КІБЕРБЕЗПЕКА ТА ЗАХИСТ КРИТИЧНОЇ ІНФОРМАЦІЙНОЇ ІНФРАСТРУКТУРИ / CYBERSECURITY & CRITICAL INFORMATION INFRASTRUCTURE PROTECTION

DOI: 10.18372/2225-5036.28.17368

УДОСКОНАЛЕННЯ МЕТОДУ ВИЯВЛЕННЯ НЕПРАВДИВОЇ ІНФОРМАЦІЇ ЗА ДОПОМОГОЮ БАЙЕСОВСКОГО КЛАСИФІКАТОРА

Наталія Лукова-Чуйко, Тетяна Лаптєва

Київський національний університет імені Тараса Шевченка , Україна

ЛУКОВА-ЧУЙКО Наталія Вікторівна , д. т. н., завідувачка кафедри кібербезпеки та захисту інформації
Дата та місце народження: 1982, Сєвєро-Курільськ, Сахалінська область.
Освіта: Київський національний університет імені Тараса Шевченка, Україна, 2006.
Посада: завідувачка кафедри кібербезпеки та захисту інформації факультету інформаційних технологій, Київського національного університету імені Тараса Шевченка, Україна ,
Наукові інтереси: кібербезпека та захист інформації.
Публікації: більше 100 наукових публікацій, включаючи монографії, підручники, статті та патенти.
E-mail: lukova@knu.ua.
ORCID ID: 0000-0003-3224-4061.

ЛАПТЄВА Тетяна Олександрівна, аспірантка, кафедри кібербезпеки та захисту інформації, факультету інформаційних технологій
Дата та місце народження: 1986, Київ, Київська область, Україна.
Освіта: Національний авіаційний університет, Україна, 2010.
Посада: аспірантка кафедри кібербезпеки та захисту інформації
Наукові інтереси: захист інформації, виявлення неправдивої інформації, інформаційне протиборство.
Публікації: більше 10 наукових публікацій, з них 5 публікацій у наукометрічної базі Scopus
E-mail: tetiana1986@ukr.net.
ORCID ID: 0000-0002-5223-9078.

Анотація. *У статті наведено аналіз рішень прикладних задач, що вирішуюся шляхом застосування методів кластерного аналізу. Це і класичні задачі Data Mining: кластеризація, класифікація, і характерні тільки для текстових документів завдання: автоматичне анотування, витяг ключових понять та ін. Проведено скорочений аналіз методів кластерізації інформації. Підтверджується постулат про те, що кластеризація передбачає розбиття множини елементів на кластери, кількість яких визначається локалізацією елементів заданої множини в околицях деяких природних центрів кластерів. Проведено аналіз застосування байесовського класифікатора. доведено, що байесовського класифікатор при наявності апріорних ймовірностей працює з високою точністю при визначенні неправдивої інформації. Однак відповідь про отримання цій ймовірності цей метод не дає. За рахунок використання наївного байесовського класифікатора для виявлення неправдивої інформації, удосконалено метод виявлення неправдивої інформації. Цей метод дозволяє вирішити проблему невизначеності апріорної ймовірності. Запропонований наївний байесовський класифікатор для обробки текстів виявився досить ефективним. Зроблена оцінка ефективності алгоритму запропонованого методу розробки класифікатора для визначення правдивості інформації. На основі первинних даних, які отримали з мережи Інтернет, були розраховані чисельні значення оцінки*

алгоритму удосконаленого методу визначення неправдивої інформації. Отримані таки значення метрик: Recall = 0,853; Precision = 0,869; F-measure = 0,861; Accuracy = 0,855. Отримані результати доводять, що удосконалений метод (без додаткового навчання) одразу має гарні результати. Це доводить адекватність розробленого методу, та дає результативний науковий метод для виявлення неправдивої інформації. Особливо актуальність удосконаленого методу виявлення неправдивої інформації існує у теперішний час, у умовах інформаційної війни.

