УДК 004.27

СИСТЕМА ВИЯВЛЕННЯ КІБЕРЗАГРОЗ НА ОСНОВІ АНАЛІЗУ ДАНИХ СОЦМЕРЕЖ З ВИКОРИСТАННЯМ МЕТОДІВ МАШИННОГО НАВЧАННЯ

Москаленко Д.М., здобувач вищої освіти, гр. мнПі-191
Таценко А.С., здобувач вищої освіти, гр. мнПі-191
Науковий керівник: **Дорош М.С.**, д.т.н., доцент
Національний університет «Чернігівська політехніка»

У сучасному світі одна з основних тенденцій розвитку глобальної мережі є активний розвиток соціальних мереж. Все більше користувачів долучаються до найбільш популярних соціальних мереж, таких як Facebook, Twitter, Instagram. Кожна з цих платформ дозволяє користувачам ділитися думками, новинами, давати свої власні оцінки щодо подій у світі, тощо. Оскільки використання соціальних мереж стає все більш впровадженим в повсякденне життя користувачів, соціальні мережі є одним з найнебезпечніших місць у кіберпросторі, оскільки ними користуються не тільки доброзичливі для користувача люди. В соціальних мережах існують загрози, що стосуються дуже багатьох складових благополуччя життя користувача. Недоброзичливці можуть нанести шкоду матеріальному статусу, репутації, технічному/програмному забезпеченню, фізичному та ментальному здоров'ю користувача соціальних мереж. Велика кількість людей діляться своєю думкою через Інтернет, тому в епоху інформації даний підхід до аналізу думки громадськості та окремих користувачів вимагає своєї уваги і є достатньо актуальним. В області класифікації настроїв аналізуються думки або настрої людей на предмет особистих образ, нецензурної лексики, токсичної поведінки, персональних погрозах[1]. Системи аналізу настроїв застосовуються в соціальних платформах і майже в кожному бізнесі, тому що думки або настрої відображають переконання, вибір і діяльність людей. За допомогою цих систем також можна приймати рішення для бізнесу, в політичних цілях і тому подібне [2].

Об'єктом дослідження є перш за все задача інтелектуального аналізу повідомлень соціальної мережі Twitter, а також різні методи і архітектури машинного навчання для вирішення завдань аналізу природної мови. Метою є розробка методики виявлення потенційно небезпечних повідомлень у соціальних мережах на прикладі розробки класифікатора емоційного тону повідомлень користувачів соціальної мережі Twitter за змістом їх публічних повідомлень (твітів) для аналізу безпеки аккаунтів користувачів.

Методами дослідження стали: метод обробки текстових даних до векторного виду та згорткова нейронна мережа. Задача класифікації настроїв була вирішена за допомогою алгоритму Випадкового Лісу[3].

Згорткові нейронні мережі були спочатку розроблені для комп'ютерного зору, наприклад для аналізу зображень. З 2011 року вони стають все більш популярними в NLP. У 2014 році CNN була успішно використана в задачах обробки текстів, таких як класифікація пропозицій і моделювання пропозицій [4]. CNN складається з згортаючого і об'єднуючого шарів. Вхідні дані передаються в ланцюжок з одного або декількох згортальних шарів, за якими розташовується шар пулу. Ця структура шару зі згорточним пулом може повторюватися для створення більш глибоких змін, які утворюють глибокі CNN. Згорточний шар застосовує шаблон фільтра до вхідних даних. Це допомагає зіставити вхідні дані з базовою картою об'єктів і знайти локальні зв'язки між ними. Шаблон фільтра зменшує кількість ваг в мережі, зіставляючи частини вхідних даних однієї функції. Крім того, шар об'єднання скорочує карту об'єктів і об'єднує семантично подібні об'єкти разом. Згорточний шар може складатися з більш ніж одного шаблону фільтра, що показано наведеною структурою на рисунку 1. Слова передаються в мережу і фільтруються згорточними шарами. Після об'єднання FNN з шаром softmax прогнозує клас повідомлення.

