МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего образования «Самарский национальный исследовательский университет имени академика С.П. Королева»

(Самарский университет)

Институт информатики и кибернетики

Кафедра информационных систем и технологий

Отчет по лабораторной работе №2

по курсу

«Интеллектуальные системы управления»

Выполнил:

студент группы

6132

Иванов И.И.

Проверил:

к.т.н., доцент Солдатова О.П.

Самара 2023

Целью данной лабораторной работы является демонстрация способности нейронной сети решать задачи классификации. Реализовать модель на языке Python в среде Jupyter Notebook с использованием библиотек Numpy, Pandas, Matplotlib, Sclearn, Keras и Tensorflow.

Используемая модель: многослойный персептрон. Вариант №5.

Порядок выполнения работы:

1. Считать файл с данными для классификации итальянских вин.
2. Написать программу, имитирующую работу многослойного персептрона с настраиваемым числом скрытых слоёв от 1 до 3 в соответствии с вариантом.
3. Провести обучение сети с использованием заданных алгоритмов обучения SGD, SGD с импульсом (моментом) Нестерова, Adagrad , Adadelta, RMSProp , Adam , Adamas, Nadam .
4. Провести тестирование сети, обученной при помощи заданных алгоритмов.
5. Продемонстрировать работу многослойного персептрона преподавателю.
6. Исследовать зависимость функции ошибки при обучении и тестировании, а также метрику качества accuracy от объёма обучающей выборки для всех алгоритмов.
7. Исследовать зависимость функции ошибки при обучении и тестировании, а также метрики качества accuracy от числа эпох обучения для заданных алгоритмов.
8. Построить графики функции ошибки при обучении и тестировании, а также метрики качества accuracy для обучения и тестирования.
9. Рассчитать метрики precision, recall, TPN, FPN, F1.
10. Построить ROC кривые и посчитать площадь AUC под кривой.
11. Сделать вывод о лучшем алгоритме обучения и лучших параметрах топологии сети и обучения.
12. Рассчитать и построить графическое изображение для Confusion matrics.
13. Составить отчёт, который должен содержать постановку задачи, текст программы, результаты исследований по пунктам 5-12 и сдать его преподавателю.

Описание сети

В данной работе для решения задачи классификации используется сеть на основе многослойного персептрона.

Многослойный персептрон состоит из множества входных узлов, которые образуют входной слой; одного или нескольких скрытых слоев вычислительных нейронов и одного выходного слоя. Входной сигнал распространяется по сети в прямом направлении от слоя к слою. Многослойные персептроны успешно применяются для решения разнообразных сложных задач. При этом обучение с учителем выполняется с помощью такого популярного алгоритма, как алгоритм обратного распространения ошибки.

Многослойный персептрон имеет три отличительных признака:

Каждый нейрон имеет нелинейную функцию активации. Данная функция должна быть гладкой (то есть всюду дифференцируемой). Самой популярной гладкой функцией активации является сигмоидальная функция.

Сеть содержит один или несколько слоев скрытых нейронов. Эти нейроны позволяют сети обучаться решению сложных задач, последовательно извлекая наиболее важные признаки из входного вектора.

Сеть обладает высокой степенью связности, реализуемой посредством синаптических соединений.

Структура многослойного персептрона с одним скрытым слоем изображена на рисунке 2.

Файл с данными для классификации итальянских вин

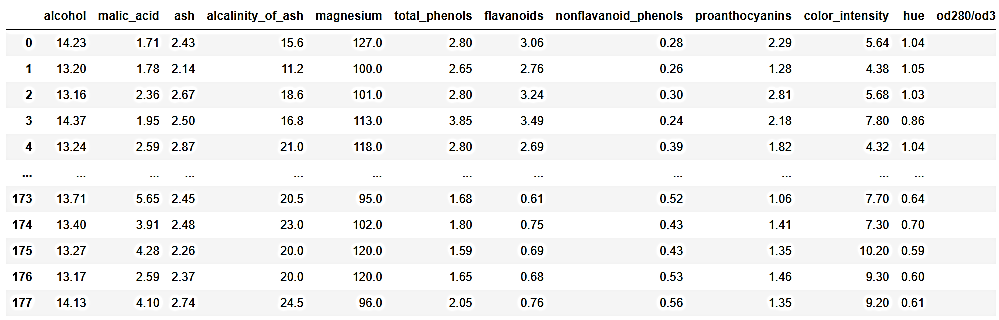
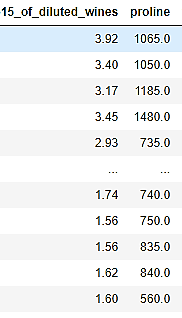
 

Рисунок 1 – Пример данных датасета.