Ключові слова: *кластерізація, класифікація, теорема Байеса, неправдива інформація, загроза, протиборство.*

Постановка проблеми

Зростаючий об'єм інформації, призвело до поняття автоматизації отримання та сортування інформації – Data Mining. Це напрямок пов'язаний з широким спектром завдань – від розпізнавання розмитих образів до створення пошукових машин[1]. Важливою складовою Data Mining є обробка інформації у інформаційному просторі. Такі завдання описуються на поняття класифікації та кластеризації. Класифікація полягає у визначенні приналежності деякого елемента одному з заздалегідь створених класів. Кластеризація передбачає розбиття множини елементів на кластери, кількість яких визначається локалізацією елементів заданої множини в околицях деяких природних центрів кластерів. Реалізація задачі класифікації спочатку повинна описуватися на задані постулати, основні з яких – апріорна інформація про первинній множині об'єктів і міра близькості елементів та класів [2, 3].

Іншими словами, з'являється можливість розподілити накопичені дані відповідних закладів по певних групах, за деякими схожими критеріями. На разі існує велика кількість алгоритмів кластеризації. Умовно методи кластеризації розбиваються на два класи – ієрархічні та неієрархічні [4]. У неієрархічних алгоритмах присутня наявність умови зупинки та кількості кластерів. Основою цих алгоритмів є гіпотеза про порівняно невелику кількість прихованих факторів, які визначають структуру зв'язку між ознаками. Ієрархічні алгоритми не зав'язані на кількості кластерів. Ця характеристика визначається за динамікою злиття та поділу кластерів під час побудови дерева вкладених кластерів (дендрограми). У свою чергу, ієрархічні алгоритми діляться на агломеративні, які будуються шляхом об'єднання елементів, тобто зменшенням кількості кластерів, та дивізимні, засновані на поділі (розщепленні) існуючих груп (кластерів).

Неієрархічні алгоритми набули великої популярності, зважаючи на того, що в їх основі лежить та чи інша задача оптимізації, тобто групування вихідної множини об'єктів у кластери є рішенням деякого екстремального завдання. Розглянемо кілька найбільш популярних методів.

Одним із найпопулярніших методів чисельного аналізу є метод найменших квадратів. Метод k-середніх, метод Fuzzy k-means алгоритм є узагальненням попереднього, у випадку, коли кластери є нечіткими множинами, і, елемент може наложити різним кластерам з різної ступені надійності. Метод Кластеризація Гюстафсона-Кесселя та інші [5, 6].

Усі ці методи відрізняються, щонайменше, набором вхідних параметрів і в залежності від обраних параметрів результати кластеризації можуть відрізнятись. Виникає складність в оцінці результату кластеризації. Один із способів оцінити результат кластеризації –розташувати об'єкти та центри кластерів на координатній площині, але тут виникає інша проблема – у випадку, коли дані багато розмірні відпадає можливість розташування їх на координатній площині. В такому випадку виникає наукове завдання по розробці методики оцінки результатів кластеризації та класифікації.

Аналіз останніх досліджень і публікацій

На даний час в літературі описано багато прикладних задач, що вирішуюся шляхом застосування методів кластерного аналізу. Це і класичні задачі Data Mining: кластеризація, класифікація, і характерні тільки для текстових документів завдання: автоматичне анотування, витяг ключових понять та ін.

Так у роботі [1] загалом описано завдання класифікації, та традиційні підходи, які використовуються для класифікації текстових документів. Описано, як працює найпростіший, але водночас один із найчастіше використовуваних при обробці натуральних мов алгоритм класифікації – наївний байєсівський класифікатор. Але сам опис вирішення завдання класифікації для текстових документів не наводиться , питання побудови самого класифікатора не розглядаються.

У роботі [2] автори розглядають один з методів вирішення задачі обробки даних за допомогою наївного та дереводоповненого байєсівський класифікаторів. Дається визначення класифікаторів, а саме наївний та дереводоповнений байєсівські класифікатори – це ймовірнісні графічні моделі, що використовуються для моделювання великих масивів даних, які мають багато невизначеностей серед своїх взаємозалежних наборів характеристик.

Дані моделі широко використовуються для сегментації зображень, медичної діагностики і інших задач кластеризації та класифікації даних. Нажаль завдання створення моделей та методів побудови класифікаторів для обробки новинних даних на предмет правдивості, на базі розроблених методів не розглядають.

У роботах [3-4] визначають сутність класифікаторів, але обґрунтовувано власне бачення того, за якими ознаками та яким чином повинна проводитися класифікація даних. Так, наприклад, у доцільно обґрунтовано доцільність створення та проведення класифікації взаємодіючих підприємств на основі пізнання закономірностей інтеграційних процесів. Наводяться ознаки об'єктів для класифікації, але процес визначення критерія відбору об'єктів для прийняття рішення для подальшої класифікації не наводиться.

