patirtis galėti rodytas interesas atitinkantis reklama vadovaujantis slapukas politika nustatymas galėti pakeistas čia slapukas nustatymas sutikti draudžiama lošti nuotolinius stalo lošimus <u>lošimus</u> kategorijos automatais asmenims nėra sukakę metai draudžiama lošti asmenims nėra sukakę metų išpėjame neatsakingas lošimas gali tapti priklausomybės azartinių lošimų priežastimi atnaujintas greitas patogus atsisiųsk <u>betsafe</u> programėlę atsisiųsk atnaujintas greitas patogus atsisiųsk <u>betsafe</u> programėlę atsisiųsk draudžiama lošti nuotolinius stalo lošimus <u>lošimus</u> kategorijos automatais asmenims nėra sukakę metai draudžiama lošti asmenims nėra sukakę metų išpėjame neatsakingas lošimas gali tapti priklausomybės azartinių lošimų priežastimi lt lietuvių lt english en <u>русский ru</u> slapukas nustatymas prisijungti <u>prisijungti</u> registracija sportas lažybos gyvas rezultatas statistika patikrinti lažybos kortelė <u>check cash out</u> artima šalis <u>europa</u> <u>vilnius</u> visos šalys kraunasi dešimtainiai amerikietiškas trupmeniniai kazino <u>live games</u> kazino gyvas pokeris <u>stars club</u> mokėjimai <u>mobile</u> naujiena atnaujintas greitas patogus atsisiųsk <u>betsafe</u> programėlę atsisiųsk atnaujintas greitas patogus atsisiųsk <u>betsafe</u> programėlę atsisiųsk sportas įvykis prasidėjus viskas prasidėti nuolat kintantys koeficientas suteikti galimybė matant vyksta aikštė įvertintas koeficientas priimtas sprendimas matant įvykis eiga <u>cash out</u> funkcija skirtas apsaugotas dalis statomas pinigas sumoti nuimtas mažesnis dalis potencialas laimėjimas didus pirminis stadija atliktame statyme statymas gyvas įvertinus komanda pajėgumas nusiteikimas pirmas minutė bendras žaidimas kokybė komanda diktuojamą tempas prisitaikymas ekipa žaidimas filosofija susikirtimą koeficientas galėti būtas peržiūrėtas statymas gyvas sekcija ekipa pajėgus įmuštas daug įvartis priešininkas žaisti uždaresnį saugesnį futbolas komanda pelnys daug įvartis ekipa įveikti varžovas įvartis pranašumas mačo eiga sekti vertas rinktis įvartis azijietiška pranašumas potencialiai stiprus ekipa mušti komanda treneris dažnas iškelti ekipa žaidėja tikra užduotis vieni treneris mėgti užklupta oponentas nepasiruošęs didelis spaudimas varžovas skirtas pirma mačo minutė žaidimas tempas užkeliamas aukštuma silpnas ekipa atsilaikytas tapti labas sunkus komanda pasiekus įvartis žaidimas tempas kristi ekipa pereiti žaistas antras numeris žaidimas grindžia kontratakėmis saugesniu kamuolys palaikymas pastebint niuansas koeficientas galėti būtas peržiūrėtas statymas gyvas sekcija mėgstamas krepšinis komanda platus lygus pasirinkimas įskaitant <u>lietuvo nacionalinis rinktinė</u> varžybas pradedant kontroliniais susitikimais baigiant svarbus didus metas renginys <u>europa</u> pasaulis čempionatais olimpinis žaidynės pereinant vietinį frontas koeficientas galėti būtas patikrintas visose rungtynės įskaitant vietinį frontas <u>betsafe</u> <u>lkl</u> krepšinis lyga žalgiris rytas <u>lietkabelis</u> utena pasvalys <u>kėdainiai</u> likęs <u>lietuvo</u> komanda nuožmiausi mūšis derbiai susitikimas lemiantys ekipa išsidėstymas turnyrinis lentelė koeficientas galėti būtas peržiūrėtas visose <u>eurolygos</u> <u>europa</u> taurėti <u>nba</u> kova žvelgiant skirtingas krepšinis lygas skirtingas
--------------	--

Lyginant lentelę 12 iš 3.2 poskyrio ir 13 lentelę galime matyti, kad tekstas normalizavosi, tam tikri žodžiai skirtinguose tekstuose tapo identiški ir gali būti suvienodinti. Nebeliko jokių papildomų simbolių, kaip pavyzdys pirmame analizuojame puslapyje „Lažybos gyvai + TV transliacijos“.

3.3. Žalingo turinio duomenų rinkinio apdorojimas prototipe

Mašininio mokymosi metodams teksto įvestis nėra tinkama, būtinas konvertavimas į skaitinę reikšmę, todėl analizėje aprašytas (1.11 poskyryje) TF-IDF panaudojamas papildomam išankstiniam duomenų rinkinio apdorojimui, kad būtų galima išskirti požymius pasirinktoms etiketėms ir būtų galima panaudoti, kaip įvesti pasirinktam klasifikavimo metodui. Papildomas

apdorojimo procesas vektorizuojant duomenų rinkinį pavaizduotas 15 pav. Etape „Apdoroti iš anksto duomenis“ naudojama Python bibliotekos scikit-learn funkcija TfidfVectorizer. Šios funkcijos parametras max_df užtikrina, kad per dažnai pasirodantys žodžiai būtų pašalinami. Pavyzdžiui:

- max_df = 0,50 reiškia, kad bus nepaisomi terminai, kurie yra daugiau nei 50% dokumentuose.
- max_df = 25 reiškia, kad nepaisyti kurie yra daugiau nei 25 dokumentuose.

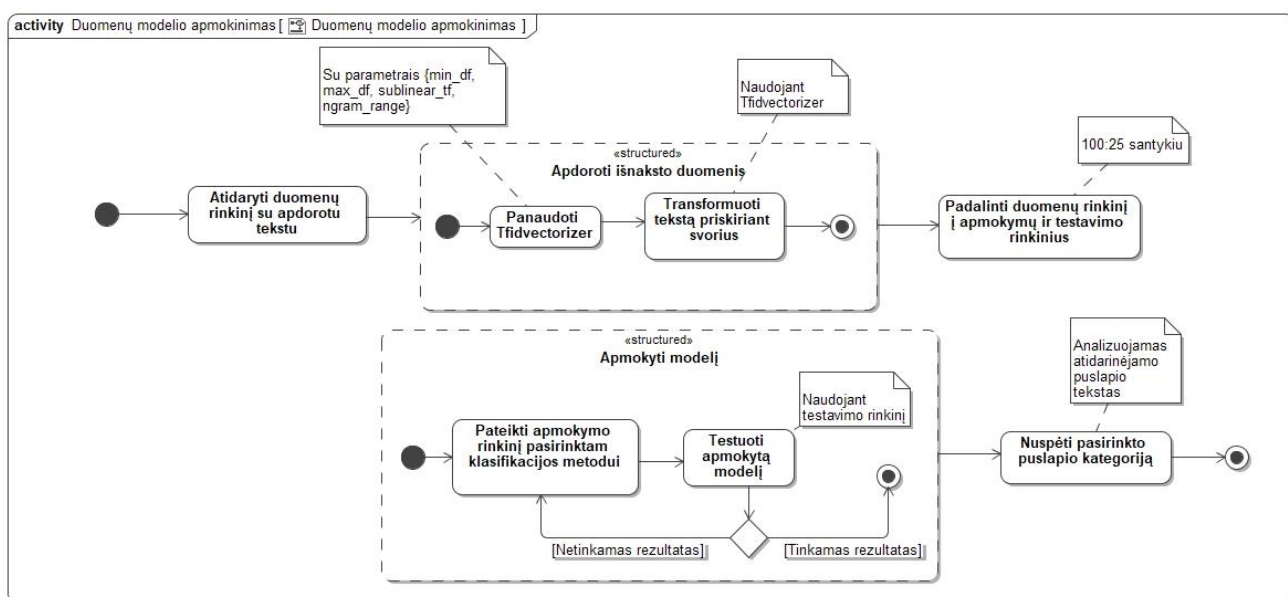
Numatytas pagal nutylėjimą max_df – 1, kuris reiškia, kad nepaisomi terminai kurie – daugiau nei 100% dokumentų, taiga numatytasis nustatymas nepaiso jokių terminų.

