МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Системы обработки информации и управления»

ОТЧЕТ

**Лабораторная работа № 6**

по дисциплине«Разработка нейросетевых систем»

Тема: «Классификация LiDAR»

ИСПОЛНИТЕЛЬ: \_\_\_\_Гаврилов Л.Я.\_\_

ФИО

группа ИУ5-25М \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись

"20" мая 2024 г.

ПРЕПОДАВАТЕЛЬ: \_\_\_\_\_\_\_\_

ФИО

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись

"\_\_"\_\_\_\_\_\_\_\_\_2024 г.

Москва – 2024

Задание

Обучите модель нейронной сети для классификации пород деревьев, используя файлы облака точек.

1. Укажите файл по вашему варианту и измените список классов, которые относятся к вашему варианту. Обучите модель.
2. Измените гиперпараметры обучения для улучшения модели: количество эпох, размер батча, скорость обучения.
3. Укажите, какие действия помогли улучшить метрики ваших моделей и объясните почему.

Ход работы

Обучим базовый вариант модели. Графики точности и функций потерь представлены на рисунках 1 и 2.

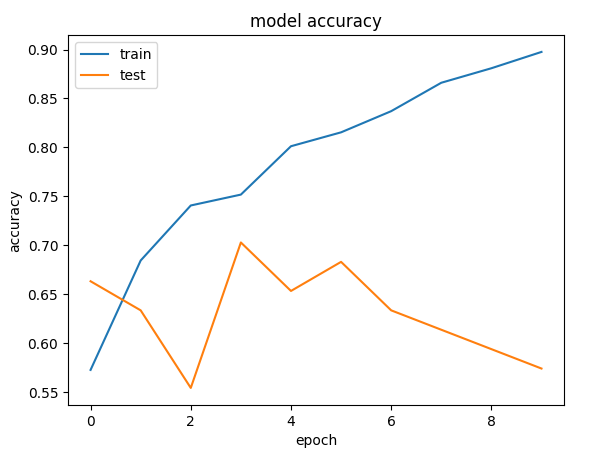


Рисунок 1 – График точности

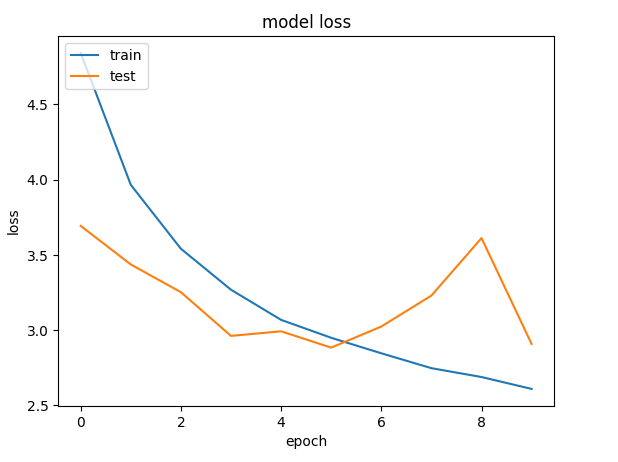


Рисунок 2 – График функции потерь

Получим значения метрики *accuracy* на обучающей выборке = 89,75% и тестовой выборке = 57,43%.

На рисунке 3 представлена матрица ошибок.

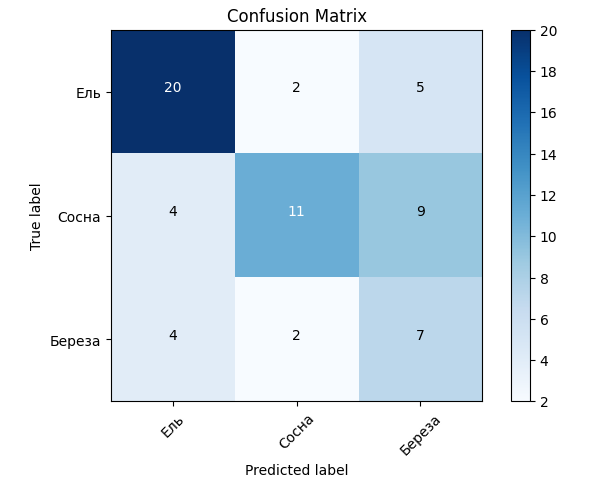


Рисунок 3 – Матрица ошибок

Уменьшим количество эпох до 5, чтобы избавиться от переобучения. Графики точности и функций потерь представлены на рисунках 4 и 5.