Многослойный персептрон

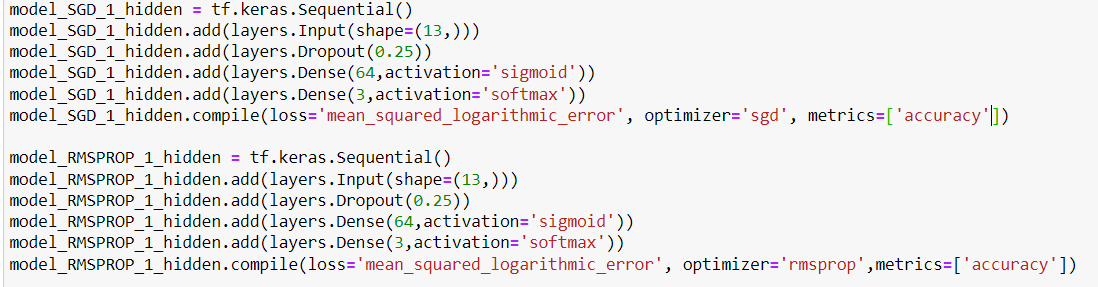


Рисунок 2 – Структура многослойного персептрона для алгоритмов SGD и RMSPROP

Структура многослойного персептрона с тремя скрытыми слоями изображена на рисунке 3

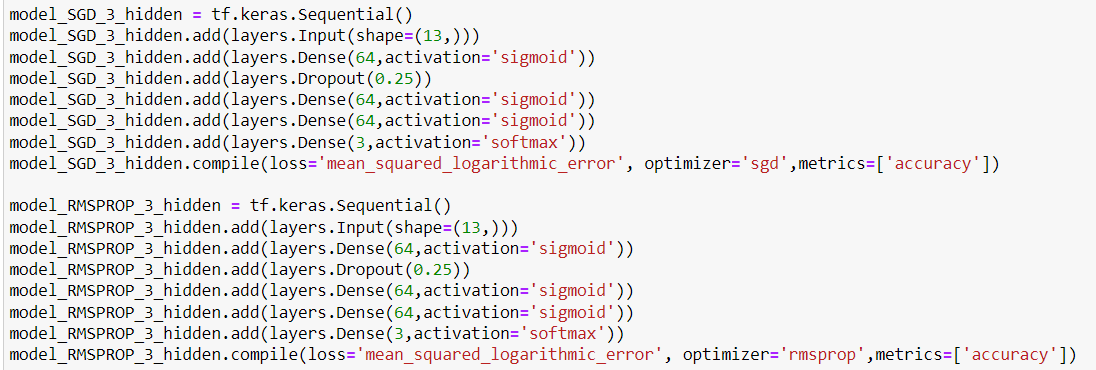


Рисунок 3 – Структура многослойного персептрона для алгоритмов SGD и RMSPROP.

В программе реализованы 2 модели многослойного персептрона.

Первая модель включает входной слой, 1 скрытый полносвязный слой. В качестве функции активации каждого слоя используется «sigmoid», для последнего слоя выбрана функция «softmax». Выбраны такие параметры: размер батча – 1, количество эпох обучения – 50, доля обучающей выборки из общей выборки = 20%. В качестве метрики погрешности прогнозирования взято accuracy, precision, recall, TPN, FPN, F1.

Вторая модель включает входной слой, 3 скрытых полносвязных слоя. В качестве функции активации каждого слоя используется «sigmoid», для последнего слоя выбрана функция «softmax». Выбраны такие параметры: размер батча – 1, количество эпох обучения – 50, доля обучающей выборки из общей выборки = 20%. В качестве метрики погрешности прогнозирования взято accuracy, precision, recall, TPN, FPN, F1.

