Технічні науки

УДК 004.9

**Федейко Юрій Володимирович**
*студент*
*Інституту прикладного системного аналізу*
*Національного технічного університету України*
*«Київський політехнічний інститут імені Ігоря Сікорського»*
**Федейко Юрий Владимирович**
*студент*
*Института прикладного системного анализа*
*Национального технического университета Украины*
*«Киевский политехнический институт имени Игоря Сикорского»*
**Fedeiko Yurii**
*Student of the*
*Institute of applied systems analysis of the*
*National technical university of Ukraine*
*"Ihor Sikorskiy Kyiv Politechnical Institute"*

***Науковий керівник:***
**Кухарєв Сергій Олександрович**
*асистент*
*Національний технічний університет України*
*«Київський політехнічний інститут імені Ігоря Сікорського»*

# ПОРІВНЯННЯ МЕТОДІВ ІДЕНТИФІКАЦІЇ СПАМУ
# СРАВНЕНИЕ МЕТОДОВ ИДЕНТИФИКАЦИИ СПАМА
# COMPARISON OF SPAM IDENTIFICATION METHODS

***Анотація.*** *Висвітлено застосування класифікаторів та нейронних мереж для задачі ідентифікації спама.*

***Ключові слова:*** *спам, класифікатор, Баєс, ідентифікація.*

***Аннотация.*** *Освещены применения классификаторов и нейронных сетей для задачи идентификации спама.*

***Ключевые слова:*** *спам, класификатор, Баєс, идентификация.*

***Summary.*** *The use of classifiers and neural networks for the problem of spam identification is covered.*

***Key words:*** *spam, classifier, Bayes, identification.*

На сьогоднішній день весь світ переходить в цифрову сферу і в ній, попри активну боротьбу, з'являється все більше і більше різного роду зловмисників які намагаються нажитися на простих людях. Завданням яке стоїть перед моєю роботою є хоч в певній мірі запобігання поширенню різного роду атак на користувачів в мережі інтернет.

Спам це масова розсилка рекламних повідомлень, які приходять без згоди отримувача. Зазвичай такі повідомлення або листи часто містять в собі віруси. В спам-розсилці часто зустрічаються шахраї, які всіма можливими способами будуть виманювати в користувачів конфіденційну інформацію для отримання коштів незаконним способом. Своєчасне розпізнавання спаму є важливим для безпеки людей, щодня сотні тисяч людей отримують спам повідомлення.

Перший зареєстрований в історії приклад спам розсилки відбувся в 1864 році, тоді деякі британські політики отримали неочікувану телеграму, що рекламувала стоматологічні послуги. Це відбулося в наслідок того, що компанія "Western Union" добавила можливість відправки телеграм в своїй мережі зразу багатьом користувачам. Як бачимо, з збільшенням можливостей для простих людей, збільшуються можливості і для зловмисників.

Як найактуальніші на сьогоднішній час розглядаються два варіанти вирішення задачі: класифікатор та нейронна мережа. Вони відрізняється абсолютно різним підходом до навчання та розпізнавання.

В машинному навчанні класифікацію розуміють як задачу визначення класу для нового об'єкта на основі емпіричних даних, які описують досліджувані зразки і відображають присутні їм властивості і закономірності. Існує залежність між зразками і класами, але вона невідома. Множина прецедентів, пар зразок-клас, складає навчальну вибірку, по якій знаходиться залежність, тобто будується алгоритм здатний для будь-якого зразка запропонувати відповідь, до якого класу він належить. Це приклад навчання з вчителем. Під вчителем в даному випадку розуміється навчальна вибірка із маркірованими повідомленнями для перевірки класифікації.

Сучасні статистичні дослідження [1] показали, що на сьогодні ймовірність будь-якого повідомлення бути спамом складає 80%. Однак більшість баєсових програм розпізнавання спама роблять припущення щодо відсутності апріорних переваг у повідомленнях бути спамом, і передбачає, що у обох випадків є рівні ймовірності 50%.

Про фільтри, які використовують дану гіпотезу, говорять як про фільтри "без упереджень".

На сьогоднішній день існує багато задач, для розв'язку яких застосовують класифікатори, такі як наприклад класифікація текстів по жанровим, авторським гендерним і іншим стилям чи розпізнавання семантичного забарвлення повідомлення автора. Дана теорія прийняття рішень складає основу статистичного підходу до задачі класифікації об'єктів. Цей підхід заснований на тому, що задача вибору рішення сформульована в термінах теорії ймовірності і відомі всі ймовірнісні величини, які важливі для даної задачі.