У роботі [5] наводиться порівняльні характеристики класифікаторів. Розглядаються дереводоповнений та наївний байесовський класифікатори. Наводяться відміни класифікаторів.

Більш детально описується дереводоповнений класифікатор. Дереводоповнений класифікатор – це покращений наївний класифікатор Байеса, що враховує ще один рівень взаємодії між параметрами системи, тобто кожна змінна може залежати від однієї іншої змінної, при цьому залежність між характеристиками даної моделі є більш реалістичною, аніж у наївному класифікаторі.

Дереводоповнений класифікатор має ненабагато більшу обчислювальну складність, ніж наївний байесовський класифікатор, при цьому показує кращу точність. Нажаль, самі алгоритми побудови класифікаторів не розглядаються. Методи побудови класифікаторів не наводиться.

У роботах [6-7] говориться про те, що автоматична класифікація, розпізнавання образів без вчителя, займає одне з центральних місць серед методів аналізу даних і є сукупністю підходів, методів і алгоритмів, призначених для знаходження деякого розбиття досліджуваної сукупності об'єктів на підмножини схожих між собою об'єктів.

При цьому вихідним припущенням для виділення таких підмножин, що отримали спеціальну назву кластерів, служить лише неформальне припущення про те, що об'єкти, які відносяться до одного кластера, повинні мати більшу схожість між собою, чим з об'єктами з інших кластерів. Конкретних методів та алгоритмів побудови класифікаторів та прийняття рішень на базі отриманих значень параметрів об'єктів класифікатора не наводиться

Разом з тим, незважаючи на значну кількість публікацій щодо вирішення різноманітних аспектів оцінювання правдивості інформації на сьогоднішній день залишається невирішеною задача оцінювання правдивості інформації з урахуванням можливого інформаційного впливу на суспільство, ресурси.

Тому проблема розробки методу створення класифікатора отриманої інформації та ухвалення рішення на основі розробленого класифікатора про визначення правдивої або неправдивої інформації є актуальним науковим завданням.

Мета роботи

1. Провести аналіз застосування байесовського класифікатора.
2. Удосконалити метод байесовського класифікатора для виявлення неправдивої інформації
3. Провести оцінку ефективності алгоритму запропонованого методу розробки класифікатора для визначення правдивості інформації

Виклад основного матеріалу дослідження

З метою скорочення обсягу інформації, для видалення неінформативних словосполучень або речень розглянемо метод TF-IDF.

Алгоритм TF-IDF (term frequency/inverse document frequency). TF-IDF – статистичний захід, який використовується у нашому випадку для оцінки важливості словосполучення або речення в контексті документа, що є частиною інформації у інформаційному просторі. Вага деякого словосполучення або речення пропорційна кількості використання цього словосполучення або речення в документі і обернено пропорційна частоті використання словосполучення або речення в інших документах інформаційного простору.

Міра TF-IDF часто використовується в задачах аналізу текстової інформації та інформаційного пошуку, наприклад, як один із критеріїв релевантності документа пошуковому запиту, при розрахунку міри близькості документів під час кластеризації.

TF (term frequency – частота слова) – відношення числа входження деякого слова до загальної кількості слів документа. Таким чином, оцінюється важливість слова ti в межах окремого документа

$$TF = \frac{n_i}{\sum_k n_i}, \quad (1)$$

де n_i – число входжень словосполучення або речення в документ; $\sum_k n_i$ – загальне число словосполучень або речень в документі.

IDF (inverse document frequency – зворотна частота документа) – інверсія частоти, з якою деяке слово зустрічається в документах колекції.

Облік IDF зменшує вагу широковживаних словосполучень.

$$IDF = \log \frac{|D|}{|(d_i \supset t_i)|}, \quad (2)$$

де | D | – Кількість документів у корпусі; $|(d_i \supset t_i)|$ – кількість документів, у яких трапляється t_i (коли $n_i \neq 0$).

Таким чином, міра TF-IDF є добутком двох співмножників: TF*IDF. Велику вагу в TF-IDF отримають словосполучення або речення з високою

частотою в межах конкретного документа та з низькою частотою вживання в інших документах.