Sekantis parametras min_df naudojamas terminams, pašalinimui kurie pasirodo per retai. Pavyzdžiui:

- min_df = 0,01 reiškia, kad bus nepaisomi terminai, kurie yra mažiau nei 1% dokumentų.
- min_df = 5 reiškia, kad bus nepaisomi terminai, kurie – mažiau nei 5 dokumentuose.

Numatytasis nustatymas min_df – 1, tai reiškia, kad bus nepaisomi terminai, kurie yra mažiau nei 1 dokumente. Taigi numatytasis nustatymas nepaiso jokių terminų.

Atliekant tyrimą buvo panaudotas parametras min_df – 5, kuris reiškia, kad bus nepaisomi terminai, kurie – mažiau nei 5 tekstuose.



17 pav. Duomenų rinkinio apdorojimas naudojant TF-IDF

Kadangi išanalizuotas tekstas iš puslapių – didelės apimties, o TF-IDF sudarinėja matricas (tekstų skaičius, žodžiai), tai paanalizuosime anksčiau pateikto teksto ištraukas:

- Tekstas1: lažybos gyvas tv transliacija lažybos punktas modernus didus,
- Tekstas2: olimpinis krepšinis atletika statymas olimpinis žaidynės,
- Tekstas3: futbolas lažybos internetu statymas tiesiogiai gyvas.

Prieš atlikdami atvirkštinį dokumento dažnį, pateiksime termino dažnį, kuris veikia, kaip skaičiavimo vektorizavimas, tai suteiks mums reikšmę kiekvienam žodžiui, kur vektorius tekste – kvadratų suma = 1.

17 lentelė. Terminų dažnis naudojant TD

Id	atletika	didus	futbolas	gyvas	internetu	krepšinis	lažybos	modernus
0	0.000	0.316	0.000	0.316	0.000	0.000	0.632	0.316
1	0.354	0.000	0.000	0.000	0.000	0.354	0.000	0.000
2	0.000	0.000	0.408	0.408	0.408	0.000	0.408	0.000

18 lentelė. Terminų dažnis pritaikius atvirkštinį dokumento dažnį (TF-IDF)

Id	atletika	didus	futbolas	gyvas	internetu	krepšinis	lažybos	modernus
0	0.000	0.356	0.000	0.271	0.000	0.000	0.541	0.356
1	0.363	0.000	0.000	0.000	0.000	0.363	0.000	0.000
2	0.000	0.000	0.46	0.349	0.46	0.000	0.349	0.000

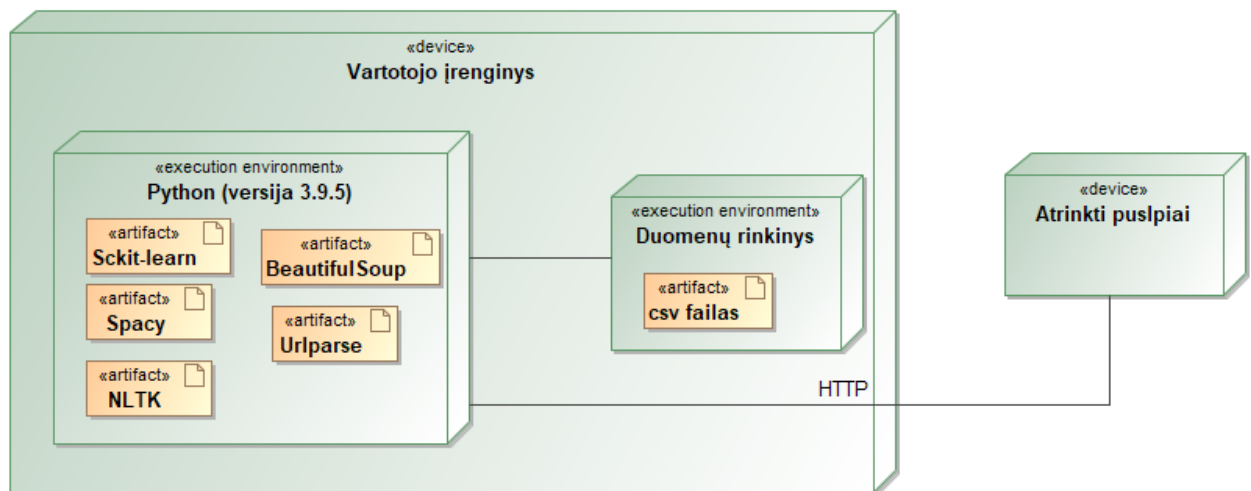
Iš 17 ir 18 lentelės naudojant tik TD ir TF-IDF metodą, matome, kad dažniausiai pasikartojantys žodžiai matricoje gauna didesnes skaitines vertes. Žodžiai kurie nesikartoja arba neturi reikšmės gauna 0 svorį.

3.3.1. Duomenų rinkinio mažinimas

Po duomenų rinkinio apdorojimo, požymių rinkinį sudaro 263146 unikalūs žodžiai bendrai susumavus visų puslapių. Didėjant puslapių kiekiui, žodžiai taip pat sparčiai plečiasi. Standartiškai atraminių vektorių algoritmas ir visi kiti aptarti analizės dalyje mašiniame mokyme naudoja visus įvesties kintamuosius, todėl gali įvykti „dimensijos prakeikimas“ kas turėtų neigiamos įtakos modelių mokymui. TF-IDF leidžia nustatyti didžiausią žodžių skaičių į kuriuos būtų atsižvelgiama. Skirtingas požymių dydis turi įtakos eksperimento rezultatams. Tyrimo metu nuspręsta naudoti 800 požymių dydį, kuris leidžia kiekvieno skirtingo puslapio išgautam tekstui jį suvienodinti ir išgauti tam tikro baigtumo tekstą puslapiui.

3.3.2. Žalingo turinio aptikimo prototipo diegimas

Projektuojamo žalingo turinio aptikimo metodo prototipo diegimo diagrama pateikta 18 pav. Iš paveikslėlio matyti, kad šio prototipo pagrindą sudaro vartotojo įrenginys palaikantis Python 3.9.5 programavimo kalbą ir šios versijos įvairios bibliotekos, duomenų rinkiniui naudojamas paprastas tekstinis failas atskirtas kableliais (angl. comma separated values). Turi būti užtikrinta interneto prieiga, kad vartotojo įrenginys atrinktus puslapius duomenų rinkinyje galėtų pasiekti.



18 pav. Žalingo turinio aptikimo prototipo diegimo diagrama.

Python programavimo kalba turi didelį paplitimą, todėl gali būti palaikomos operacinės sistemos – Windows, Linux, Mac OS X. Vartotojui užtenka turėti 2GB darbinės atminties kompiuterį. Diegime gali būti panaudotas ne tik fizinis, bet ir virtualus kompiuteris. Pagrindinis kriterijus, kad būtų išpildytas darbinės atminties reikalavimas.