При обучение использованы алгоритмы SGD, RMSPROP.

Описание предметной области

В данной работе использовался набор данных для классификации вин с 3 классами.

Пример данных на рисунке 3.

Исследование модели1 и модели2 многослойного персептрона

Исследовать зависимость функции погрешности при обучении и классификации, а также метрику качества accuracy от числа эпох обучения для всех алгоритмов.

На рисунках 4 – 7 изображены графики зависимости погрешности и метрики качества accuracy от числа эпох.

Метрика accuracy при тестировании работы нейросети для алгоритма SGD с 1 слоем составила около 100% и алгоритма SGD с 3 слоями около 90%, RMSPROP с 1 слоем составила 100% и алгоритма RMSPROP с 3 слоями около 99%.

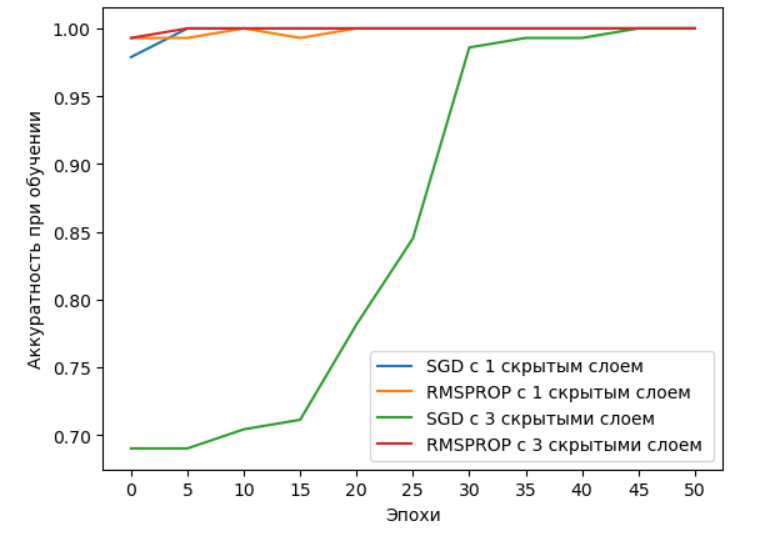


Рисунок 4 - Зависимость метрики accuracy при обучении по алгоритмам от количества эпох обучения.

