

Дипломна робота на здобуття ступеня бакалавра на тему: “Модуль розпізнавання та класифікації звуків”

1

Автор: студент 4го курсу
Групи КА-54
Бондар О. А.
Керівник: доцент, к. т. н.
Тимощук О. Л.

Актуальність роботи



- Розробка програмного продукту, що дозволить людям з проблемами зі слухом сприймати навколишнє середовище.
- На сьогодні, досягнення у сфері штучного інтелекту поступово впроваджуються у наше життя, однак досі існує невелика кількість додатків з його використанням. Розроблений модуль може бути розглянутий як приклад інтеграції штучного інтелекту та кросплатформенного додатку

Предмет та об'єкт дослідження

- Об'єкт дослідження – алгоритми в сфері машинного навчання, а саме алгоритми і методи аналізу звуків
- Предмет дослідження – моделі машинного навчання, а саме згорткові нейронні мережі.

Мета роботи

- Створення моделі, що здатна з прийнятною точністю відносити наданий звук до певного класу за допомогою сучасних методів машинного навчання
- Побудова інтелектуального модулю, що може бути використаний користувачами різних мобільних платформ, а також буде використовувати вищевказану модель і буде зручним у використанні

Постановка задачі

- Розглянути існуючі методи та прикладні завдання машинного навчання
- Детально дослідити існуючі моделі, що використовуються для аналізу звуків
- Створити та натренувати власну модель для успішної класифікації звуків
- Розробити модуль для кінцевого користувача на основі вищезгаданої моделі

Технології, що були використані в проекті



Telegram API



Keras + Tensorflow

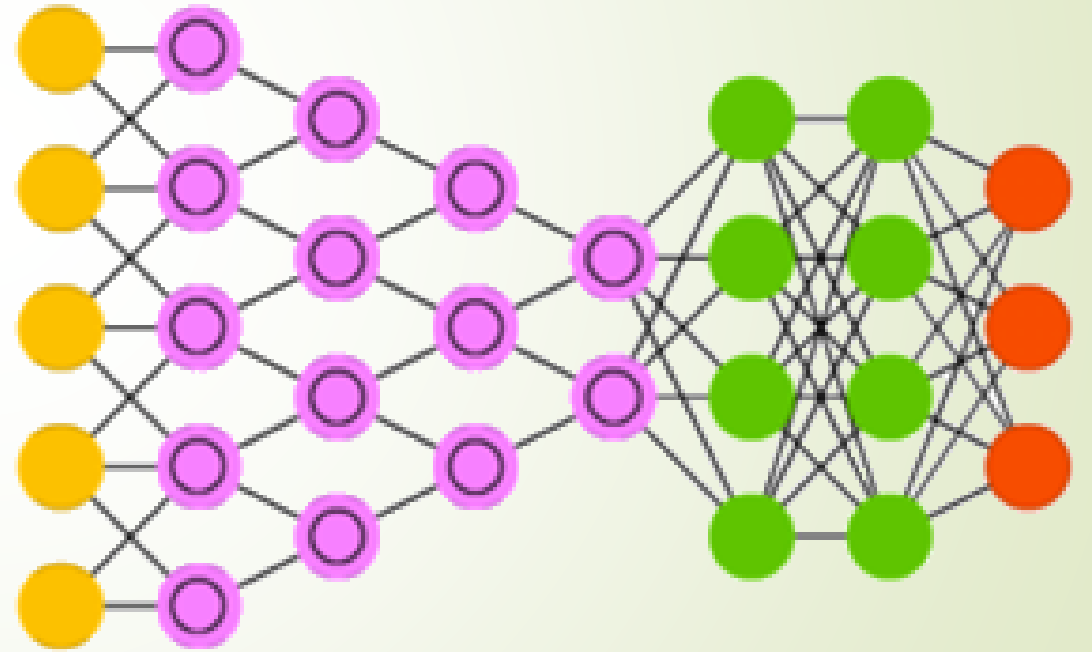
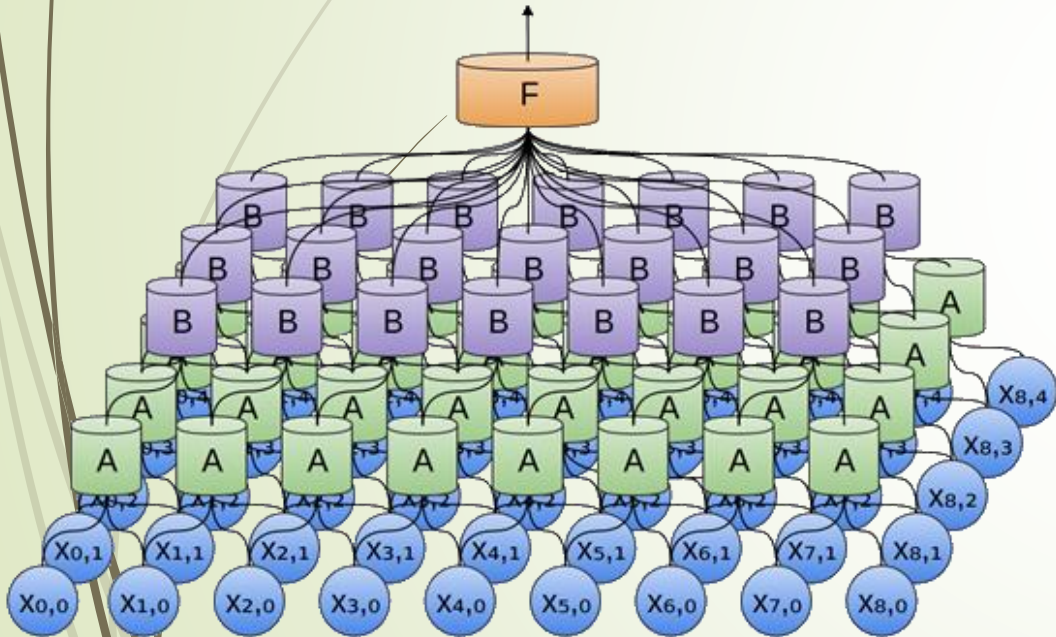
Постановка задачі класифікації