В основі даної класифікації лежить теорема Баєса. Дана теорема дозволяє визначити ймовірність будь-якої події при умові, що відбулась інша статистично пов'язана з нею подія. Іншими словами, по теоремі Баєса можна точніше перерахувати ймовірність події, врахувавши раніше відому інформацію та дані нових спостережень. Виведення теореми Баєса може бути виконане з основних аксіом теорії ймовірностей. Особливість даної теореми в тому, що для її застосування необхідна велика кількість обчислень, звідси баєсові оцінки стали активно застосовуватися тільки після розвитку обчислювальних машин [2].

Тепер повернемося до нейронної мережі, а саме розглянемо багатошаровий перцептрон. Ідею перцептрона запропонував нейрофізіолог Френк Розенблатт. Він запропонував схему, що моделювала процес людського сприйняття.

В загальному випадку перцептрон складається з трьох основних елементів: вхід, прихований шар і вихід. Основним завдання нейронної мережі є через повторювані ітерації знаходження вагових коефіцієнтів на зв'язках між елементами, щоб після навчання при поданні якогось сигналу на вхід, мережа могла дати чіткий вихід [4].

Тобто можна сказати, що баєсовий класифікатор по входженням слів рахує дві ймовірності для деякого повідомлення: бути спамом і бути звичайним повідомлення. Дальше йде порівняння того, яка ймовірність більша і на основі цього робиться фінальний висновок про повідомлення [3].

Нейронна мережа натомість, отримує на вхід одразу всі слова які присутні у навчальній вибірці і багатьма ітерація виставляє вагові коефіцієнти, щоб при поданні речення з певної кількості певних слів сигнал ніби пройшов по різних вагових зв'язках і на вихід подалась відповідь: є повідомлення спамом чи ні.

Порівнюючи ефективність роботи, можна сказати що нейронна мережа видає трохи кращий результат (порядку 1% точності в розпізнавання), але коли діло доходить до часу затраченого на навчання, то мережа показує в декілька тисяч разів гірший варіант за наївний баєсовий класифікатор (справляється з навчанням за декілька мілісекунд, а нейронній мережі для цього потрібно до хвилини часу).

Можна сказати що хвилина часу це не критично, але з спамом найчастіше стикаються онлайн платформи де іде швидкий потік інформації та в великих об'ємах і за хвилину часу може розгорнутися повноцінна спам атака, тому в таких випадках вигідніше буде використовувати наївний баєсовий класифікатор для розпізнання спаму та динамічного підлаштовування під нові загрози.

**Література**

1. "Більше 90% електронних листів є спамом**".** URL: http://certmag.com/more-than-90-percent-of-e-mails-in-third-quarter-were-spam/ (дата звернення: 30.05.2020)
2. Х. Цанг Оптимальність наївного баєса. URL: https://www.cs.unb.ca/~hzhang/publications/FLAIRS04ZhangH.pdf (дата звернення: 03.06.2020)
3. Н. Фрідман, Д. Гейгер, і М. Голдшмідт Класифікатор баєсової мережі. Машинне навчання. 1997. Vol. 29. С. 131-163.
4. Ерік Б. Баум Про можливості багатошарових перцептронів // Journal of Complexity. 1988. No 4. P. 193-215. URL: https://core.ac.uk/download/pdf/82339409.pdf (дата звернення: 04.06.2020)

## References

1. "More Than 90% of E-Mails Were Spam**".** URL: http://certmag.com/more-than-90-percent-of-e-mails-in-third-quarter-were-spam/ (date of application: 30.05.2020)
2. H. Zhang The Optimality of Naive Bayes. URL: https://www.cs.unb.ca/~hzhang/publications/FLAIRS04ZhangH.pdf (date of application: 03.06.2020)
3. N. Friedman, D. Geiger, and M. Goldszmidt Bayesian network classifiers. Machine Learning. 1997. Vol. 29, P. 131–163.
4. Eric B. Baum On the Capabilities of Multilayer Perceptrons // Journal of Complexity. 1988. No 4. P. 193-215. URL: https://core.ac.uk/download/pdf/82339409.pdf (date of application: 04.06.2020)