Федеральное государственное бюджетное образовательное учреждение высшего образования
«МИРЭА – Российский технологический университет»
РТУ МИРЭА

Институт кибербезопасности и цифровых технологий
Кафедра КБ-14 «Цифровые технологии обработки данных»

Классификация сетевого трафика с использованием различных архитектур нейронных сетей

Доклад
по выпускной квалификационной работе бакалавра по направлению подготовки
09.03.02 – Информационные системы и технологии

Студент: Челышев С.Р. БСБО 09-18

Научный руководитель: Оцоков Ш.А.

Москва 2022 г.

Цель и задачи, решаемые в работе

Целью выпускной квалификационной работы является разработка моделей машинного обучения для классификации сетевого трафика.

Для достижения поставленной цели в работе сформулированы следующие основные задачи:

- Сделать обзор существующих решений для анализа трафика;
- Исследовать методы машинного обучения, используемые для классификации трафика;
- Реализация моделей машинного обучения для классификации трафика.

Актуальность и проблема

- Одна из существующих проблем сетей – это предоставление недостаточного качества обслуживания и сложности с детектированием атак или трафика вредоносного ПО.
- Главной задачей является реализация моделей машинного обучения для классификации трафика.
- Классификация трафика важна для правильного распределения сетевых ресурсов, имеет большое значение для сложных задач управления сетью, таких как обеспечение надлежащего качества обслуживания (Quality of Service), обнаружение аномалий и детектирования атак, этим обусловлена актуальность работы.

Пример набора данных

Flow.ID	Source.IP	Source.Port	Destination .IP	Destination .Port	Protocol	Timestamp
172.19.1.46-10.200.7.7-52422-3128-6	172.19.1.46	52422	10.200.7.7	3128	6	26/04/2021 11:11:17
172.19.1.46-10.200.7.7-52422-3128-6	10.200.7.7	3128	172.19.1.46	52422	6	26/04/2021 11:11:17
10.200.7.217-50.31.185.39-38848-80-6	50.31.185.39	80	10.200.7.217	38848	6	26/04/2020 11:11:17

Реализация

Программные средства:

1. Язык программирования Python

2. Библиотеки - Pandas, Numpy, Sklearn

3. Jupyter Notebook – среда разработки, где можно сразу увидеть результат работы кода или отдельных его компонентов

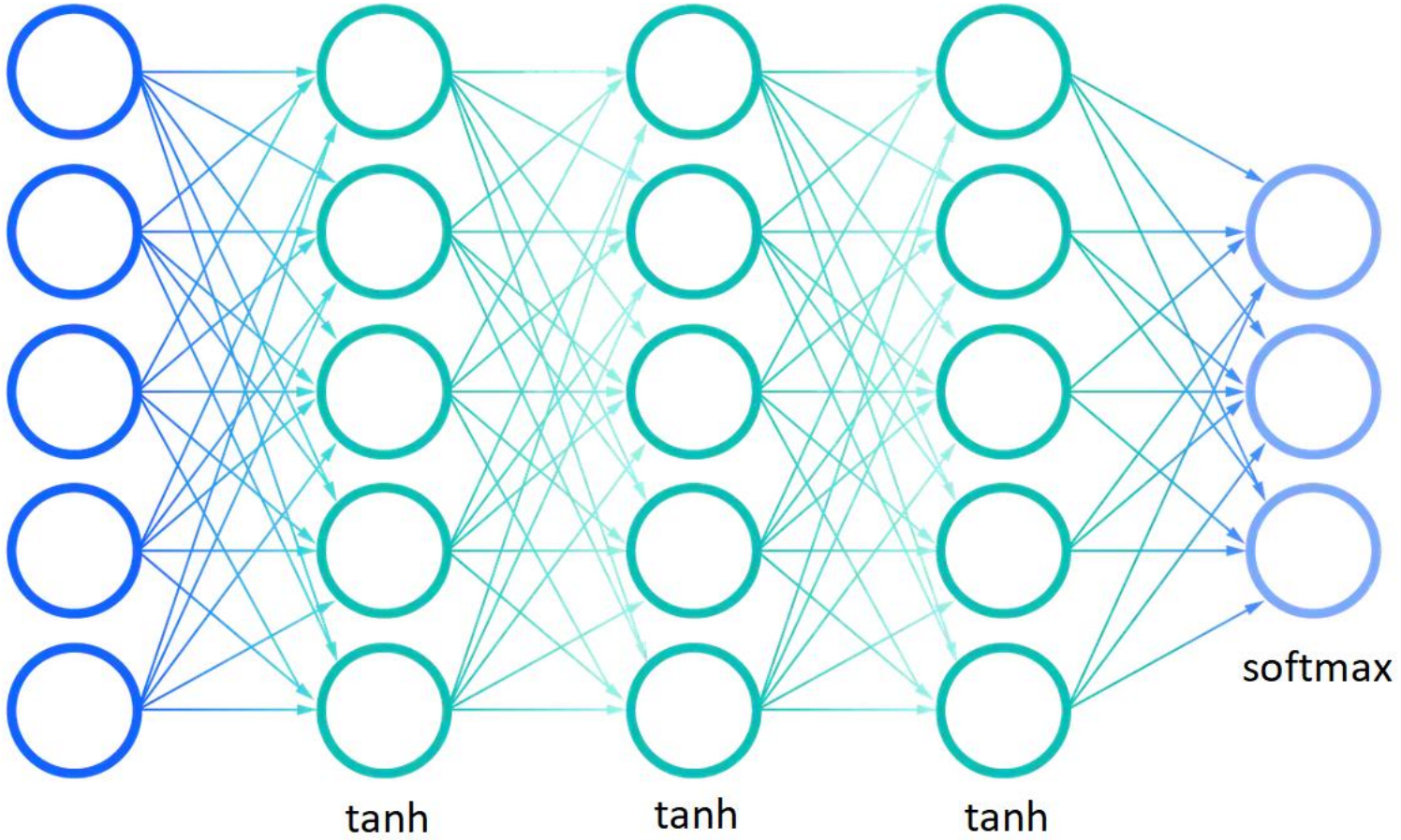
4. Matplotlib, seaborn, networkx – для визуализации