- Необхідно при надходженні на вхід *.wav файлу, визначити його належність до певного класу

Дані для навчання

- Для навчання моделі був використаний датасет із платформи Kaggle під назвою “Urban Sound Classification”
- Всього у датасеті наявно 8732 класифікованих звука тривалістю менше 4 секунд, які належать до одного з 10 класів.

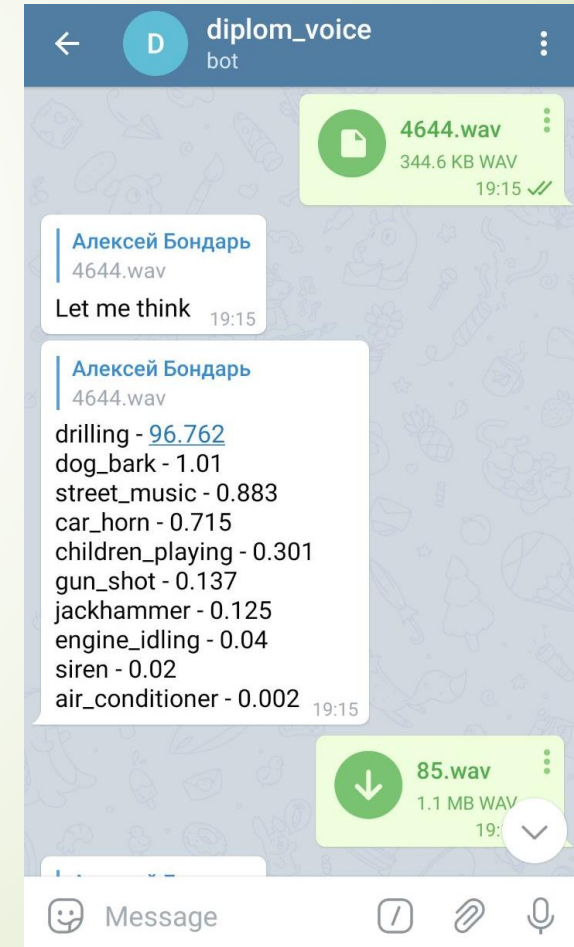
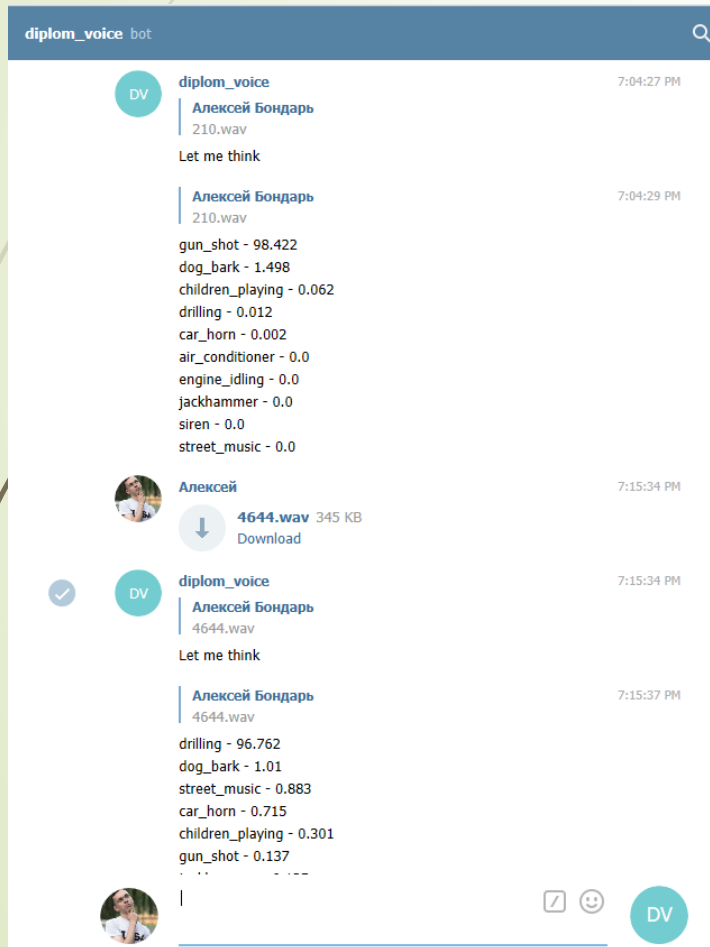
Модель на основі згорткової нейронної мережі



Активаци́йні функції нейронів

- ▶ Relu: $f(x) = x^+ = \max(0, x)$
- ▶ Softmax: $\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$; $i = 1, \dots, K$; $z = (z_1, \dots, z_K) \in R^K$

Інтерфейс користувача



Результати роботи

Характеристика	Модель
Точність	96.3%
Час, затрачений на навчання (хв)	75
Швидкість обробки (мс)	1221
Розміри моделі (МБ)	10.5

Подальша робота

Подальша робота може проводитись у кількох напрямках:

- Покращення моделі для коректної роботи із зашумленими записами.
- Впровадження нових команд модулю, додання більшої інформативності результату, його візуалізація, збільшення кількості класів.

Дякую за увагу!