3.3.3. Žalingo turinio atpažinimo prototipo realizacijos išvados

Surinktas duomenų rinkinys, panaudojant aprašytą metodologiją su nurodytais įrankiais. Duomenų rinkinyje tekstas normalizuotas ir etiketėms išskirtos savybės vektorizuojant tekstą. Realizuoti analitinėje dalyje aprašyti mašininio mokymosi algoritmai, kad būtų galima įvertinti skirtingus algoritmus, išrinkti labiausiai tinkantys duomenų rinkiniui.

3.4. Žalingo turinio atpažinimo tyrimas

Tyrimo metu bus bandoma įvertinti analitinėje dalyje aprašyti mašininio mokymosi algoritmai:

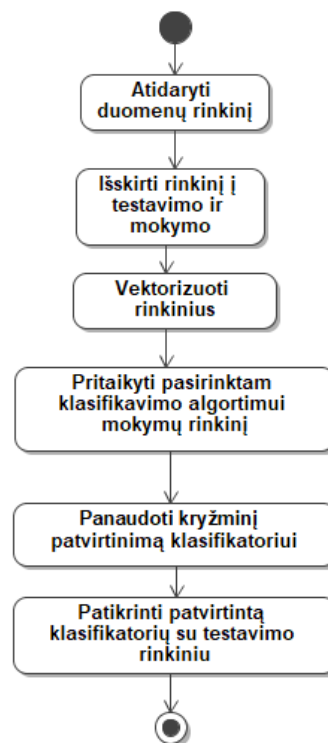
- Atsitiktinio miško,
- Atraminių vektorių,
- Multinominiu Naiviu Bajeso,
- Gausiniu Naiviu Bajeso.

Atraminių vektorių algoritmas pasirinktas dėl galimo didelio tikslumo gavimo ir mažesnės skaičiavimų galios naudojimo. Likusieji panaudoti mašininio mokymosi algoritmai parinkti palyginimui tikslumų.

Sekančiuose skyreliuose pateikiami bandymai ir algoritmų įvertinimai.

3.4.1. Žalingo turinio atpažinimo metodo klasifikacijos žingsniai ir numatytieji parametrai

Visi pagrindiniai klasifikacijos žingsniai pateikiami 19 pav. šiame poskyryje bus aprašyti parametrai, kurie – naudojami šiuose žingsniuose. Pirmas žingsnis atidaryti rinkinį, šiuo metu taip pat etiketėms – prilyginami skaičiai, kaip pvz. Pornografinis – 0, azartinis - 1, normalus – 0.



19 pav. Žalingo turinio aptikimo metodo klasifikacijos apmokymo žingsniai.

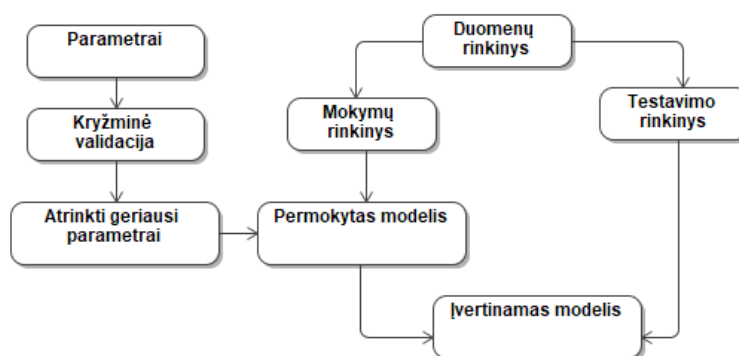
Duomenų rinkinys išskiriamas į atsitiktinius testavimo ir mokymo rinkinius. Testavimo rinkinį sudaro 20% pradinio rinkinio, o likusieji 80% puslapių naudojama, kaip mokymų rinkinys. Taikant duomenų rinkinio padalijimą pritaikytas maišymas, kuris rinkinio būseną visada vienodai išmaišo, tolygiai.

Sudarinėjant testavimo ir mokymo rinkinių vektorius naudojamas analizės dalyje aprašytas TF-IDF, kurio parametrai pateikti 19 lentelėje.

19 lentelė. TF-IDF parametrai

Kodavimas	UTF-8
N-gram diapazonas	1,2
Maksimalios savybės	800
Normalizavimas	L2
Sutiesinimas	tiesa

Maksimalios savybės reiškia, kad kiekvienas puslapio tekstas turės maksimalų savybių skaičių. Kai taikoma l2 normalizavimas, vektorinių elementų kvadratų suma – 1. Dviejų vektorių kosinuso panašumas – jų taškinė sandauga. Sutiesinimas (angl. sublinear) terminų dažnio mastelį.



20 pav. Klasifikatoriaus kryžminė validacija [37].

Pritaikius kryžminę validaciją pagal 20 pav. yra išrenkami geriausių rezultatą davę klasifikatoriaus parametrai. Kadangi naudojama ši validacija, pasirinktiems klasifikatoriams nėra pateikiami jokie pradiniai parametrai apart atsitiktinio miško klasifikatoriaus. Šiam minėtam klasifikatoriui parenkamas medžių skaičius miške – 100, didžiausias medžio gylis – 5 ir atsitiktinė būsena – 0.

3.4.2. Sudarytos n-gramos ir b-gramos

Apmokinant modelį – sudarinėjama kiekvienai kategorijai n-gramos, b-gramos, po vieną žodį ir žodžių junginį sudarytą iš dviejų žodžių. 20 lentelėje pateikiami atrinkti žodžiai pagal kategoriją, kuriuos išrinko klasifikatorius.

20 lentelė. Daugiausiai pasikartojančios n-gramos ir b-gramos

	Azartinis	Pornografinis	Normalus
n-gramos	azartinis, žaidimas, lošimas, žaidėja, lažybos	nuogas, mergintis, sex, seksas, moteris	receptas, maistas, kelionė, patiekalas, šeima
b-gramos	azartinis žaidimas, azartinis lošimas, foto filmas, lošimas automata	nuogas mergintis, oralinis seksas, safe sex	maistas produktas, sandėlis krepšelis, sirgti ligos,

Lentelėje pateikti atraminio vektoriaus klasifikatoriaus atrinkti žodžiai. Pagal pasirinktą klasifikatorių atrinkti žodžiai ir jų junginiai skiriasi nežymiai.

3.4.3. Algoritmų įvertinimai

Mašininio mokymosi modelių veikimą įvertinti galima pagal tai, kaip yra klasifikuojamos internetinių puslapių kategorijos. Vertinimo balai – rodiklis, kuris leidžia sužinoti mašininio mokymosi galimybę, nuspėti kategorijas pagal pirminį mokymo požymių rinkinį. Mašininio mokymosi modeliai gali nuspėti svetainių kategorijas, kai mokymo požymių rinkinys pritaikomas pagal modelį. Pagal apmokymo duomenų rinkinį į modelį įvedami požymių duomenys, kurie pateikia prognozuojamų internetinių svetainių kategorijų rinkinį. Šios prognozuotos kategorijos palyginamos su mokymo etikečių rinkiniu. Prognozių rinkinys ir mokymo etikečių rinkinys suteikia galimybę sudaryti vertinimo balus ir analizuoti, kaip modelis apmokytas nuspėti svetainių kategorijas. Skaičiuojant modelių prognozių balus formulėse – 4 terminai:

- Tikrai teigiamas (angl. true positive) – rezultatas, kai modelis teisingai prognozuoja teigiamą klasę. Pavyzdys: Sergančiam asmeniui, buvo diagnozuota liga.