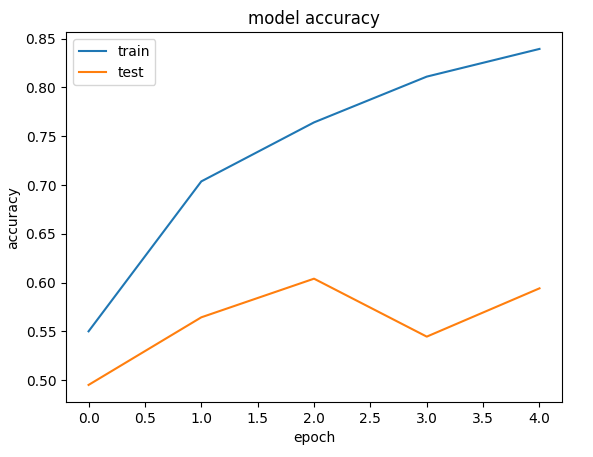


Рисунок 4 – График точности

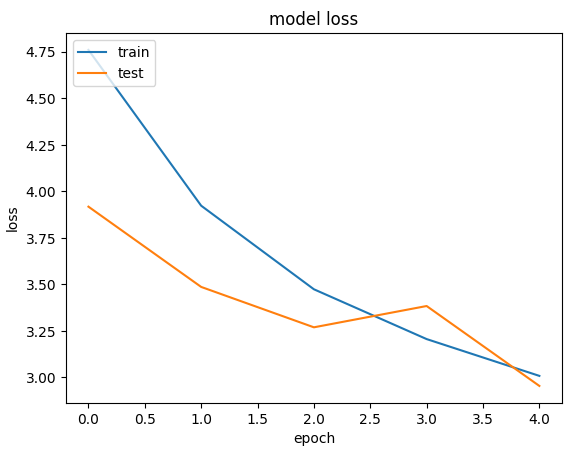


Рисунок 5 – График функции потерь

Получим значения метрики *accuracy* на обучающей выборке = 83,95% и тестовой выборке = 59,41%.

На рисунке 6 представлена матрица ошибок.

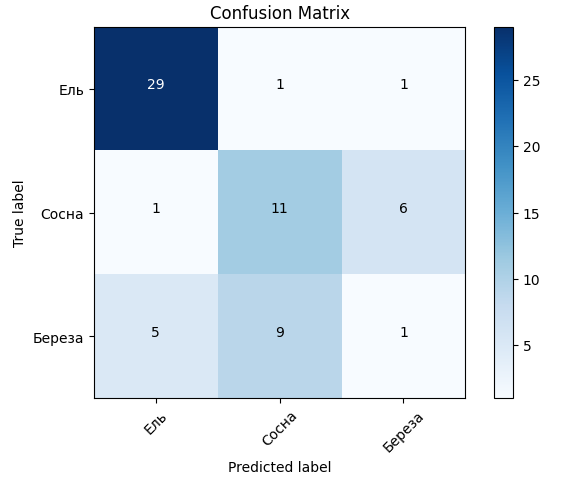


Рисунок 6 – Матрица ошибок

Уменьшим в модели базового варианта скорость обучения в 2 раза (lr=0.0005) и увеличим количество эпох в 2 раза (epochs=20). Графики точности и функций потерь представлены на рисунках 7 и 8.