Описание программной части

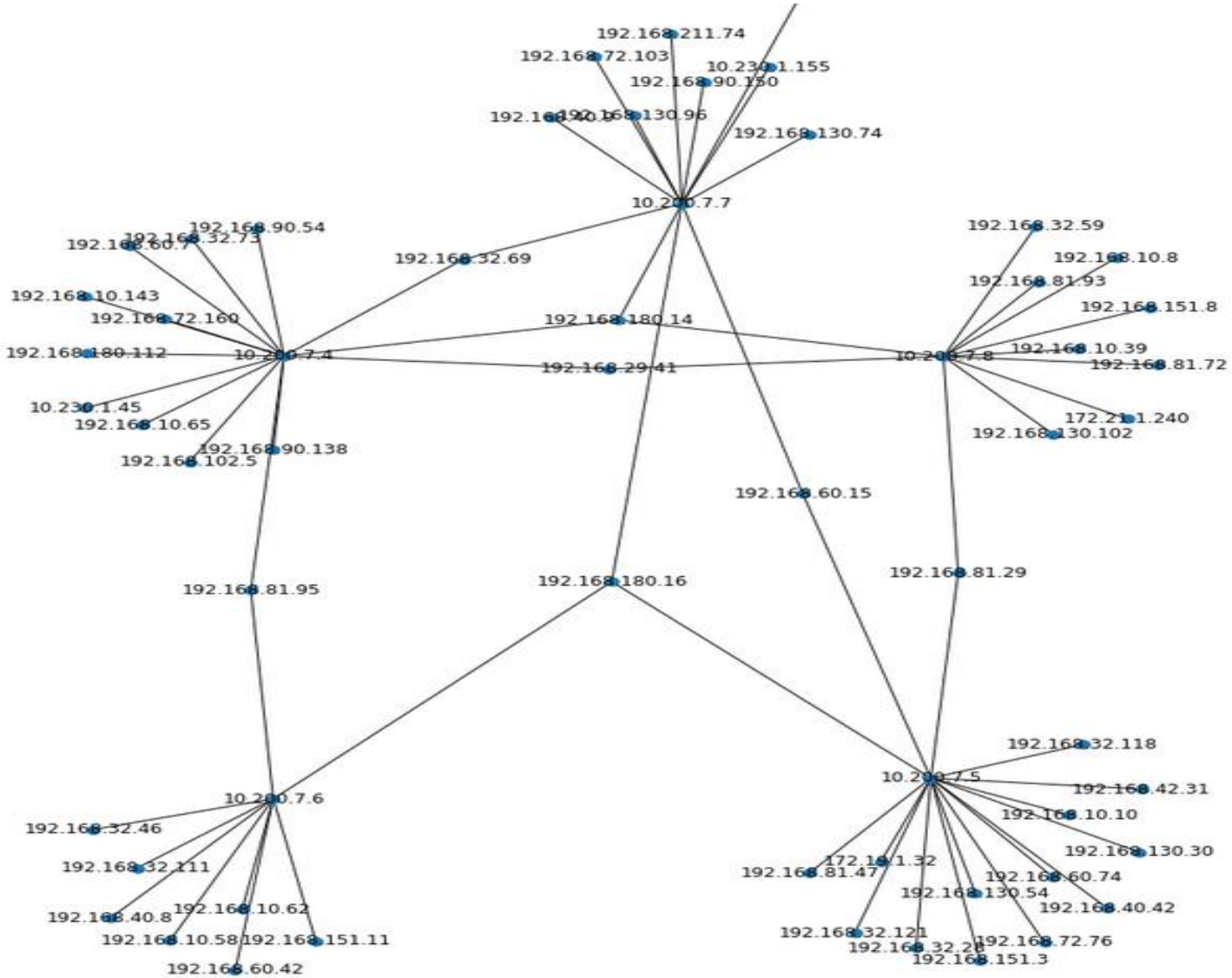
- При использовании метода опорных векторов используется класс sklearn.svm.SVC из библиотеки sklearn. Эта реализация использует стратегию один-к-одному для мультиклассовой классификации, поскольку метод опорных векторов поддерживает только бинарную классификацию.

- Реализованная нейронная сеть состоит из 5 слоёв, в первых 4 из которых 100 нейронов, а в последнем число нейронов равно количеству классов. На выходном слое используется функция активации softmax. На скрытых слоях используется функция активации tanh.