Напрямок дослідження проблеми класифікації

За головне будемо вважати, що звичайний випадок ознак знаходиться у евклідовому просторі, а якісність вирішального правила вимірювати частотою появи правильних рішень. Як правило, його оцінюють, наділяючи безліч об'єктів Ω деякою вірогідністю мірою. Байесовський підхід [14] виходить із статистичної природи спостережень. За основу береться припущення про існуванні ймовірнісної міри на просторі образів, яка або відома, або може бути оцінена. Мета полягає у розробці такого класифікатора, який буде правильно визначати найімовірніший клас для пробного образу. Тоді завдання полягає у визначенні "найвірогіднішого" класу. Байесовський підхід ґрунтується на припущенні про існування деякого розподілу ймовірностей для кожного параметра. Недоліком цього методу є необхідність припущення існування апріорного розподілу для невідомого параметра, так і знання його кількісних показників. Байесовський класифікатор ґрунтується на тому, що відомі апріорні ймовірності гіпотез P(c_i), тобто, ймовірності випадання класів c_i (i = 1,2, ..., k). Відповіді на запитання: як їх знайти, класифікатор Байеса не дає [8,10-12]. Наївний байесівський класифікатор дозволяє вирішити цю проблему. Найбільш галузь застосування наївного байесовського класифікатора, це завдання класифікації текстової інформації, де цей метод е ефективним [9,13-15].

Для розгляду наївного байесовського методу робиться передумова, що алгоритм класифікації працює на множині інформаційних даних D={b_i}. Вся множина інформації розбивається на не перетинаючи підмножини класів:

$$C = \{c_i\}, \bigcup b_i = D, c_i \bigcap c_j = \varnothing \ (i \neq j). \quad (3)$$

Завданням класифікації є визначення класу, до якого відноситься цей документ. Кожному елементу b ставиться у відповідність набір ознак b = {wi}. Набір документів, що визначає клас, надалі ми будемо називати навчальною вибіркою.

Далі застосовується алгоритм класифікації виділення документів найбільш відповідних заданому класу.

Для того, щоб застосувати теорему Байеса. Яка описується виразом (4):

$$P(y \mid x) = \frac{P(x \mid y)P(y)}{P(x)}, \quad (4)$$

де P(x) – апріорна ймовірність спостереження x; P(y) – апріорна ймовірність появи події y.

Для класифікації документів, треба зробити припущення:

$$P(c_j) = \frac{n(c_j)}{\sum_j n(c_j)}, \quad (5)$$

де $n(c_j)$ – кількість термів у класі c_j.

Словосполучення незалежні одне від іншого:

$$P(w_i|c_j) = \frac{n(w_i, c_j)}{n(c_j)}, \quad (6)$$

де { w_i} – набір термів в документі b; n (w_i,c_i) – кількість термів w_i у класі у класі c_j.

Для визначення відповідної категорії для заданого документа, потрібно отримати відповідну множину словоформ. За безліччю словоформ будується структура з неповторних словосполучень та їх лічильників – (w_i, n_i).

Визначення відповідної категорії починається з кореня дерева множин статистики. Через M позначимо кількість множин статистики в даному вузлові дерева. Категорії, на належність до яких ми перевіряємо документ, позначимо через c_j(j =0,...,M – 1). Для кожного слова w_i в кожній множині статистики знаходимо це словосполучення та відповідний лічильник n(w_i ,C_j), j (j=0,1,..,M-1) – номер категорії (множини статистики)). n (c_i) позначимо кількість документів у j-й категорії. Тоді отримаємо:

$$N_j(w_i|c_j) = \frac{n(w_i, c_j)}{n(c_j)}. \quad (7)$$

Вираз (7) це нормований лічильник словосполучень w_i у j-й категорії.

Тоді ймовірність відповідності унікального словосполучення w_i j-й категорії буде визначатися виразом:

$$P(c_j|w_i) = \frac{N_j(w_i)\, n(w_i, c_j)}{S(w_i)}, \quad (8)$$

де добуток береться за всіма словами досліджуваної множини словоформ.

Апріорна ймовірність зустрічі категорії c_i буде мати вигляд:

$$P(c_j) = \frac{n(c_j)}{\sum_{j=0}^{M-1} n(c_j)}. \quad (9)$$

Слід відмітити, що якщо документ має дуже велику кількість слів які зустрічаються у категорії c_j(j =0,...,M – 1), то значення $P(c_j|w_i)$ може вийти за обмеження визначення змінної. Тому значення $P(c_j|w_i)$ треба контролювати і якщо воно вийде за обмеження значення змінної, то обмежувати малим числом, або брати рівному 0.