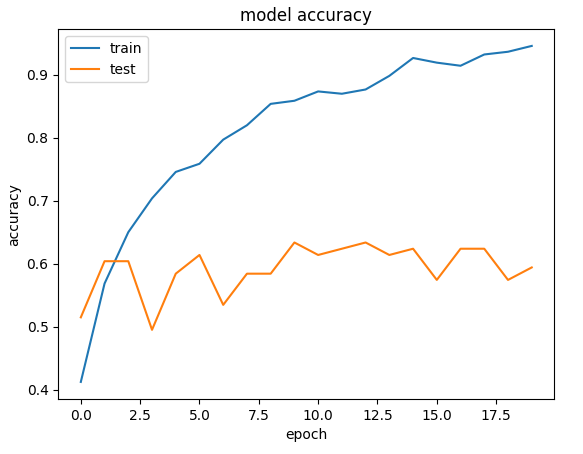


Рисунок 7 – График точности

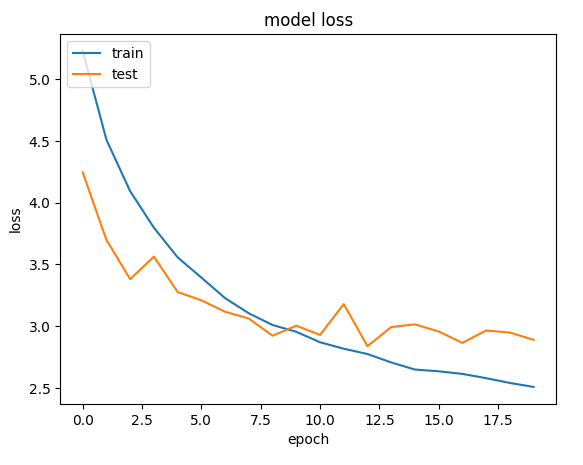


Рисунок 8 – График функции потерь

Заметим, что на 10-ой эпохе происходит переобучение. Получим значения метрики *accuracy* на обучающей выборке = 94,57% и тестовой выборке = 59,41%.

На рисунке 9 представлена матрица ошибок.

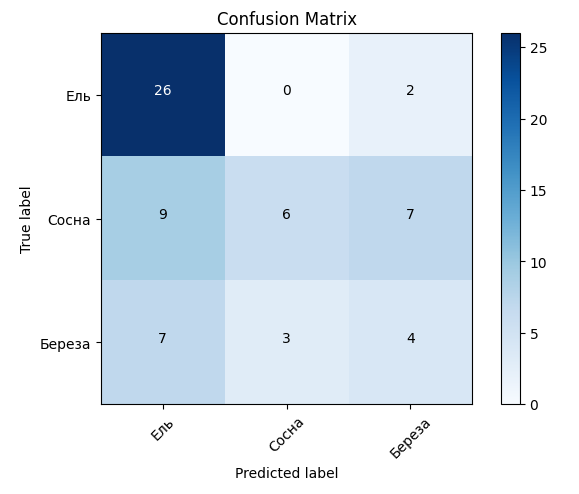


Рисунок 9 – Матрица ошибок

Уменьшим количество эпох для модели №3 до 10, чтобы избавиться от переобучения. Графики точности и функций потерь представлены на рисунках 10 и 11.