- Tikrai neigiamas (angl. true negative) – rezultatas, kai modelis teisingai prognozuoja neigiamą klasę. Pavyzdys: Nesergančiam asmeniui, nebuvo diagnozuota liga.
- Klaidingai teigiamas (angl. false positive) – rezultatas, kuris parodo, kad sąlyga egzistuoja, nors ji neegzistuoja. Pavyzdys: Sveikam žmogui buvo diagnozuota tikra liga.
- Klaidingai neigiamas (angl. false negative) – testo rezultatas, kad sąlyga nepasitvirtino, o iš tikrųjų ji galioja. Pavyzdys: Sergančiam asmeniui, nustatyta, kad ligos nėra.

$$\text{Tikslumas} = \frac{TP+TN}{TP+TN+KT+KN} \quad (6)$$

$$\text{Atšaukimas} = \frac{TP+TN}{TP+KT} \quad (7)$$

$$f1 = \frac{2 \cdot \text{preciziškumas} + \text{atšaukimas}}{\text{preciziškumas} + \text{atšaukimas}} \quad (8)$$

$$\text{Preciziškumas} = \frac{TP}{TP+KN} \quad (9)$$

Klasifikavimo tikslumas – teisingų prognozių skaičius, kuris išreikštas visų prognozių santykiu. Tikslumas apskaičiuojamas pagal formulę (3). Atšaukimas – teisingų teigiamų rezultatų skaičius, padalytas iš teisingai teigiamų ir klaidingai neigiamų rezultatų skaičiaus, apskaičiuojamas formulėje (4). Preciziškumo vertinimo metodas nustato kiek tikslus arba netikslus – mašininio mokymosi modelis pagal kiek teigiamų prognozių buvo prognozuota iš visų prognozių, apskaičiuojamas pagal (6) formulę. F1 rezultatas – tikslumo ir atšaukimo vidurkis, vertinant abi metrikas (5) formulę.

Sumaišymo matricos naudojimas leidžia pateikti mašininio mokymosi modelio prognozavimo rezultatus. Naudojant daugiau nei dvi etiketes, tikslumas dažnu atveju gali būti klaidinantis.

21 lentelė. Sumaišymo matricos paaiškinimas

		Faktinės vertės	
		Teigiamas	Neigiamas
Prognozuojamos vertės	Teigiamas	Tikrai teigiamas (TP)	Klaidingai neigiamas (FN)
	Neigiamas	Klaidingai teigiamas (FP)	Tikrai neigiamas (TN)

Sumaišymo matricos išvesties pavyzdys dvejetaini klasių klasifikacijai. TP – atspėta reikšmė – teigiama ir modelis nuspėjo, kad tai yra teigiama. TN – nuspėta reikšmė neigiama ir ji yra teisinga. KN – numatytas teigiamas rezultatas, o tai yra klaidinga. KT – nuspėta reikšmė neigiama ir tai yra klaidinga.

3.4.4. Išanalizuotų algoritmų palyginimas su surinktu duomenų rinkiniu

Mašininio mokymosi algoritmai apmokyti, panaudojant darbo metu sudaryti duomenų rinkinį, įvertintas jų vidutinis tikslumas (angl. accuracy) ir standartinis nuokrypis. Palyginimas pateiktas 22 lentelėje.

22 lentelė. Ištirtų algoritmų kryžminės validacijos vidutiniai įvertinimai

Algoritmo pavadinimas	Vidutinis tikslumas, %	Standartinis nuokrypis
-----------------------	------------------------	------------------------

Algoritmo pavadinimas	Vidutinis tikslumas, %	Standartinis nuokrypis
Gausinis Naivus Bajesas	86,57	0,09
Multinominis Naivus Bajesas	87,90	0,08
Atsitiktinis miškas	86,28	0,11
Atraminis vektorius (SVM)	87,61	0,10

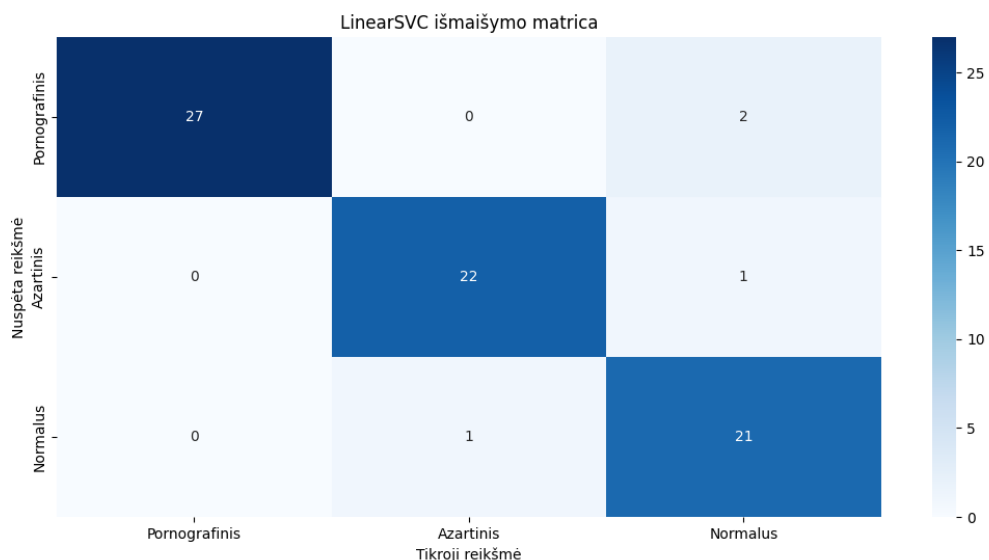
Didžiausią vidutinį tikslumą per kryžminę validaciją turi Multinominis Naivus Bajesas. Didžiausią standartinį nuokrypį ir mažiausią tikslumą turi atsitiktinis miškas, kurio tikslumas 86,28%, lyginant su atraminio vektoriaus mažesnis 1,3%.

Sekančioje 23 lentelėje pateikiami atraminio vektoriaus analizės matavimų įvertinimai ir vidurkiai. Preciziškumas skiriasi 4,29~7,14% skirtingų reikšmių. Tarp reikšmių skirtumo neturi likę įvertinimai. Bendras atraminio vektoriaus tikslumas – 93,69%.

23 lentelė. Atraminio vektoriaus klasifikatoriaus įvertinimas

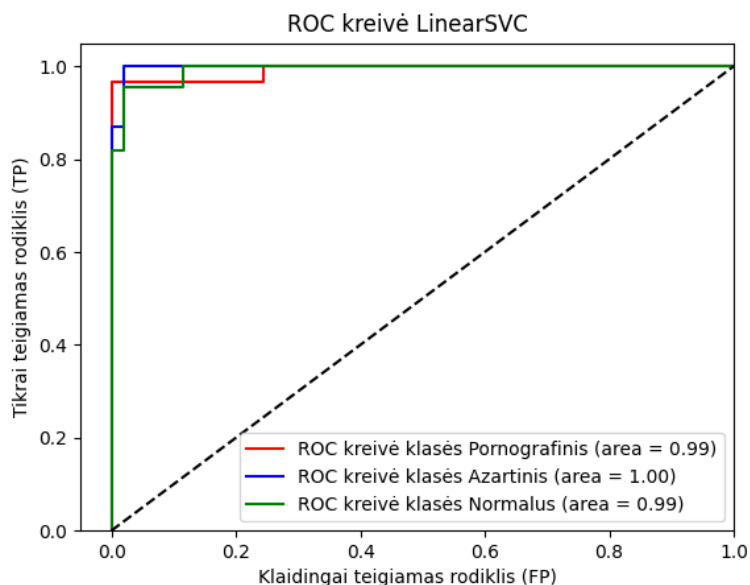
Algoritmo pavadinimas	Preciziškumas, %	Atšaukimo koeficientas, %	F1 reikšmė, %
Azartinis	94,29	97,06	95,65
Pornografinis	97,14	89,47	95,65
Normalus	90,24	94,87	92,50
Vidurkis	93,89	93,80	93,77
Svertinis vidurkis	93,84	93,69	93,69

21 pav. pateikta algoritmo išmaišymo matrica, kurioje matosi, kiek kurios reikšmės nuspėtos algoritmo teisingai. Pagal tikrai teigiamas (TP) reikšmes normalaus turinio buvo nuspėta mažiausiai ir preciziškumas tarp visų kategorijų – mažiausias.