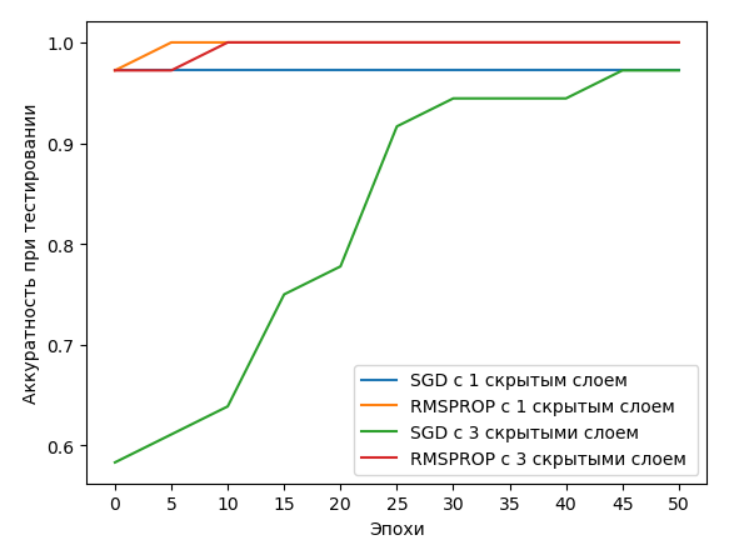
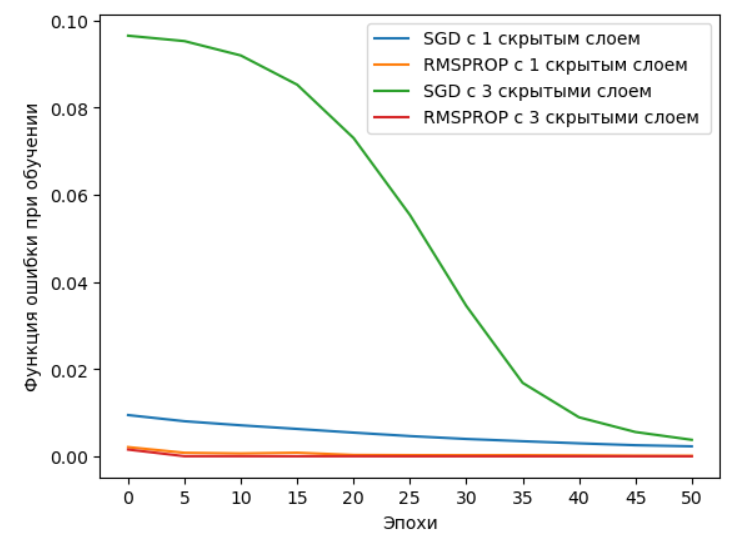
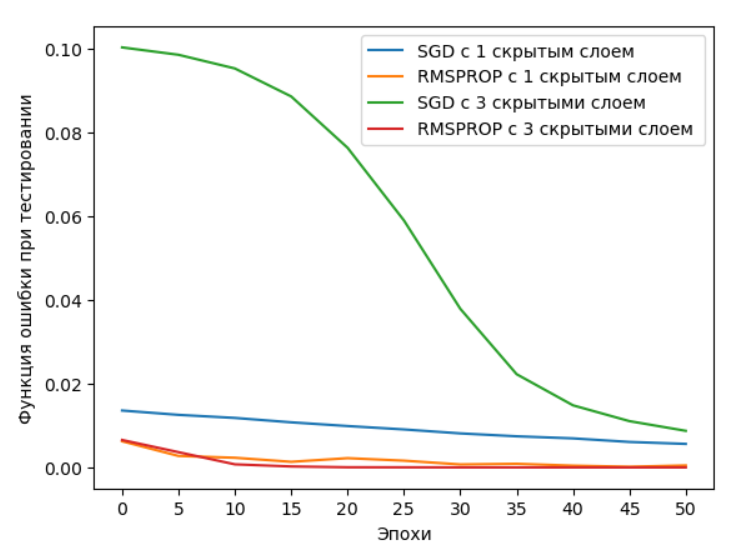


Рисунок 5 – Зависимость метрики accuracy при классификации по алгоритмам от количества эпох обучения.

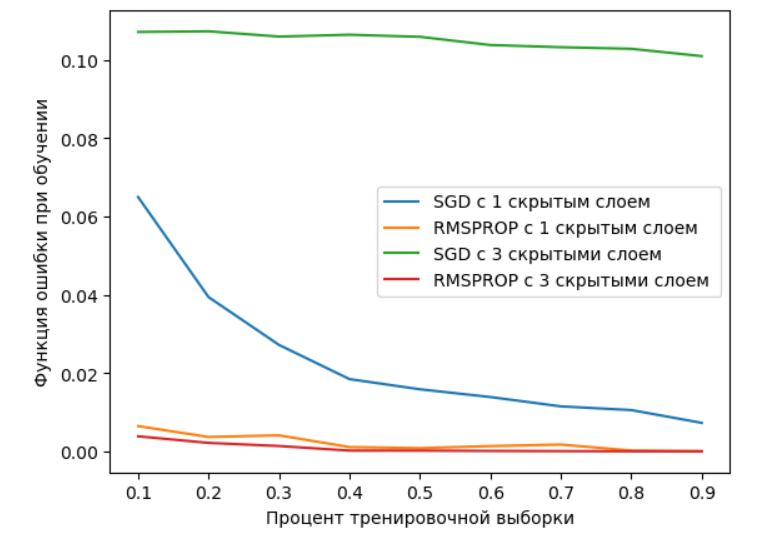
Рисунок 6 – Зависимость погрешности при обучении от количества эпох.