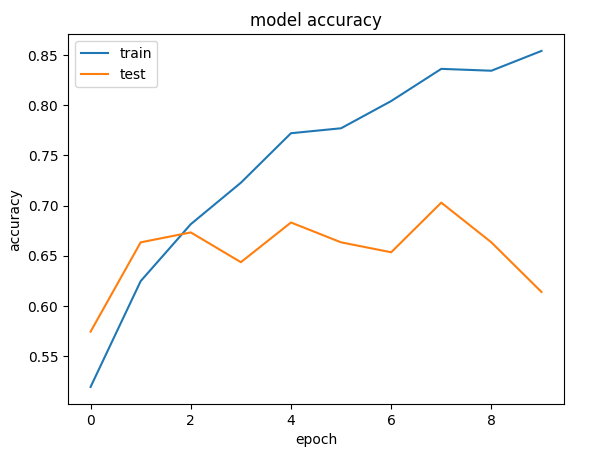


Рисунок 10 – График точности

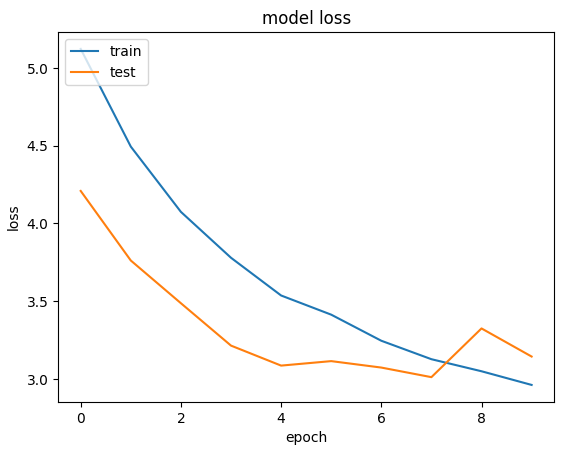


Рисунок 11 – График функции потерь

Получим значения метрики accuracy на обучающей выборке = 85,43% и тестовой выборке = 61,39%.

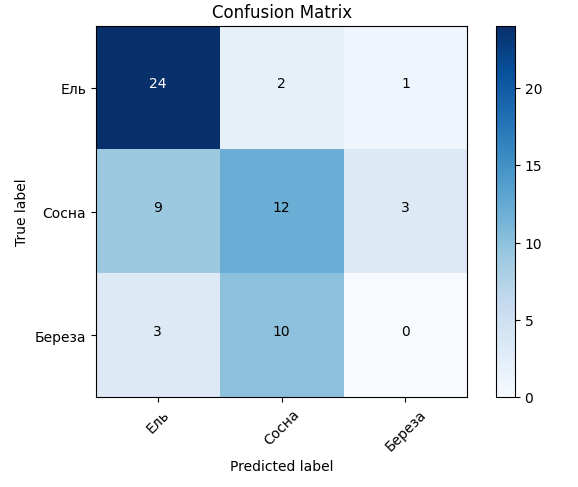
На рисунке 12 представлена матрица ошибок. 

Рисунок 12 – Матрица ошибок

Увеличим размер батча и количество эпох базовой модели в 2 раза. Графики точности и функций потерь представлены на рисунках 13 и 14.