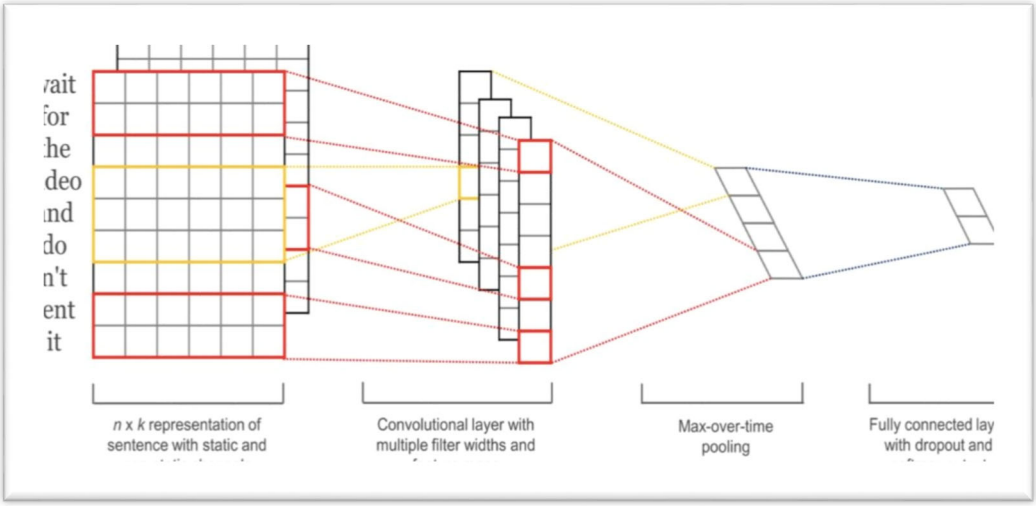

Рисунок 1 – Структура згорточного шару

Для реалізації системи було обрано мікросервісну архітектуру[4], яка складається з двох основних сервісів. Перший для отримання інформації про аккаунти за допомогою Twitter API, та перетворення, очистки текстових даних для подальшої обробки другим сервісом - нейромережевим модулем (рисунок 2)

Нейромержевий модуль представляє собою сервіс для класифікації текстових даних на основі моделі, побудованої на розміченій вибірці з більш ніж 160 000 заголовків новин. В подальшому оброблена інформація про настрої користувачів класифікується та передається до інтерфейсу користувача.