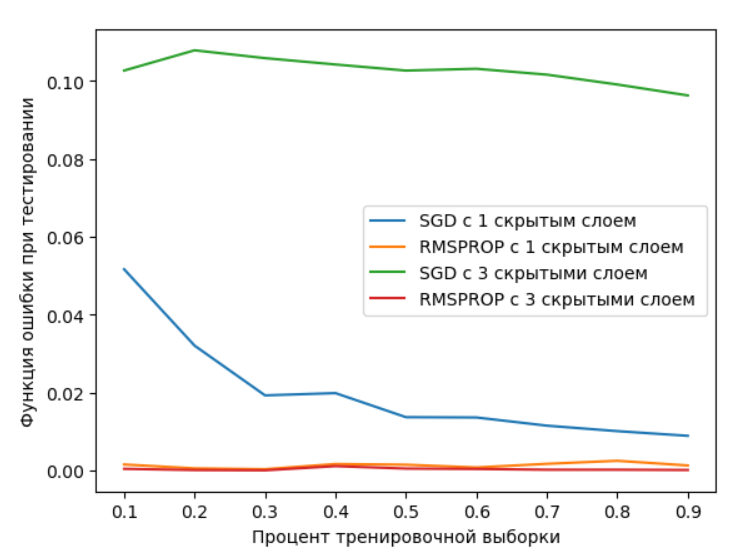
Рисунок 7 – Зависимость погрешности при классификации от количества эпох.

Исследование функции погрешности при обучении и классификации, а также метрики качества accuracy для всех исследований.

Дополнительно было проведено исследование погрешности прогнозирования от способа разделения обучающей выборки на две части: обучающую и тестирующую.

Результаты экспериментов приведены ниже. По оси X находится доля выборки, используемая для обучения, от 10% до 90%, по оси Y – погрешность прогнозирования.

Рисунок 8 – Зависимость погрешности при обучении от разделения обучающей выборки на 2 части: обучающую и тестирующую.

Рисунок 9 – Зависимость погрешности при классификации от разделения обучающей выборки на 2 части: обучающую и тестирующую.

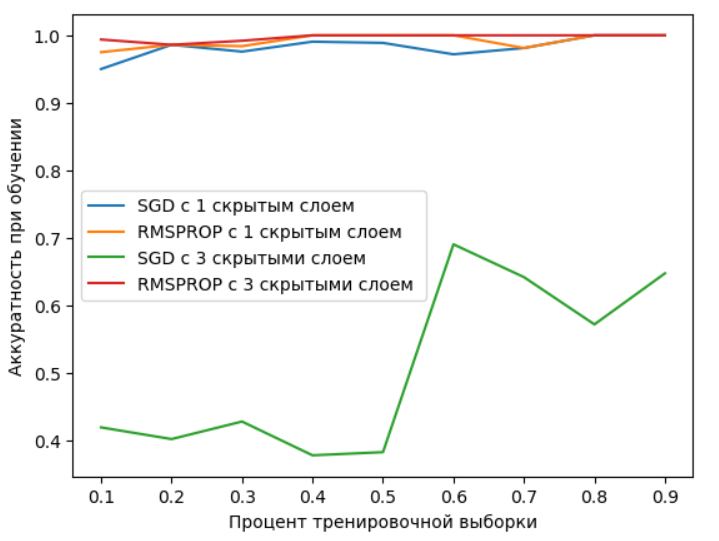


Рисунок 10 – Зависимость метрики accuracy при обучении от разделения обучающей выборки на 2 части: обучающую и тестирующую.

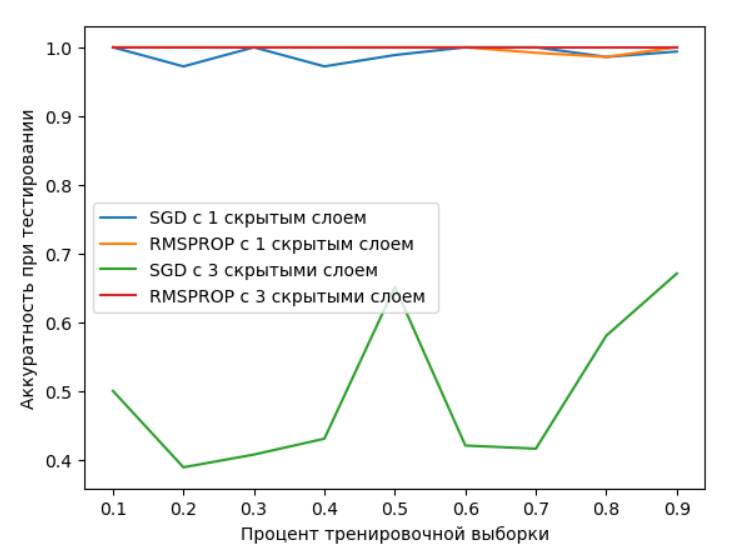


Рисунок 11 - Зависимость метрики accuracy при классификации от разделения обучающей выборки на 2 части: обучающую и тестирующую.