У випадку коли не відкидати слова які не несуть інформації, тоді пропонується використовувати вираз:

$$P(c_j|w_i) = \frac{N_j(w_i)\, n(w_i, c_j)}{S(w_i)} \log \frac{M + 1/2}{M(w_i) + 1/2}, \quad (10)$$

де M(w_i) – кількість категорій у яких зустрічаються слова.

Після першого кроку визначаємо k-штук категорій з найбільшим значенням $P(c_j|w_i)$. Зберігаємо

їх назву та значення $P(c_j|w_i)$. Відповідно з назвою кожної з цих множин, заходимо у відповідний пункт і проводимо обробку. Якщо інформація відсутня, то цей пункт пропускається. Після того, як будуть оброблені всі дочірні (за визначеними на першому кроці) категорії, з них і збережених на попередньому кроці, обирається k-штук категорій з найбільшим значенням $P(c_j|w_i)$. Зберігаємо їх звання та значення $P(c_j|w_i)$, після чого, переходимо до наступного кроку, але тільки за тими категоріями, які не були збережені на попередньому етапі. Процес продовжуємо доти, доки не буде категорій, за якими можна здійснювати перевірку. Результатом виконання програми буде k-штук категорій, найбільш можливих досліджуваного документа, з погляду критерію Байеса.

Проведемо оцінку ефективності роботи алгоритму запропонованого методу розробки класифікатора. Оцінка ефективності здійснювалася за стандартними метриками Recall (повнота), Precision (точність), F-measure (міра є гармонійним середнім між точністю і повнотою), Accuracy (доля вірних відповідей).

Основою перевірки є тестова вибірка, в якій проставлено відповідність між документами та їх класами. Залежно від ваших конкретних умов, отримання подібної вибірки може бути утруднене, оскільки часто її складають люди. Але іноді її можна отримати без великого обсягу ручної роботи, якщо виявити винахідливість. Якихось конкретних рецептів, на жаль, не існує.

При наявності тестової вибірки, достатньо нацькувати класифікатор на документи і співвідносити його рішення з відомим правильним рішенням. Але для того, щоб приймати рішення гірше або краще справляється з роботою, нова версія алгоритму нам необхідна чисельна метрика його якості.

Такою метрикою може бути частка документів, за якими класифікатор прийняв вірне рішення.

$$Accuracy = \frac{P}{N}, \qquad (11)$$

де P – кількість документів якими класифікатор прийняв вірне рішення; N – розмір навчальної вибірки.

Тим не менш, у цієї метрики є одна особливість, яку необхідно враховувати. Вона надає всім документам однакову вагу, що може бути не коректно у разі якщо розподіл документів у навчальній вибірці сильно зміщений у бік якогось одного чи кількох класів. У цьому випадку класифікатор має більше інформації щодо цих класів і відповідно в рамках цих класів він прийматиме більш адекватні рішення. На практиці це призводить до того, що ви маєте accuracy, скажімо, 80%, але при цьому в рамках якогось конкретного класу класифікатор працює дуже погано не визначаючи правильно навіть третину документів.

Точність (precision) та повнота (recall) є метриками, які використовуються при оцінці більшої частини алгоритмів вилучення інформації. Іноді вони використовуються самі по собі, іноді як базис для похідних метрик, таких як F-міра або R-Precision.

Точність системи в межах класу – це частка документів, що дійсно належать даному класу щодо всіх документів, які система віднесла до цього класу. Повнота системи – це частка знайдених класифікатором документів, що належать класу, щодо всіх документів цього класу в тестовій вибірці.

Ці значення легко розрахувати виходячи з таблиці контингентності, що складається кожного класу окремо.

Таблиця 1

Категорія		Експертна оцінка	
		позитивна	негативна
Оцінка системи	позитивна	TP	FP
	негативна	FN	TN

де TP- істино-позитивне рішення; TN- істино-негативне рішення; FP- хибно -позитивне рішення; FN- хибно -негативне рішення.

Тоді, точність та повнота визначаються таким чином:

$$Precision = \frac{TP}{TP + FP}; \qquad (12)$$

$$Recall = \frac{TP}{TP + FN}. \qquad (13)$$

На практиці значення точності та повноти набагато зручніше розраховувати з використанням матриці confusion matrix (неточностей).