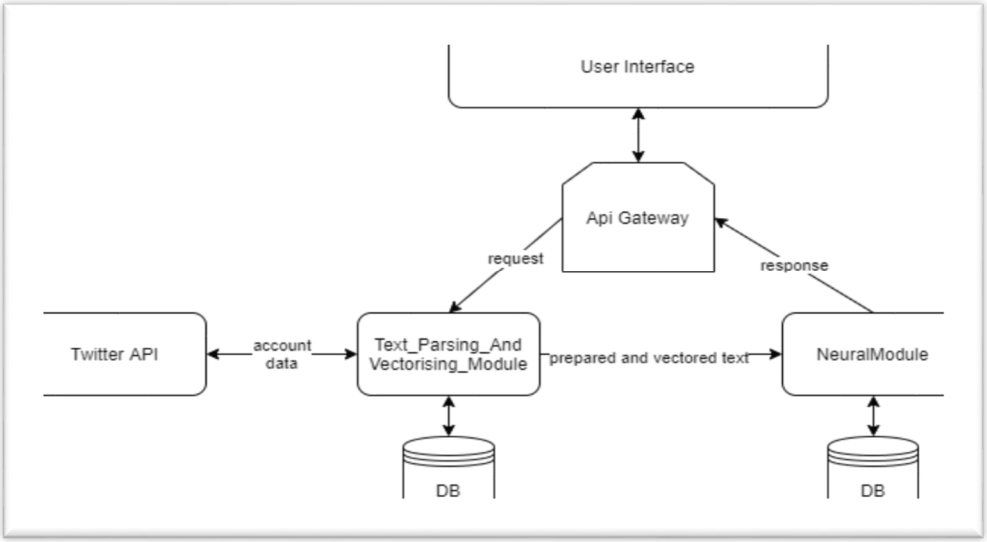

Рисунок 2. – Структура нейромережевого модуля

Завдяки використанню мікросервісної архітектури система може відображати користувацький інтерфейс як на мобільних застосунках так і на веб-браузерах, а також дозволяє налаштовувати параметри нейромережі, моделей і обирати API інших соціальних мереж, таких як Instagram та Facebook [5]. Вибір налаштувань та параметрів обмежуються тільки бажанням замовника та доступності API інших соціальних мереж, як дата сетів для навчання моделі нейромережі.

Список використаних джерел

1. Кібербулінг [Електронний ресурс]. – Режим доступу http://security.mosmetod.ru/internet-zavisimosti/98-kiberbulling

2. Online Social Networks: Threats and Solutions https://arxiv.org/abs/1303.3764
3. A Comprehensive Guide to Convolutional Neural Networks [Електронний ресурс]. – Режим доступу: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53
4. Microservice Architecture [Електронний ресурс]. – Режим доступу: http://microservices.io/
5. Преимущества микросервисной архитектуры [Електронний ресурс]. – Режим доступу: https://otus.ru/nest/post/1012

УДК 004.658.2

ПОРІВНЯННЯ СУЧАСНИХ CASE-ТЕХНОЛОГІЙ ПРОЄКТУВАННЯ БАЗ ДАНИХ

Нагорний П. В., здобувач вищої освіти групи ПІ-181
Науковий керівник: **Білоус І. В.**, к.т.н., доцент
Національний університет «Чернігівська політехніка»

На сучасному етапі розвитку IT-технологій бази даних та системи роботи з ними посідають провідне місце у розробці складних інформаційних систем. При цьому спостерігається тенденція до ускладнення структури проєктованих баз даних. Зважаючи на це, особливої актуальності набувають CASE-технології проєктування баз даних. Під CASE-технологією розуміємо методологію та набір конкретних інструментів аналізу, моделювання, проєктування, окремих аспектів розробки та супроводження інформаційних систем програмного забезпечення [1]. Одним з головних призначень CASE-технологій є відокремлення та автоматизація процесу проєктування від розробки.

Сучасні CASE-технології проєктування баз даних можна порівнювати за кількома критеріями. В першу чергу, це функціональні можливості. Одним з найбільш важливих компонент засобів проєктування баз даних є візуальний конструктор, який надає можливість зручної графічної побудови моделі бази даних та подальшої реалізації самої бази даних на основі побудованої моделі. Не всі сучасні CASE-технології пропонують вказану можливість, тому можемо виділити дві групи засобів проєктування баз даних: неповнофункціональні (реалізують лише обмежений набір можливостей, візуальний конструктор відсутній), повнофункціональні (реалізують майже всі необхідні розробнику інструменти проєктування, візуальний конструктор наявний) [1].

Інший критерій порівняння засобів проєктування баз даних – вартість технології. Комерційні технології часто пропонують дещо ширший перелік функціональних можливостей, проте для більшості проєктувальників та розробників переважно достатньо функціонала некомерційних технологій.

Як вже було зазначено, існує велика кількість комерційних та некомерційних засобів проєктування баз даних. Розглянемо деяких популярних представників, їх переваги та недоліки [2].

1) *MySQL Workbench*. Інструмент для візуального проєктування та подальшої розробки баз даних MySQL. Передбачаємо можливості проєктування, розробки, адміністрування баз даних. Існує в декількох версіях: Community, Standard, Enterprise. Видання Community є безкоштовним та включає більшість функцій, необхідних користувачу. Основним практичним обмеженням можна назвати підтримку виключно MySQL.

2) *dbFrogeStudio.* Створені окремі програмні продукти для MySQL, PostgreSQL, Oracle, SQL Server. Передбачає підтримку розробки запитів, дизайну баз даних, адміністрування, роботи з даними, аналізу даних. Безкоштовна версія Express досить обмежена та містить лише