ROC-кривые и дополнительные метрики

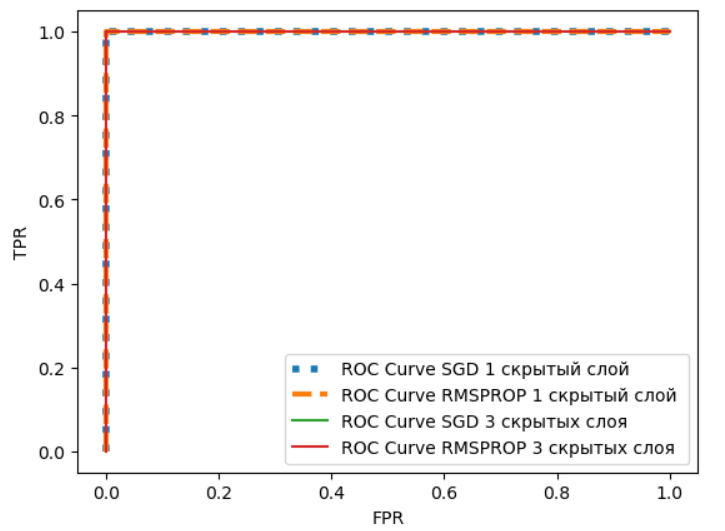


Рисунок 12 – Зависимость TPR от FPR

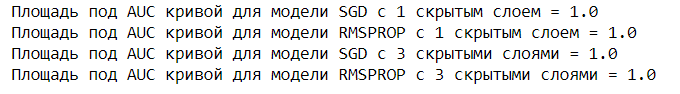
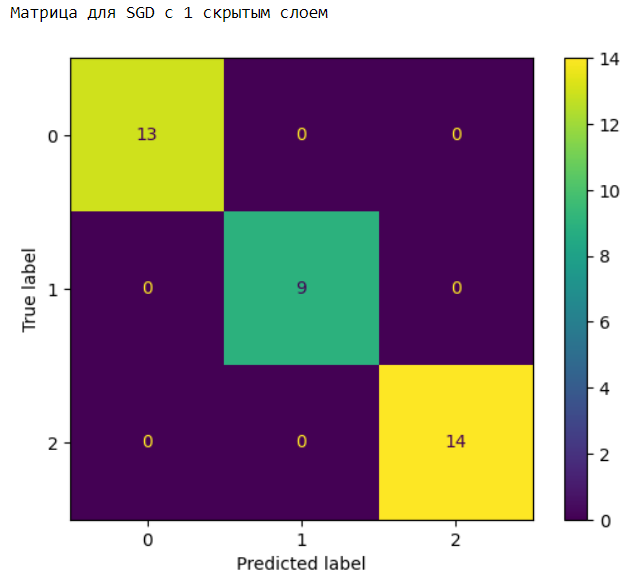
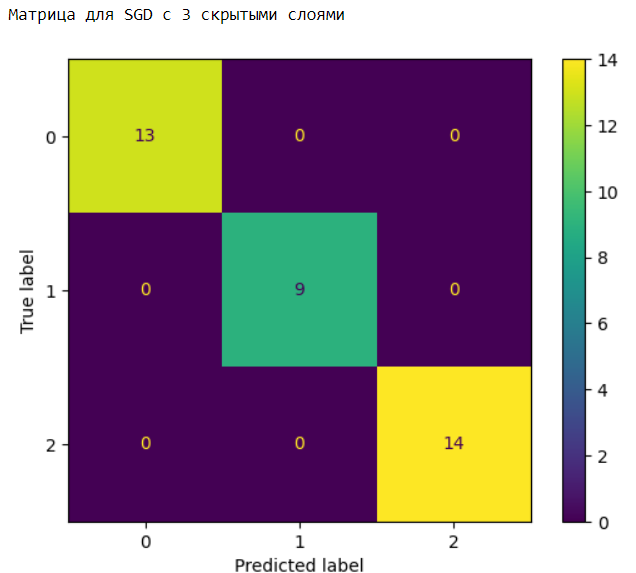
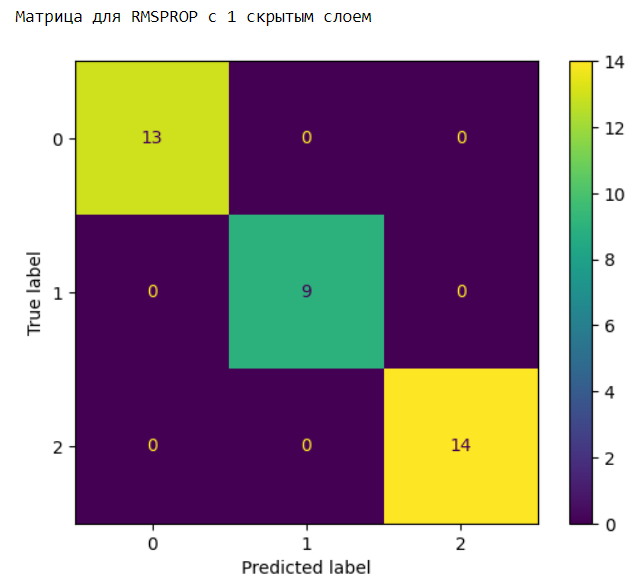
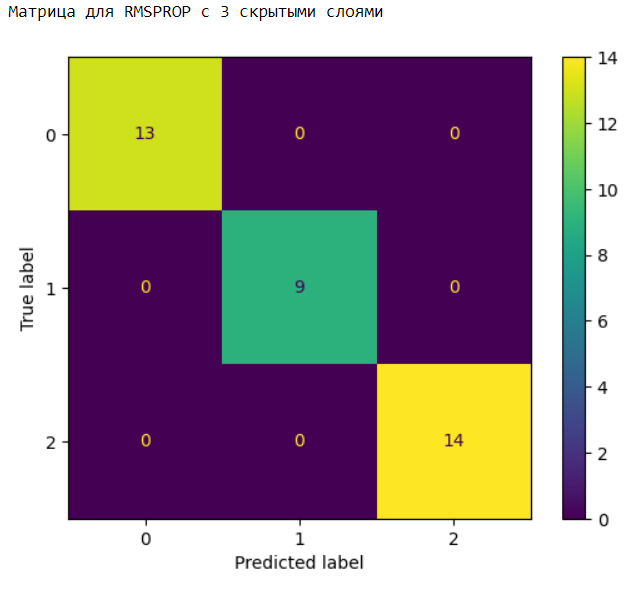


Рисунок 13 – Площади под ROC-кривыми

Рисунок 14 – Confusion matrix для SGD с 1 слоем

Рисунок 15 – Confusion matrix для SGD с 3 слоями

Рисунок 16 – Confusion matrix для RMSPROP с 1 слоем

Рисунок 17 – Confusion matrix для RMSPROP с 3 слоями

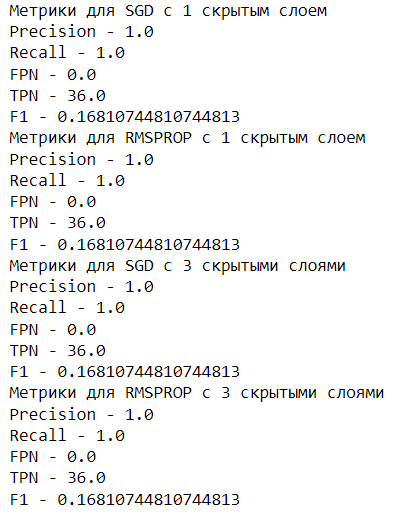


Рисунок 18 – Метрики Precision, Recall, FPN, TPN, F1.