Чим вище точність та повнота, тим краще. Але в реальному житті максимальна точність і повнота не можна досягти одночасно і доводиться шукати якийсь баланс. Тому, треба мати метрику, яка б поєднувала в собі інформацію про точність і повноту алгоритму запропонованого методу розробки класифікатора. Саме такою метрикою є F-мера

F-мера є гармонійним середнім між точністю і повнотою. Вона прагне нуля, якщо точність чи повнота прагне нуля.

$$F = 2 \times \frac{Precision \times Recall}{Precision + Recall}. \qquad (14)$$

Ця формула надає однакову вагу точності і повноті, тому F-мера падатиме однаково при зменшенні і точності і повноти. Можна розрахувати F-міру надавши різну вагу точності і повноті, якщо ви усвідомлено віддаєте пріоритет однієї з цих метрик при розробці алгоритму.

$$F = (\beta^2 + 1) \times \frac{Precision \times Recall}{\beta^2 \times Precision + Recall}, \qquad (15)$$

де β набуває значення в діапазоні 0<β<1 якщо ви хочете віддати пріоритет точності, а при β>1 пріоритет надається повноті. При β=1 формула зводиться до попередньої і ви отримуєте збалансовану F-міру (також її називають F1).

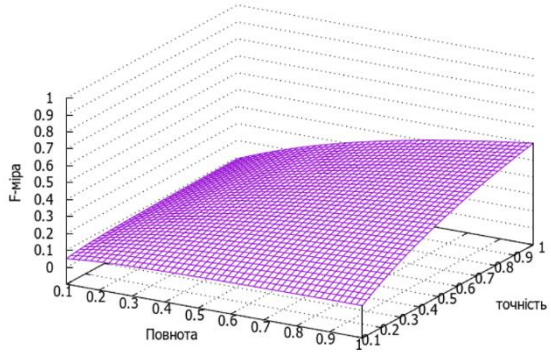

Рис.1.Збалансована міра

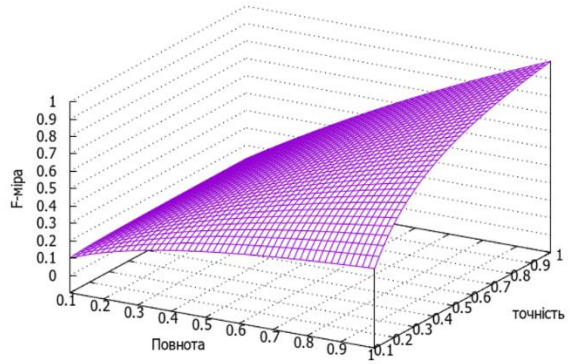

Рис.2. Пріоритет точності $\beta^2=0,25$

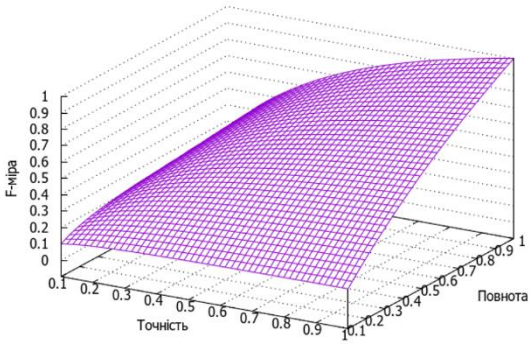

Рис.3. Пріоритет повноти $\beta^2=2$

Проводячи аналіз наведених графіків, бачимо, що функція F-міра падатиме однаково при зменшенні і точності і повноти. Що є вірним результатом. Результати підтверджують, що F-міра зводить до одного дві основні метрики: точність і повноту. Тобто F-міра вважається головною функцію оптимізації. При обранні пріоритету (точність або повнота) можливо отримати необхідний алгоритм забезпечення точності класифікатора.

На основі первинних даних, які отримали з мережі Інтернет, були розраховані чисельні значення оцінки алгоритму удосконаленого методу визначення неправдивої інформації. Отримані таки значення метрик: Recall = 0,853; Precision = 0,869; F-measure = 0,861; Accuracy = 0,855. Отримані результати доводять, що удосконалений метод (без додаткового навчання) одразу має гарні результати. Це додатково доводить адекватність розробленого методу.

Висновки. Проведено аналіз застосування байесовського класифікатора. Підтвердили, що байесовського класифікатор при наявності апріорних ймовірностей працює з високою точністю при визначенні неправдивої інформації. Однак відповідь про отримання цій ймовірності цей метод не дає.