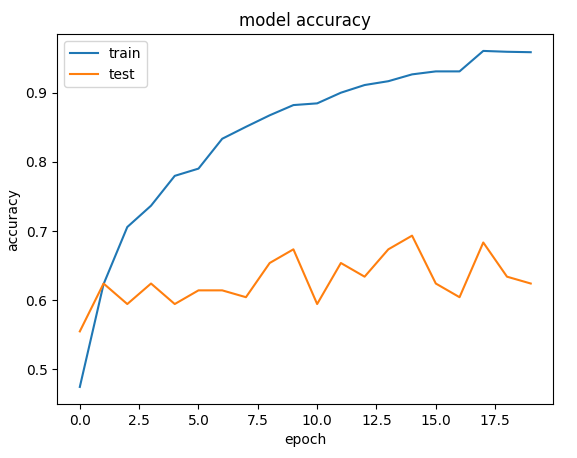


Рисунок 13 – График точности

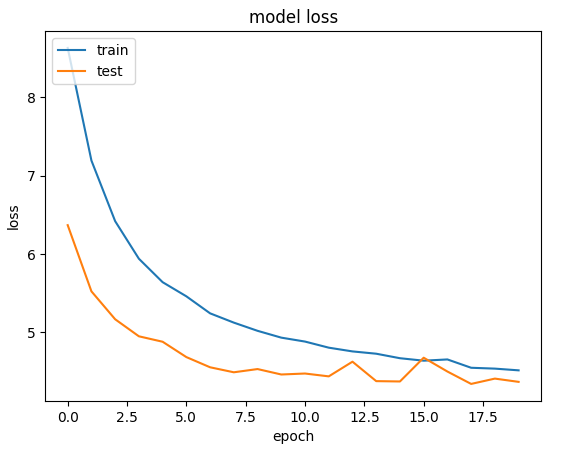


Рисунок 14 – График функции потерь

Получим значения метрики accuracy на обучающей выборке = 95,86% и тестовой выборке = 62,38%.

На рисунке 15 представлена матрица ошибок.

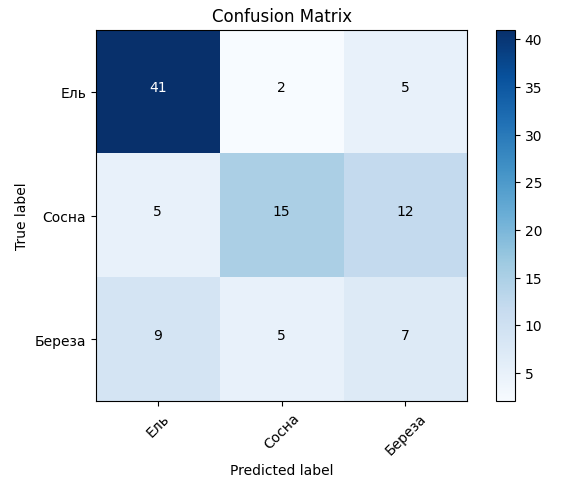


Рисунок 15 – Матрица ошибок

Для модели №5 увеличим количество эпох (epochs=40) и уменьшим скорость обучения (lr=0.0005) в 2 раза. Графики точности и функций потерь представлены на рисунках 16 и 17.

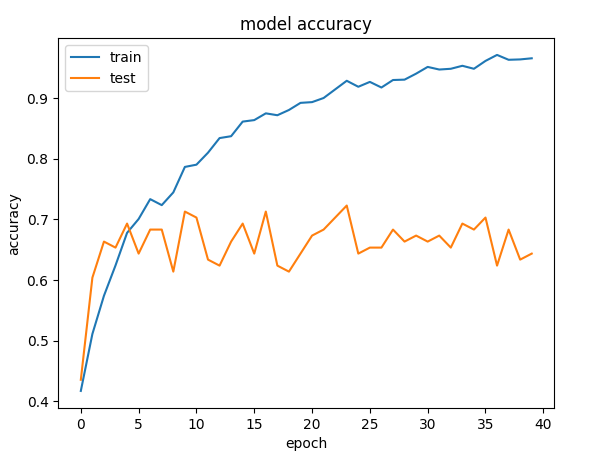


Рисунок 16 – График точности

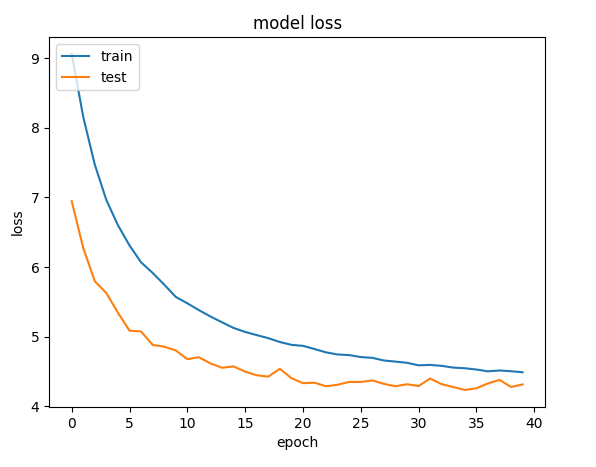


Рисунок 17 – График функции потерь

Получим значения метрики accuracy на обучающей выборке = 96,54% и тестовой выборке = 64,36%.