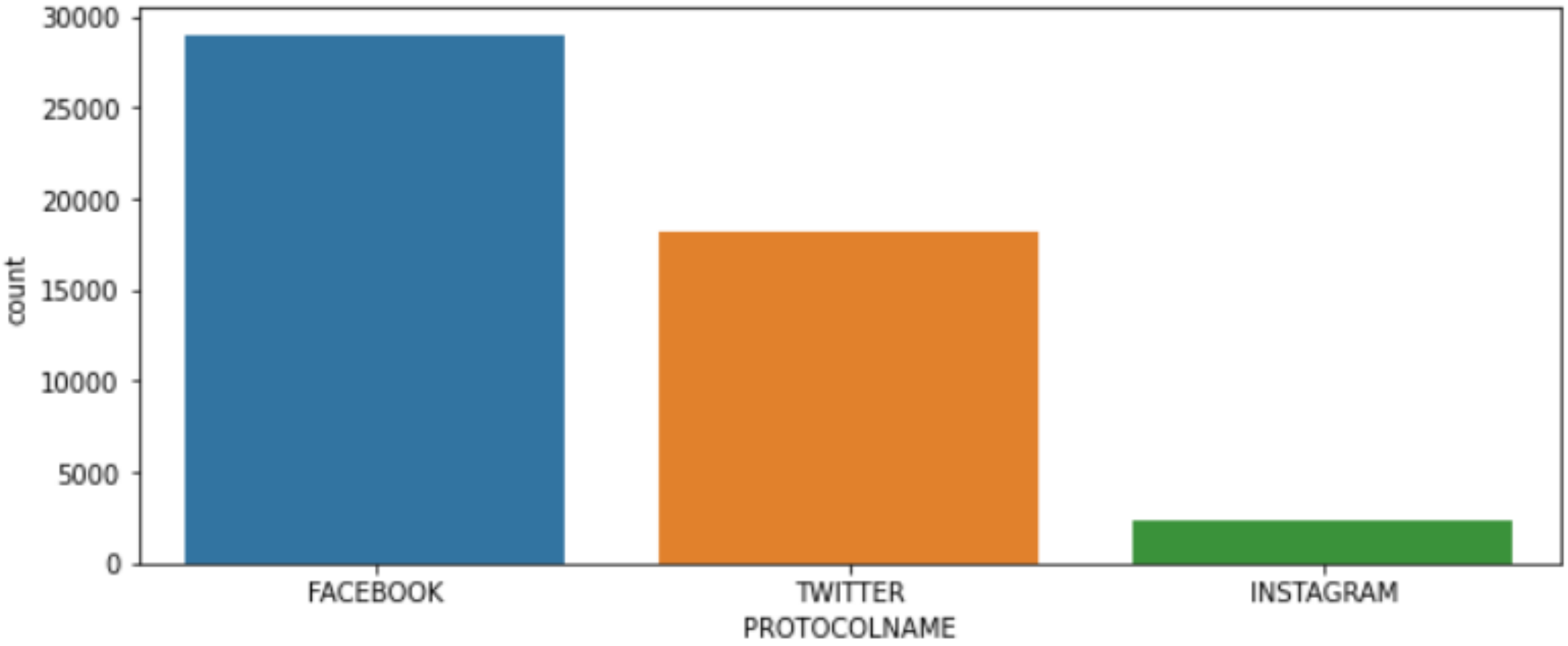
Архитектура нейронной сети



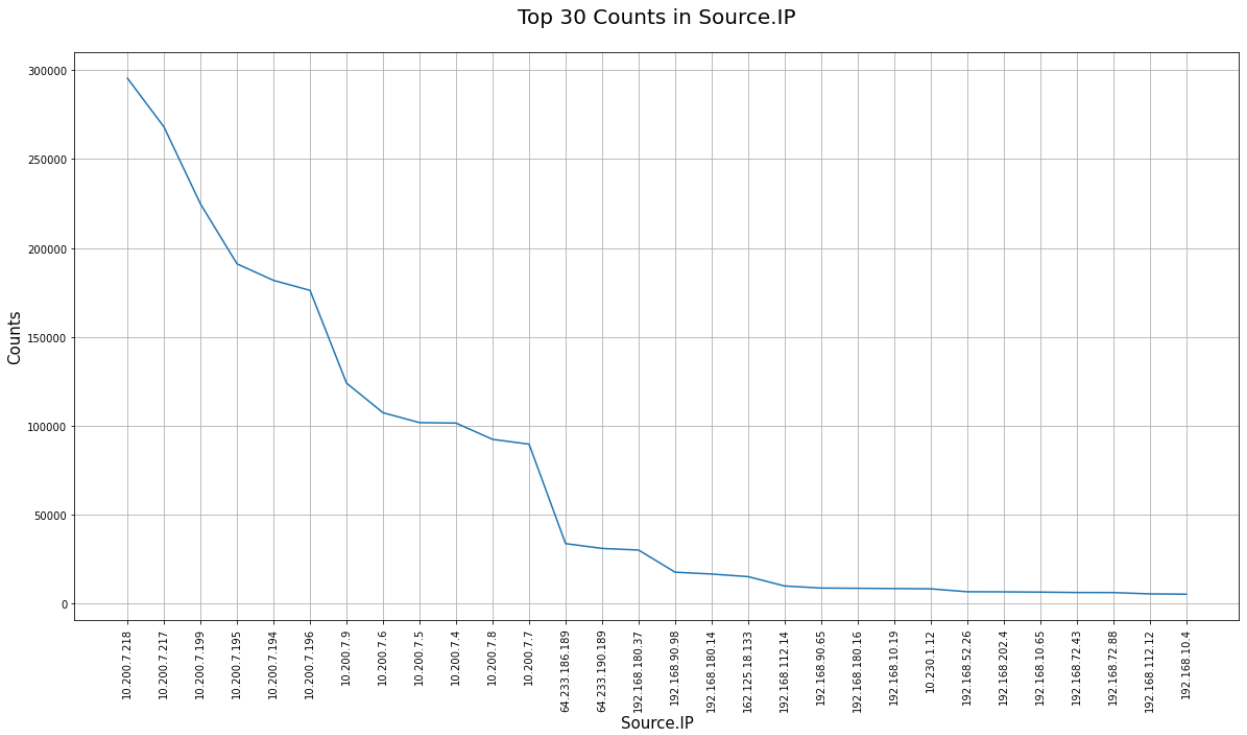
Граф связей сети

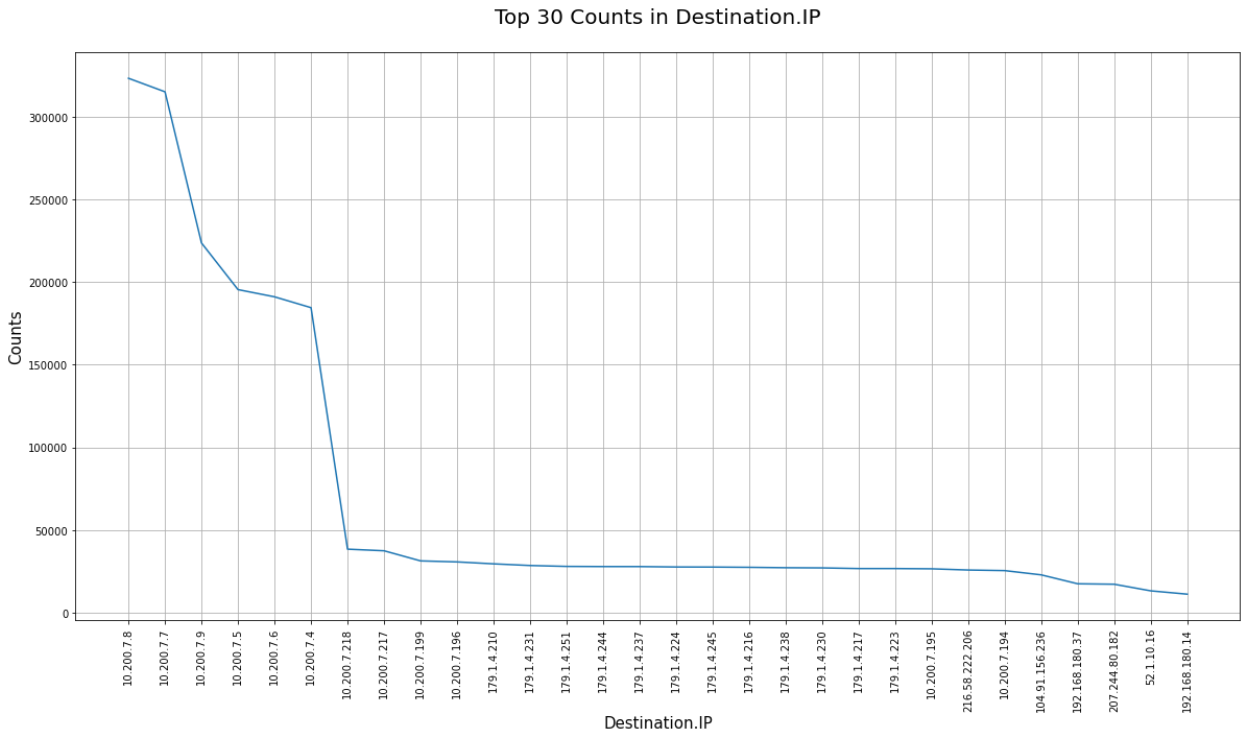


Наиболее часто встречаемые приложения

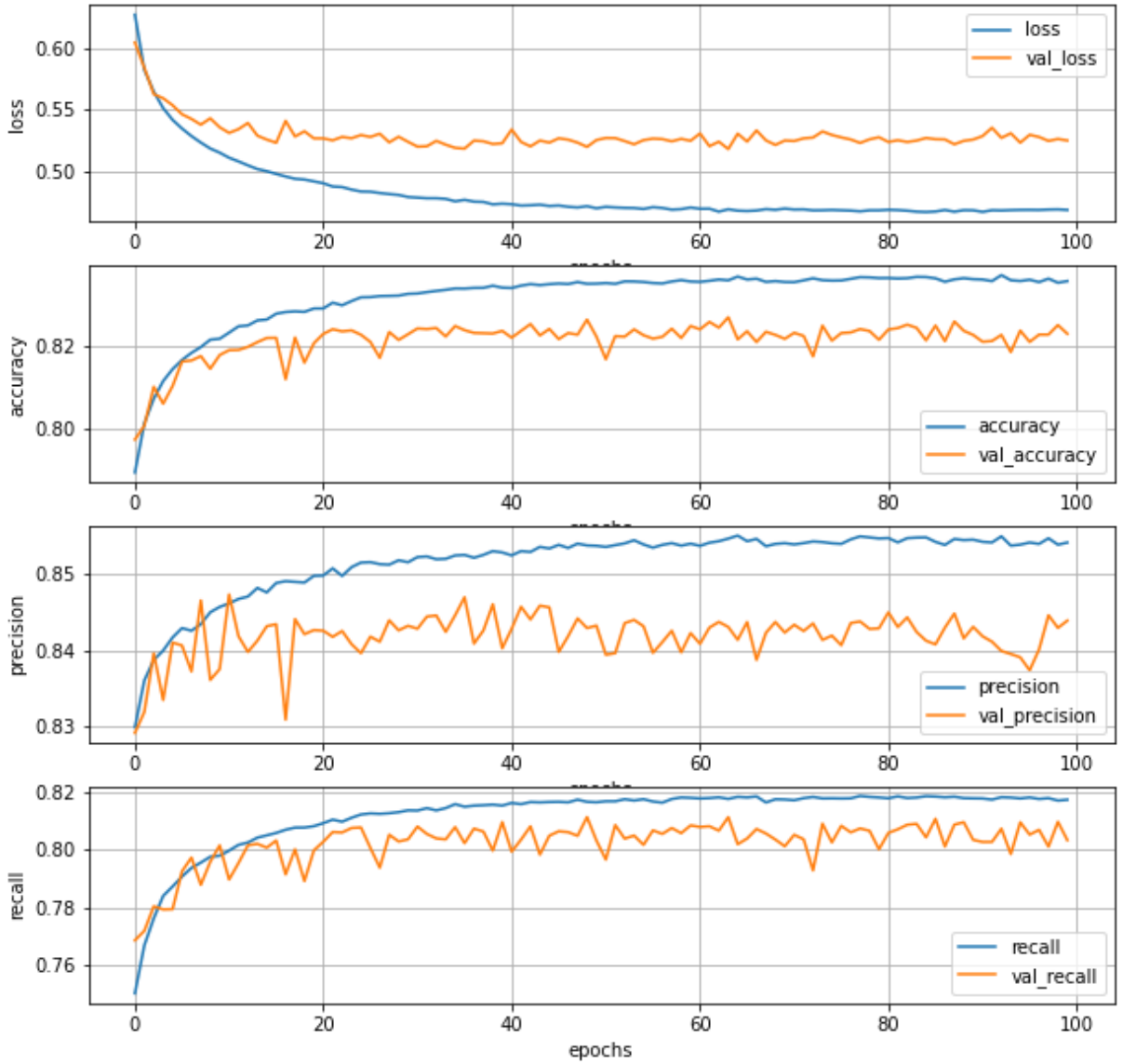


Наиболее часто встречающиеся IP





Результаты обучения нейронной сети



Точность определения конкретных приложений

Приложений	Точность определения
Amazon	1.0
Wikipedia	0.65
WhatsApp	0.33

Заключение по ВКР

В рамках ВКР было выполнено:

1. Сделан обзор существующих решений для анализа трафика.
2. Проведено исследование методов машинного обучения, используемых для классификации трафика.
3. Реализованы модели машинного обучения для классификации трафика.

Спасибо за внимание