Запропоновано удосконалення методу застосування байесовського класифікатора за рахунок використання наївного байесовского класифікатора для виявлення неправдивої інформації. Цей метод дозволяє вирішити цю проблему. Запропонований наївний байесовський класифікатор для обробки текстів виявився досить ефективним.

Проведена оцінку ефективності роботи алгоритму запропонованого методу розробки класифікатора. На основі зведеного критерія. F-міра є зведеним критерієм, або формальною метрикою оцінки якості алгоритму запропонованого методу розробки класифікатора. F-міра зводить до одного дві інші основні метрики: точність і повноту. F-міра вважається головною функцію оптимізації. В залежності від критерія, який обирається (повнота або точність) визначається F-міра та пропонується алгоритм. Можливо розраховувати F-міру надавши різну вагу точності і повноті, якщо ви усвідомлено віддаєте пріоритет однієї з цих метрик при розробці алгоритму запропонованого методу розробки класифікатора.

Проведене моделювання оцінки якості алгоритму запропонованого методу розробки класифікатора. Результати моделювання наведені у графічному вигляді. Аналіз наведених графіків, довів, що функція F-міра падатиме однаково при зменшенні і точності і повноти. Надавши різну вагу точності і повноті можливо видавати пріоритет однієї з цих метрик при розробці алгоритму запропонованого методу розробки класифікатора. Отримані результати доводять адекватність оцінки якості алгоритму запропонованого методу розробки класифікатора.

ЛІТЕРАТУРА

[1] Лаптєва Т.О. Алгоритм визначення міри існування недостовірної інформації в умовах

інформаційного протиборства. Кібербезпека: освіта, наука, техніка. No. 2 (14), 2021, С. 15-25.

[2] Наконечний В. С., Барабаш О. В., Лаптєва Т. О., Міщенко А. В. Удосконалення методу виявлення та кластеризації джерел неправдивої інформації. Наукоємні технології. Інформаційні технології, кібербезпека. Том 54 № 4 (2022) С. 105-111.

[3] Лаптєва Т.О., Лукова-Чуйко Н.В., Собчук А.В. Дослідження основних загроз і оцінка безпеки інформаційних систем. Математика. Інформаційні технології. Освіта. 2022 рік: збірка тез допов. учасник. XI Міжнар. наук.–практ. конф., 3–5 червня 2022 р. Луцьк–Світязь: СНУ імені Лесі Українки, 2022. С. 101-103.

[4] Поліщук Ю. Я., Гнатюк С. О., Сейлона П. А. Мас медіа як канал маніпулятивного впливу на суспільство. Інформаційна безпека. 2015. Т. 21, Ч. 3. С. 301-308.

[5] V. Theocharis, W. Lowe, J. W. van Deth, G. Garcla-Albacete. Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. Information, Communication & Society. 2015. 18. pp. 202-220.

[6] Butko, T., Prokhorchenko, A., Muzykin, M. An improved method of determining the schemes of locomotive circulation with regard to the technological peculiarities of railcar traffic. Eastern-European Journal of Enterprise Technologies. 2016. 5(3 (83)), pp. 47–55.

[7] Молодецька К. В. Підхід до виявлення організаційних ознак інформаційних операцій у соціальних інтернет-сервісах. Пріоритетні напрямки розвитку телекомунікаційних систем та мереж спеціального призначення. Застосування підрозділів, комплексів, засобів зв'язку та автоматизації в АТО : збірн. матер. IX наук.-практ. конф., 25 листоп. 2016 р. Київ: ВІТІ. 2016. С. 130-131.

[8] Лаптєв О.А., Бабенко Р.В., Правдивий А.М., Зозуля С.А., Стефурак О.Р. Удосконалена методика вибору послідовності пріоритетів обслуговування потоків інформації. Науково-практичний журнал «Зв'язок». К. : ДУТ, 2020. №4 (146), С. 27 – 31.

[9] O. Svynchuk, O. Barabash, J. Nikodem, R. Kochan, O. Laptiev. Image compression using fractal functions.Fractal and Fractional, 2021, 5(2), 31. pp. 1-14.

[10] Zamrii I., Sobchuk V., Laptiev O., Savchenko V., Shkapa V., Kovalenko V. and Kotok V. Fractal Functions and Their Application to Source Data Coding. ARPN Journal of Engineering and Applied Sciences. Vol. 17, No. 4, 2022. pp. 424-435.