21 pav. Atraminio vektoriaus algoritmo išmaišymo matrica.

Atraminio vektoriaus ROC, AUC įvertinimas – 0,99, kreivė pateikiama 22 pav. Pateikiama tikimybės kreivė, kurioje nubraižoma TPR ir FPR esant įvairioms slenkstinėms vertėms atskiria signalą nuo triukšmo. Plotas po kreive (AUC) – klasifikatoriaus gebėjimas atskirti klases, kuris naudojamas, kaip santrauka.



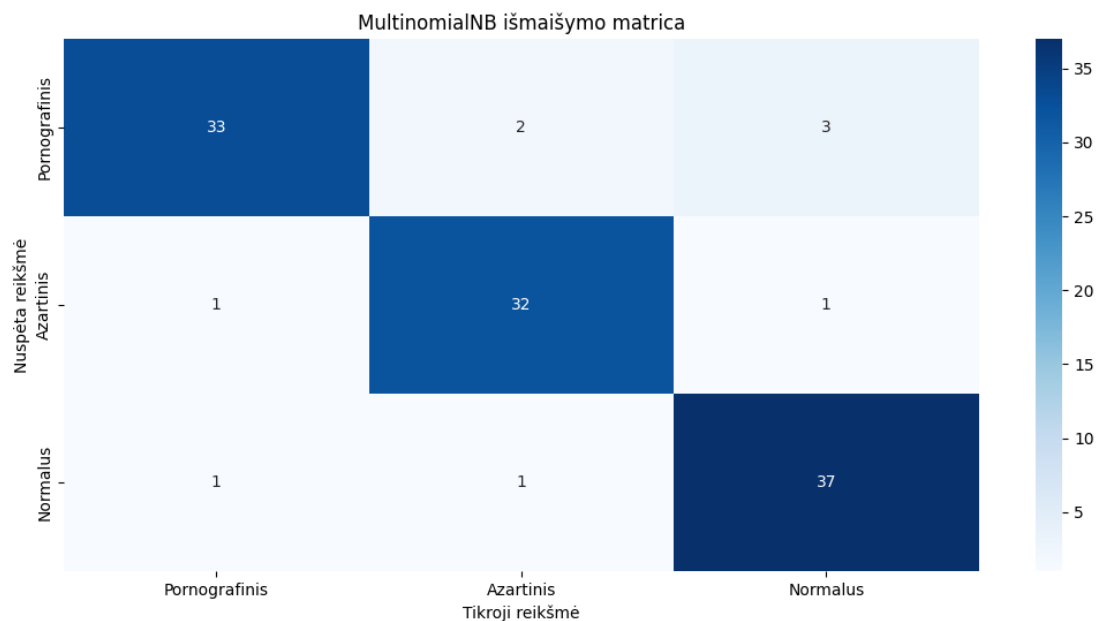
22 pav. Atraminio vektoriaus ROC kreivė.

Toliau pateikiamas multinominio Naiviojo Bajeso klasifikatoriaus įvertinimas 24 lentelėje. Matoma, kad tik pornografinio turinio kategorijai preciziškumas buvo 94%, tačiau atšaukimo ir f1 reikmės buvo normalaus turinio mažesnės nei pornografinio. Bendras klasifikatoriaus tikslumas 91,89%.

24 lentelė. Multinominio Naivaus Bajeso algoritmo metodo įvertinimas

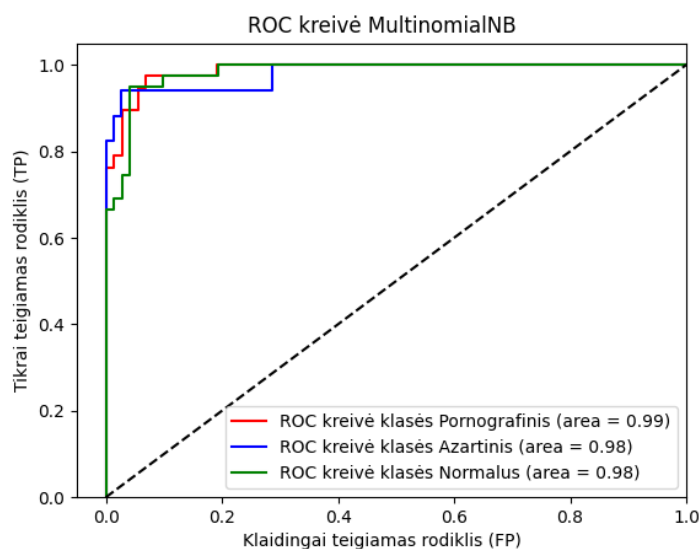
Algoritmo pavadinimas	Preciziškumas, %	Atšaukimo koeficientas, %	F1 reikšmė, %
Azartinis	91,43	94,12	92,75
Pornografinis	94,29	86,84	90,41
Normalus	90,24	94,87	92,50
Vidurkis	91,99	91,94	91,89
Svertinis vidurkis	91,99	91,89	91,86

23 pav. pateikta multinominio Naiviojo Bajeso išmaišymo matrica, kurioje matosi, kad normalus turinys buvo nuspėtas teisingai daugiau kartų nei likusios kategorijos.



23 pav. Multinominio Naivaus Bajeso išmaišymo matrica.

Multinominio klasifikatoriaus ROC, AUC įvertinimas 0,98, kreivė pateikiama 24 pav.



24 pav. Multinominio Naivaus Bajeso išmaišymo matrica.

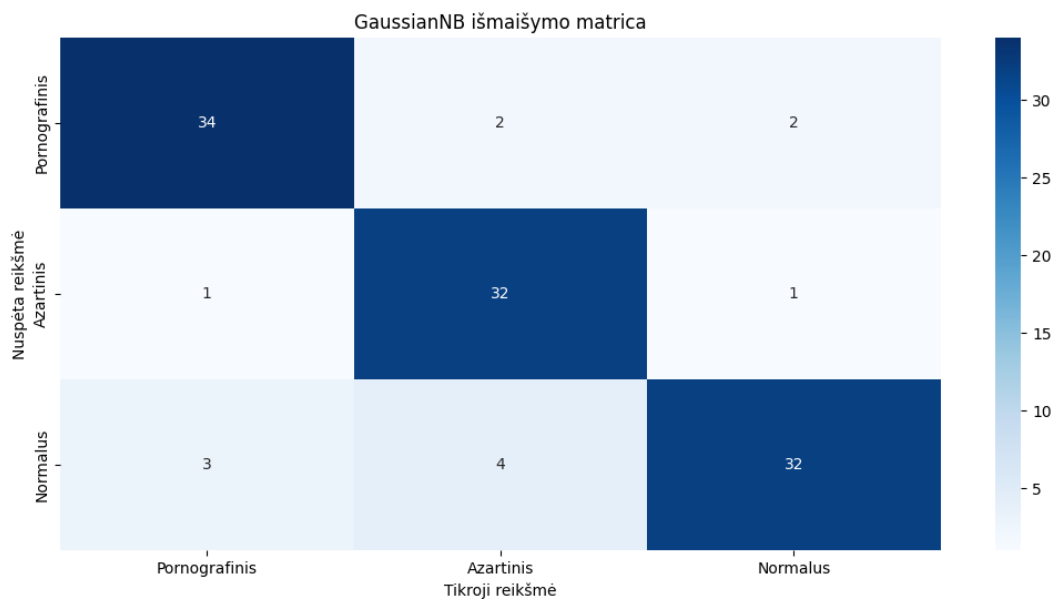
Toliau pateikiamas Gausinio Naivaus Bajeso klasifikatoriaus įvertinimas 25 lentelėje. Gausinis parodė prastesnius rezultatus nei multinominis, tačiau normalaus turinio kategorijai turi didžiausią preciziškumą lyginant su gausiniu. Bendras šio klasifikatoriaus tikslumas 88,29%.