На рисунке 18 представлена матрица ошибок.

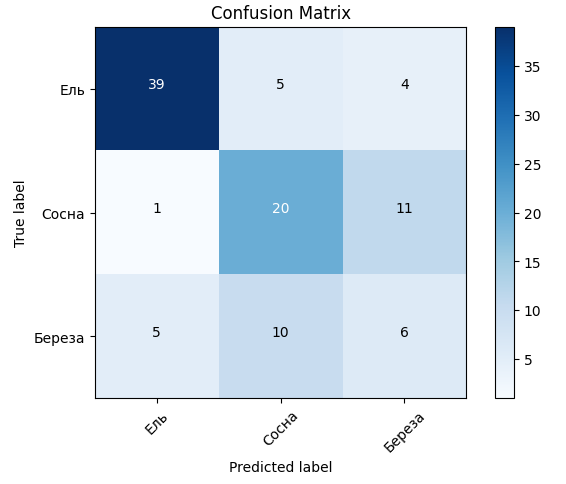


Рисунок 18 – Матрица ошибок

Дополнительное задание – улучшение за счёт изменения оптимизатора:

Для последней модели изменим оптимизатор с SGD на Adam. Все остальные параметры оставим неизменными. Графики точности и функций потерь представлены на рисунках 19 и 20.