[11] Serhii Yevseiev, Volodymir Ponomarenko, Oleksandr Laptiev, Oleksandr Milov and others/ Synergy of building cybersecurity systems. Kharkiv. Publisher PC TECHNOLOGY CENTER. 2021 – 188 c.

[12] Наталія Лукова-Чуйко, Тетяна Лаптєва. Виділення та відбір ознак для визначення неправдивої інформації. V Міжнародна науково-практична конференція. "Проблеми кібербезпеки інформаційно-телекомунікаційних систем" (PCSITS)"27-28 жовтня 2022 р. Київ, Україна. Збірник матеріалів доповідей та тез. С. 13-15.

[13] Lenkov, S., Kubyavka, M., Kubiavka, L., Lenkov, Y., Shevchuk, V. Reflex Intellectual text processing systems: Natural language text addressing (2019) CEUR Workshor Proceedings, ISSN: 16130073. - 2386, pp. 85-95.

[14] S. Lenkov, G. Zhyrov, D. Zaytsev, I. Tolok, E. Lenkov, T. Bondarenko, Y. Gunchenko, V. Zagrebnyuk, O. Antonenko/ Features of modeling failures of recoverable complex technical objects with a hierarchical constructive structure. Східно-Європейський журнал передових технологій. ISSN 1729-3774. – Харків, 2017. - №4/4(88). – С. 34-42.

[15] Oleg V. Barabash, Andrii P. Musienko, Valentyn V. Sobchuk, Nataliia V. Lukova-Chuiko, Olga V. Svynchuk. Distribution of Values of Cantor Type Fractal Functions with Specified Restrictions. Chapter in Book "Contemporary Approaches and Methods in Fundamental Mathematics and Mechanics". Editors Victor A. Sadovnichiy, Michael Z. Zgurovsky. Publisher Name: Springer, Cham, Switzerland AG 2021. pp. 433-455.

УДК 336.71:004.056

Lukova-Chuiko N., Laptieva T. Improving the false information detection method using the bayesian classifier

Abstract. *The article provides an analysis of solutions to applied problems, which are solved by applying methods of cluster analysis. These are classic Data Mining tasks: clustering, classification, and tasks typical only for text documents: automatic annotation, extraction of key concepts, etc. An abbreviated analysis of information clustering methods was conducted. The postulate that clustering involves dividing a set of elements into clusters, the number of which is determined by the localization of the elements of a given set in the vicinity of some natural centers of clusters, is confirmed. An analysis of the application of the Bayesian classifier was carried out. it is proven that the Bayesian classifier, in the presence of a priori probabilities, works with high accuracy in determining false information. However, this method does not give an answer about obtaining this probability. Due to the use of a naive Bayesian classifier to detect false information, the method of detecting false information has been improved. This method allows you to solve the problem of the uncertainty of the a priori probability. The proposed naive Bayesian classifier for text processing turned out to be quite effective. The effectiveness of the algorithm of the proposed method of developing a classifier for*

determining the veracity of information was evaluated. On the basis of the primary data obtained from the Internet, the numerical values of the evaluation of the algorithm of the improved method of determining false information were calculated. The obtained metric values are: Recall = 0.853; Precision = 0.869; F-measure = 0.861; Accuracy = 0.855. The obtained results prove that the improved method (without additional training) immediately has good results. This proves the adequacy of the developed method and provides an effective scientific method for detecting false information. The improved method of detecting false information is especially relevant at the present time, in the conditions of an information war.
Keywords: *clustering, classification, Bayes theorem, false information, threat, confrontation.*

Лукова-Чуйко Наталія Вікторівна, доктор технічних наук, професор, завідувачка кафедри кібербезпеки та захисту інформації, Київського національного університету імені Тараса Шевченка.
Nataliia Lukova-Chuiko, doctor of Technical Sciences, Professor, Department of Cyber Security and Information Protection Faculty of Information Technology Taras Shevchenko National University of Kyiv.

Лаптєва Тетяна Олександрівна, аспірантка, кафедра кібербезпеки та захисту інформації, Київського національного університету імені Тараса Шевченка.
Tetiana Laptieva, PhD-student, Department of Cyber Security and Information Protection**,** Taras Shevchenko National University of Kyiv.

Отримано *3 жовтня 2022 року,* затверджено редколегією *14 листопада 2022 року*