25 lentelė. Gausinio Naivaus Bajeso metodo įvertinimas

Algoritmo pavadinimas	Preciziškumas, %	Atšaukimo koeficientas, %	F1 reikšmė, %
Azartinis	84,21	94,12	88,89

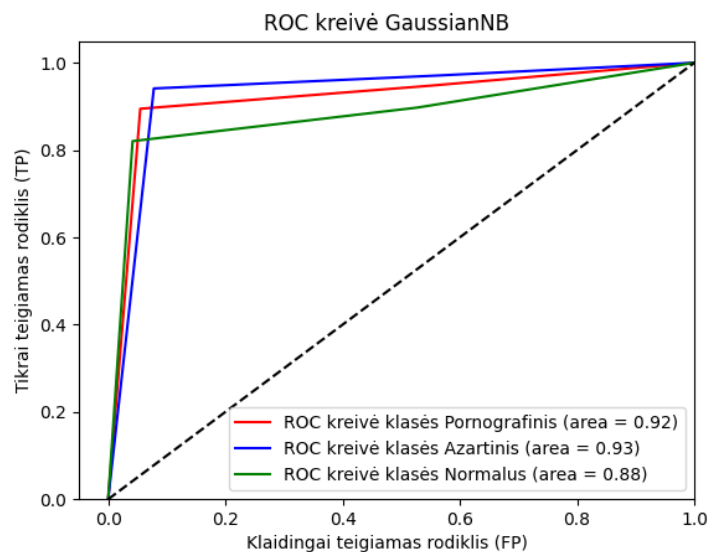
Algoritmo pavadinimas	Preciziškumas, %	Atšaukimo koeficientas, %	F1 reikšmė, %
Pornografinis	89,47	94,12	88,89
Normalus	91,43	82,05	86,49
Vidurkis	88,37	89,55	88,28
Svertinis vidurkis	89,55	88,29	88,25

25 pav. pavaizduota algoritmo išmaišymo matrica, kurioje matosi kiek, kurios reikšmės nuspėtos teisingai. Pagal tikrai teigiamas reikšmes normalus ir azartinis turinys buvo nuspėtas teisingai tiek pat kartų.



25 pav. Gausinio Naivaus Bajeso išmaišymo matrica.

Gausinio klasifikatoriaus ROC, AUC įvertinimas – 0,91, kreivė pateikiama 26 pav.



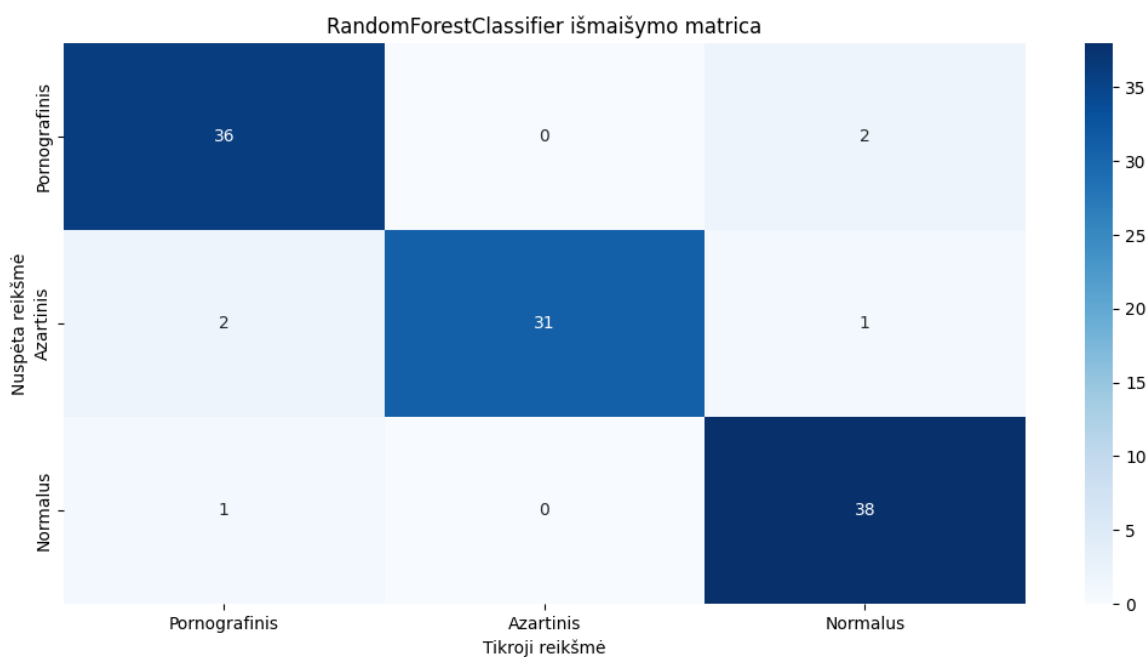
26 pav. Gausinio Naivaus Bajeso ROC kreivė.

Toliau pateikiamas 26 lentelėje atsitiktinio miško algoritmo matavimų įvertinimai, vidurkiai. Bendras šio klasifikatoriaus tikslumas – 89,26%. Lyginant su kitais testuojamais mašininio mokymosi algoritmais šio tikslumas – mažiausias.

26 lentelė. Atsitiktinio miško metodo įvertinimas

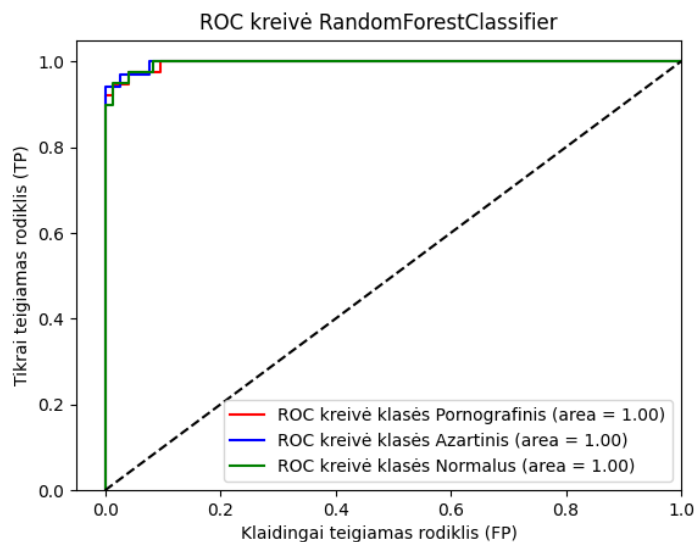
Algoritmo pavadinimas	Preciziškumas, %	Atšaukimo koeficientas, %	F1 reikšmė, %
Azartinis	100	91,18	95,38
Pornografinis	87,80	94,74	93,51
Normalus	92,23	92,31	92,31
Vidurkis	93,31	92,74	92,94
Svertinis vidurkis	93,12	92,79	92,85

27 pav. pateikta algoritmo išmaišymo funkcija, kurioje matosi kiek kokios reikšmės atspėtos teisingai.



27 pav. Atsitiktinio medžio išmaišymo matrica.

Atsitiktinio medžio klasifikatoriaus ROC, AUC įvertinimas – 0,99, kreivė pateikta 25 pav.



