

РЕФЕРАТ

Розмір пояснювальної записки – 109 аркушів, містить 23 ілюстрації, 38 таблиць, 3 додатки, 38 посилань на джерела.

Актуальність теми. У роботі розглянуто проблему в області медицини, а саме діагностики захворювань. Проаналізувавши наявні дослідження обраного об'єкта, а саме програмного забезпечення для класифікації рентгенівських знімків, було виявлено, що існуючі рішення проблеми фокусуються на точності діагностування захворювань. Питання зменшення часу тренування моделей наразі не досліджене, незважаючи на важливість даної теми – налагодження параметрів моделі прийнятної точності потребує багато часу, від декількох місяців і до року. В зв'язку з цим виявлено потребу в розробці архітектурного рішення, що пришвидшить час тренування моделей.

Мета дослідження. Основною метою є зменшити час тренування моделей діагностування захворювань по рентгенівських знімках за рахунок використання розподілених технологій.

Об'єкт дослідження: програмне забезпечення для класифікації рентгенівських знімків.

Предмет дослідження: процеси розроблення, модифікації, аналізу, забезпечення якості, впровадження і супроводження програмного забезпечення для класифікації рентгенівських знімків.

Для реалізації поставленої мети **сформульовані наступні завдання:**

- аналіз наявних досліджень методів та рішень діагностики захворювань по рентгенівських знімках;
- аналіз наявних математичних методів класифікації зображень та обґрунтування вибору методу класифікації;
- аналіз наявних програмних бібліотек, які реалізують обраний математичний метод;
- реалізація програмного коду, який вирішує поставлене завдання з використанням обраних методів;

- розробка графічного інтерфейсу для демонстрації програмного забезпечення;
- проведення маркетингового аналізу проєкту;

оцінка ефективності запропонованого рішення на основі проведення обчислювальних експериментів.

Наукова новизна результатів магістерської дисертації полягає в подальшому розвитку задачі розпізнавання захворювань по рентгенівських знімках, а саме створенні нового підходу до навчання моделей діагностики захворювань шляхом використання архітектурного рішення, яке включає в себе розподілену обробку даних на основі інтегрованого рішення Apache Spark та Tensorflow.

Практичне значення отриманих результатів полягає в тому, що розроблено високоефективне архітектурне рішення для створення, порівняння та застосування моделей діагностування декількох типів захворювань по даним рентгенівських знімків грудної клітини. Дане рішення може бути використане науковцями та програмістами для створення, тестування та покращення моделей діагностики захворювань за рентгенівськими знімками.

Зв'язок з науковими програмами, планами, темами. Робота виконувалась на кафедрі інформатики та програмної інженерії Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського» в рамках теми «Методи та технології високопродуктивних обчислень та обробки надвеликих масивів даних». Державний реєстраційний номер 0117U000924.

Апробація. Наукові положення дисертації пройшли апробацію на Третій Всеукраїнській науково-практичній конференції молодих вчених та студентів «Інженерія програмного забезпечення і передові інформаційні технології» (SoftTech-2022 Осінь) – м. Київ.

Публікації. Наукові положення дисертації опубліковані в тезах доповідей на науково-технічній конференції «SoftTech-2022 осінь».

Ключові слова: ДІАГНОСТИКА ЗАХВОРЮВАНЬ, РЕНТГЕНІВСЬКІ ЗНІМКИ, РОЗПОДІЛЕНА ОБРОБКА ДАНИХ, APACHE SPARK.