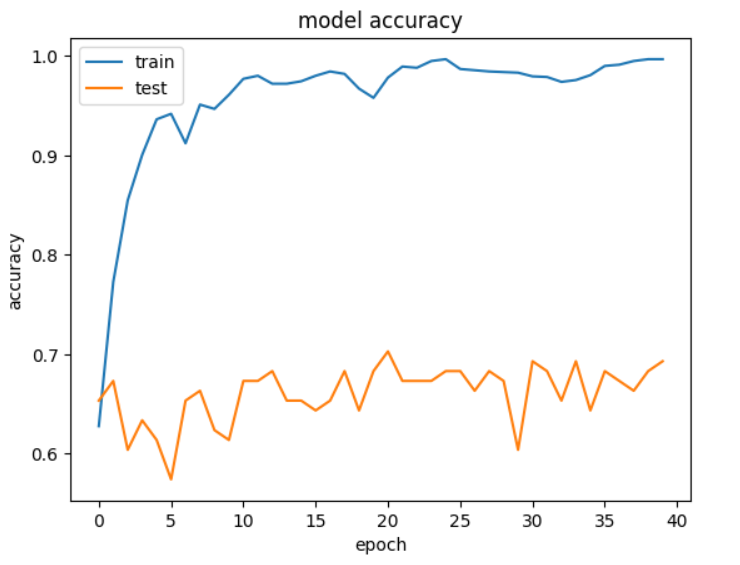


Рисунок 19 – График точности

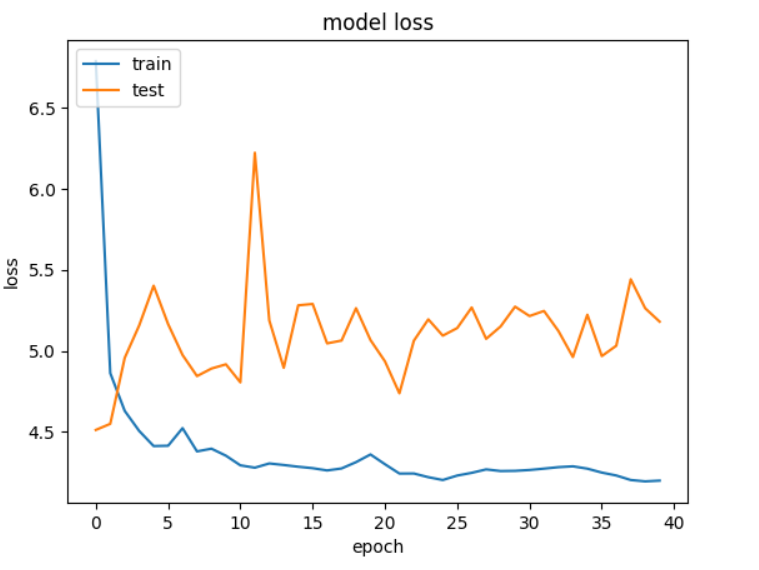


Рисунок 20 – График функции потерь

Получено значения метрики accuracy на обучающей выборке = 99,69% и тестовой выборке = 69,31%.

На рисунке 21 представлена матрица ошибок.

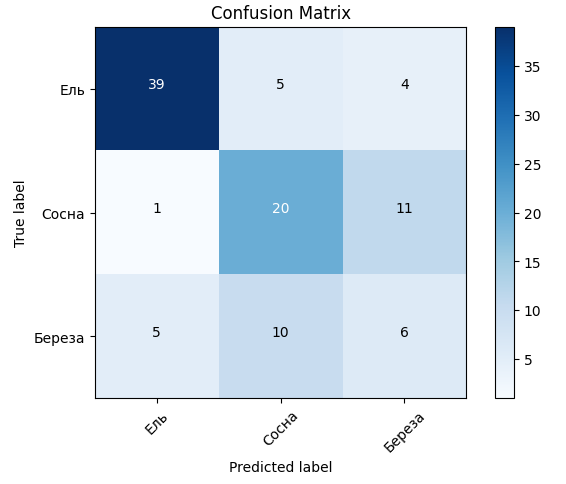


Рисунок 21 – Матрица ошибок

**Анализ результатов лабораторной работы**

Результаты моделей приведем в сводной таблице 1.

Таблица 1 – Результаты лабораторной работы

| **№** | **Изменяемые параметры** | **Точность** | **Коммент.** |
| --- | --- | --- | --- |
| 1 | batch\_size = 64,  lr = 0,001  epochs = 10 | 57,43% | Базовый вариант. Обнаружили переобучение |
| 2 | batch\_size = 64,  lr = 0,001  epochs = 19 | 59,41% | Увеличили количество эпох |
| 3 | batch\_size = 64,  lr = 0,0005  epochs = 20 | 59,41% | Уменьшили скорость обучения в 2 раза, увеличили число эпох в 2 раза. Обнаружили переобучение |
| 4 | batch\_size = 64,  lr = 0,0005  epochs = 10 | 61,39% | Уменьшаем эпохи до 10, избавляемся от переобучения |
| 5 | batch\_size = 128,  lr = 0,001  epochs = 20 | 62,38% | Увеличили размер батча в 2 раза, увеличили число эпох в 2 раза. Обнаружили переобучение |
| 6 | batch\_size = 128,  lr = 0,0005  epochs = 40 | 64,36% | Улучшение варианта |
| 7 | batch\_size = 128,  lr = 0,0005  epochs = 40  оптимизатор Adam | 69,31% | Улучшение варианта за счет использования оптимизатора Adam |

Выводы

В ходе лабораторной работы была обучена модель нейронной сети для классификации пород деревьев, используя файлы облака точек. Важно грамотно определить момент наступления переобучения, и соответственно уменьшить количество эпох до этого момента.

Уменьшение скорости обучения и увеличение числа эпох повысило качество модели, что совпадает с теорией. Увеличение размера батча улучшило качество модели. Наилучшей оказалась модель (64,36%) с увеличенным в 2 раза размером батча относительно базового варианта, уменьшенной в 4 раза скоростью обучения и соответственно увеличенным количеством эпох.

Матрица ошибок группирует TP и TN по разным классам. TP находится на главной диагонали матрицы, FN - в нижней части матрицы, а FP - в верхней части матрицы. Наилучший результат показал класс «Ель», а наихудший – класс «Береза».