28 pav. Atsitiktinio medžio ROC kreivė.

3.5. Tyrimo išvados

Tyrimo metu ištirti 4 mašininio mokymo algoritmai ir jie išdetalizuoti. Naudoti šie: atraminio vektoriaus, atsitiktinio miško, multinominio ir gausinio Naiviojo Bajeso klasifikatorių algoritmai buvo apmokinti naudojant sudarytą duomenų rinkinį, kad būtų galima įvertinti ir išrinkti geriausią mašininio mokymosi algoritmą pasirinktos problemos sprendimui.

Pagal tikslumą iš visų keturių mašininio mokymosi algoritmų tiksliausiai atspėja azartinio ir pornografinio turinio tinklapius. Atraminų vektorių klasifikatorius pagal bendrą tikslumą (93,69) pasiekė geriausią rezultatą, tačiau normalus turinys buvo aptinkamas tik 90,24 preciziškumu. Gausinis Naivus Bajesas aptinkant normalų turinį buvo geriausias, tačiau sekančios duomenų rinkinio kategorijos buvo aptinkamas prasčiau.

Naudojantis eksperimento rezultatais galime daryti išvadą, kad atraminio vektoriaus algoritmas tinka žalingo turinio aptikimui ir leidžia pasiekti geresnį rezultatą.

3.6. Įžvalgos ateičiai

Atliekant apsaugos nuo žalingo turinio internete metodo tyrimą, kilo mintys, kurias galima būtų išbandyti, tačiau dėl ribotų projekto resursų to nepavyko įgyvendinti. Atliekant panašų darbą ateityje arba bandant patobulinti metodą, būtų galima išbandyti:

- Papildyti metodą nuotraukų klasifikacijomis puslapiuose naudojant Spa-BOVW ar kitus metodus.
- Pabandyti sujungti skirtingus vektorius gautus naudojant kelis vektorizavimo metodus.

Išvados ir rezultatai

- 1) Atliktoje analizėje išanalizuota kas yra žalingas turinys remiantis LR įstatymais ir kaip jis kenkia vaikams ir ne tik. Išanalizuoti rinkoje esami sprendimai, kurių dauguma paremti juodaisiais sąrašais arba nėra apibrėžta kas būtent – naudojama sistemoje, todėl nuspręsta toliau analizuoti metodus naudojančius teksto klasifikavimo algoritmus, naudojančius mašininį mokymą, kadangi juodieji sąrašai neužtikrina sėkmingo žalingo turinio identifikavimo.
- 2) Sukurtas metodas turinio atpažinimui naudojant teksto klasifikavimą. Sudaryta duomenų rinkinio sudarymo metodologija. Duomenų rinkinys paruoštas mašininio mokymosi algoritmui pagal sudarytą veiksmų seką – teksto normalizavimas, vektorizavimas. Apmokytas atraminių vektorių algoritmas pagal sudarytą duomenų rinkinį.
- 3) Įgyvendintas žalingo turinio aptikimo metodas realizuojant prototipą. Pagal duomenų rinkinio atrinkimo metodą atrinkti puslapiai suskirstyti pagal kategorijas ir jų turinys iškrapštytas, ir paruoštas mašininio mokymosi algoritmui. Realizuoti mašininio mokymosi algoritmai: atraminių vektorių, atsitiktinio miško, multinominio ir gausinio naivaus Bajeso klasifikatoriai. Realizuoti mašininio mokymosi algoritmai, kad eksperimentinėje dalyje būtų galima palyginti klasifikatorių efektyvumą su surinktu duomenų rinkiniu.
- 4) Atliktas tyrimas ir realizuoti mašininio mokymo algoritmai.
 - a) Gausinis Naiviojo Bajeso tikslumas – 88,29%.
 - b) Multinominio Naiviojo Bajeso tikslumas – 91,89%.
 - c) Atsitiktinio miško tikslumas – 89,28%.
 - d) Atraminis vektorius – 93,69%.
- 5) Su sudarytu duomenų rinkiniu atraminių vektorių algoritmo įvertinimas tarp lyginamų – geriausias, kuris yra 93,69%, nedideliu skirtumu 1,8% prastesnį tikslumą turėjo multinominis Naiviojo Bajeso algoritmas, tačiau atliekant testavimus Bajeso algoritmas ne visų kategorijų puslapius nustatydavo tiksliai. Iš visų palygintų mašininio mokymosi algoritmų žalingo turinio aptikimui geriausiai tiktų atraminių vektorių algoritmas.

Literatūros sąrašas

1. Commission of the European Communities, Illegal and harmful content on the Internet [Tinkle]. Prieiga per internetą: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:1996:0487:FIN:en:PDF> [kreiptasi 2021-01-10].
2. DE STREEL, A., Online Platforms' Moderation of Illegal Content Online [Tinkle]. Prieiga per internetą: <http://www.crid.be/pdf/public/8590.pdf> [kreiptasi 2021-01-10].
3. Lietuvos Respublikos visuomenės informavimo įstatymas [Tinkle]. Prieiga per internetą: <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.280580> [kreiptasi 2021-01-10].
4. Summary Report on the open public consultation on the Digital Services Act Package [Tinkle]. Prieiga per internetą: <https://ec.europa.eu/digital-single-market/en/news/summary-report-open-public-consultation-digital-services-act-package> [kreiptasi 2021-01-10].
5. Children's online activities, risks and safety: A literature review by the UKCCIS Evidence Group [Tinkle]. Prieiga per internetą: <https://www.gov.uk/government/publications/childrens-online-activities-risks-and-safety-a-literature-review-by-the-ukccis-evidence-group> [kreiptasi 2020-12-01].
6. Cyber-Smart Children, Cyber Sage Teenagers: Enhancing internet Safety For Children [Tinkle]. Prieiga per internetą: <https://ieeexplore.ieee.org/document/8102292> [kreiptasi 2020-12-01].
7. Keipi, T., Näsi, M., Oksanen, A., & Räsänen, P. (2016). *Online hate and harmful content: Cross-national perspectives* (p. 154). Taylor & Francis [Tinkle]. Prieiga per internetą: https://library.oapen.org/viewer/web/viewer.html?file=/bitstream/handle/20.500.12657/22350/9781138645066_text.pdf?sequence=1&isAllowed=y [kreiptasi 2020-12-01].
8. Adekolu, A. M., & Chen, L. (2019). A Simple Approach to Develop a Parental Control via a Monitoring Home Network [Tinkle]. Prieiga per internetą: <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1292238&dswid=-9973> [kreiptasi 2020-12-01].
9. Ismail, W. H. W. (2020). Parental Control System for Children on Wireless Network. *Journal of Computing Technologies and Creative Content (JTec)*, 5(1), 14-20 [Tinkle]. Prieiga per internetą: <http://www.jtec.org.my/index.php/JTeC/article/view/236> [kreiptasi 2020-12-01].
10. Suliman, N. A. B., & Mammi, H. B. K. (2017, May). Explicit words filtering mechanism on web browser for kids. In *2017 6th ICT International Student Project Conference (ICT-ISPC)* (pp. 1-6). IEEE [Tinkle]. Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/8075322> [kreiptasi 2020-12-01].
11. Lodder, A. R., & Polter, P. (2017). ISP blocking and filtering: on the shallow justification in case law regarding effectiveness of measures. *European Journal of Law and Technology*, 8(2) [Tinkle]. Prieiga per internetą: <http://ejlt.org/index.php/ejlt/article/view/517> [kreiptasi 2020-12-01].
12. The end of Pirate Bay? Virgin media and Sky users given warning to stop using torrent site [Tinkle]. Prieiga per internetą: <https://www.mirror.co.uk/tech/end-pirate-bay-virgin-media-9765951>
13. Valstybinė vartotojų teisių apsaugos tarnyba (2017) Interneto svetainių pažeidžiančių vartotojų apsaugos teisės aktų reikalavimus, priežiūros ir prieinamumo vartotojams ribojimo galimybės [Tinkle]. Prieiga per internetą:

- <http://www.vvtat.lt/data/public/uploads/2019/02/2017-m.-tyrimas-del-interneto-svetainiu-prieziuros-ir-prieinamumo-vartotojams-ribojimo-galimybiu.pdf#page=36&zoom=100,109,76> [kreiptasi 2020-12-01].
14. Ryšių reguliavimo tarnybos aprobuoto turinio filtravimo priemonių rekomendacija [Tinkle]. Prieiga per internetą: <https://www.esaugumas.lt/lt/turinio-filtravimo-programos/316> [kreiptasi 2020-12-01].
 15. Neil J. Rubenking, Ben Moore (2020) The Best Parental Control Software [Tinkle]. Prieiga per internetą: <https://www.pcmag.com/picks/the-best-parental-control-software> [kreiptasi 2020-12-01].
 16. „Net Nanny“ [Tinkle]. Prieiga per internetą: <https://www.netnanny.com/> [kreiptasi 2020-12-01].
 17. Neil J. Rubenking, Ben Moore, „Net Nanny“ review (2020) [Tinkle]. Prieiga per internetą: <https://sea.pcmag.com/security/19380/net-nanny> [kreiptasi 2020-12-01].
 18. Neil J. Rubenking, Ben Moore (2020), „Qustodio“ review [Tinkle]. Prieiga internetą: <https://www.pcmag.com/reviews/qustodio> [kreiptasi 2020-12-04].
 19. Neil J. Rubenking (2020), SafeDNS review [Tinkle]. Prieiga per internetą: <https://www.pcmag.com/reviews/safedns> [kreiptasi 2020-12-01].
 20. Rusmansyah, R., Mizuardy, H., & Yusuf, B. (2018). DNS FILTERING: A CLEAN AND POSITIVE INTERNET ENVIRONMENT IN UIN AR-RANIRY BANDA ACEH. *Cyberspace: Jurnal Pendidikan Teknologi Informasi*, 2(1), 8-17 [Tinkle]. Prieiga per internetą: <https://jurnal.ar-raniry.ac.id/index.php/cyberspace/article/view/2502> [kreiptasi 2020-12-01].
 21. Labrador, J. G., Michael, P., & Dinawanao, D. D. DNSfilter: An Erlang/OTP Implementation of a DNS-Based Web Filtering System [Tinkle]. Prieiga per internetą: https://www.researchgate.net/profile/Dante_Dinawanao/publication/322958697_DNSFilter_An_Erlang_OTP_Implementation_of_a_DNS-Based_Web_Filtering_System/links/5a79a52245851541ce5ceddb/DNSFilter-An-Erlang-OTP-Implementation-of-a-DNS-Based-Web-Filtering-System.pdf [kreiptasi 2020-12-01].
 22. Altinel, B., & Ganiz, M. C. (2018). Semantic text classification: A survey of past and recent advances. *Information Processing & Management*, 54(6), 1129-1153 [Tinkle]. Prieiga per internetą: <https://www.sciencedirect.com/science/article/pii/S0306457317305757> [kreiptasi 2020-12-01].
 23. Ieva Bumbulienė (2018), Automatinis lietuviškų pastoviųjų žodžių junginių atpažinimas [Tinkle]. Prieiga per internetą: https://www.vdu.lt/cris/bitstream/20.500.12259/36425/1/ieva_bumbuliene_md.pdf [kreiptasi 2020-12-01].
 24. Erikas Kuvakin (2017), Pattern recognition algorithms for data mining [Tinkle]. Kreiptasi 2020-12-01].
 25. Berniukevičius, A., & Kurilov, J. (2017). Dirbtiniai neuroniniai tinklai personalizuotam mokymuisi. *Lietuvos matematikos rinkinys. Ser. B*, 58, 22-27 [Tinkle]. Prieiga per internetą: <https://epublications.vu.lt/object/elaba:25739677/>
 26. Ptaszynski, M., Masui, F., Nitta, T., Hatakeyama, S., Kimura, Y., Rzepka, R., & Araki, K. (2016). Sustainable cyberbullying detection with category-maximized relevance of harmful phrases and double-filtered automatic optimization. *International Journal of Child-Computer Interaction*, 8, 15-30 [Tinkle]. Prieiga per internetą: <https://www.sciencedirect.com/science/article/pii/S221286891630054X> [kreiptasi 2021-01-10].
 27. Ali, F., Khan, P., Riaz, K., Kwak, D., Abuhmed, T., Park, D., & Kwak, K. S. (2017). A fuzzy ontology and SVM-based Web content classification system. *IEEE Access*, 5, 25781-25797 [Tinkle]. Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/8094233> [kreiptasi 2021-01-10].
 28. Weinberg, Z., Sharif, M., Szurdi, J., & Christin, N. (2017). Topics of controversy: An empirical analysis of web censorship lists. *Proceedings on Privacy Enhancing Technologies*, 2017(1), 42-61 [Tinkle]. Prieiga per internetą: <https://content.sciendo.com/view/journals/popets/2017/1/article-p42.xml> [kreiptasi 2021-01-10].

29. Varadharajan, V. (2010). Internet filtering-issues and challenges. *IEEE Security & Privacy*, 8(4), 62-65 [Tinkle]. Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/5523867/> [kreiptasi 2021-01-10].
30. Wendland, A., Zenere, M., & Niemann, J. (2021, September). Introduction to Text Classification: Impact of Stemming and Comparing TF-IDF and Count Vectorization as Feature Extraction Technique. In *European Conference on Software Process Improvement* (pp. 289-300). Springer, Cham [Tinkle]. Prieiga per internetą: https://link.springer.com/chapter/10.1007/978-3-030-85521-5_19 [kreiptasi 2021-01-10].
31. Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. *Information*, 10(4), 150 [Tinkle]. Prieiga per internetą: <https://www.mdpi.com/2078-2489/10/4/150/htm#sec4-information-10-00150> [kreiptasi 2021-01-10].
32. Jindal, R., Malhotra, R., & Jain, A. (2015). Techniques for text classification: Literature review and current trends. *webology*, 12(2) [Tinkle]. Prieiga per internetą: <http://www.webology.org/2015/v12n2/a139.pdf> [kreiptasi 2021-01-10].
33. Kadhim, A. I. (2019). Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*, 52(1), 273-292 [Tinkle]. Prieiga per internetą: <https://link.springer.com/article/10.1007/s10462-018-09677-1> [kreiptasi 2021-01-10].
34. List of Stopwords in Lithuanian Language [Tinkle]. <https://stopwords.net/lithuanian-lt/> [kreiptasi 2021-01-10].
35. Chen, Yang, et al. "Automatic detection of pornographic and gambling websites based on visual and textual content using a decision mechanism." *Sensors* 20.14 (2020) [Tinkle]. <https://www.mdpi.com/1424-8220/20/14/3989/htm> [kreiptasi [2021-01-10].
36. Amalia, Amalia, Dani Gunawan, and Maya Silvi Lydia. "The identification of negative content in websites by using machine learning approaches." *2019 5th International Conference on Computing Engineering and Design (ICCED)*. IEEE, 2019 [Tinkle]. Prieiga per internetą: <https://ieeexplore.ieee.org/abstract/document/9161105> [kreiptasi 2021-01-10].
37. Cross-validation: evaluating estimator performance [Tinkle]. Prieiga per internetą: https://scikit-learn.org/stable/modules/cross_validation.html [kreiptasi 2021-01-10].