Зависимость функции ошибки и метрики accuracy от размера Dropout

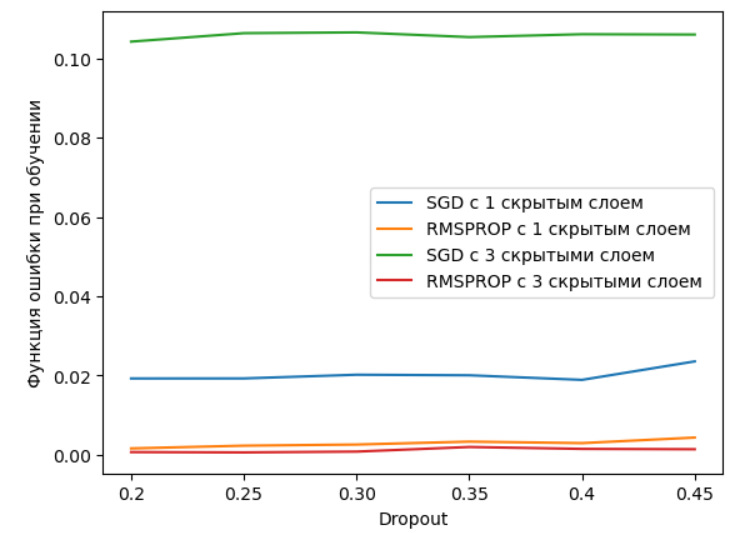


Рисунок 19 – Функция ошибки от Dropout при обучении

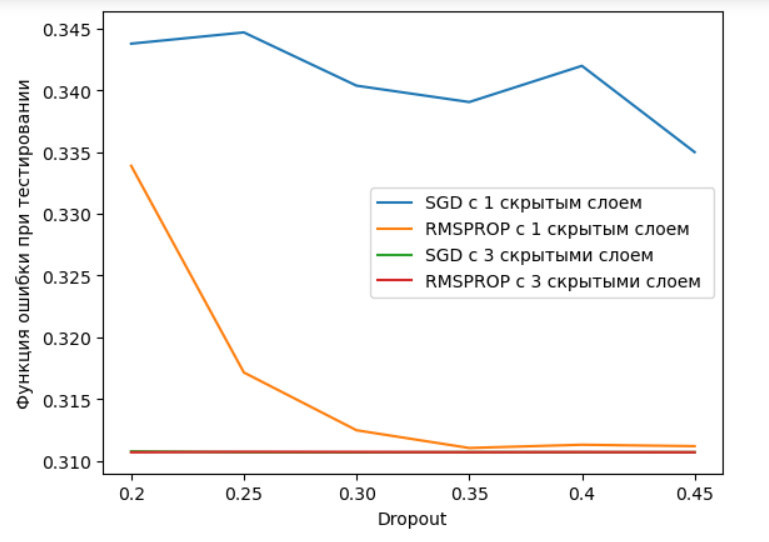
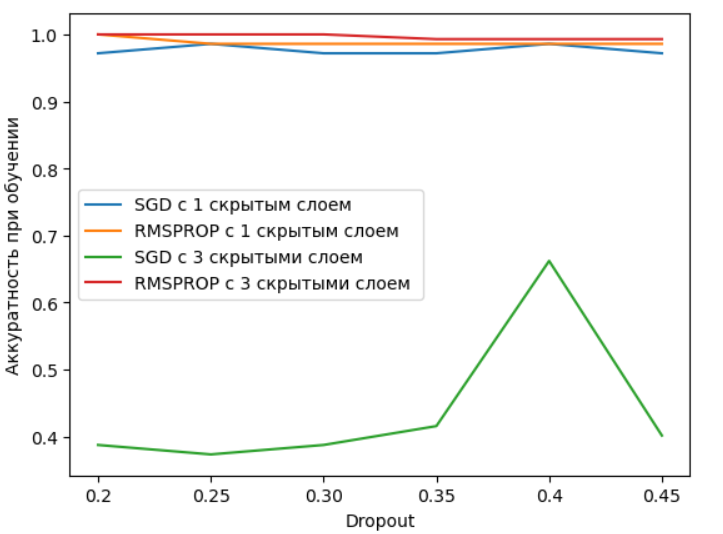


Рисунок 20 – Функция ошибки от Dropout при классификации

Рисунок 21 – Метрика accuracy от Dropout при обучении

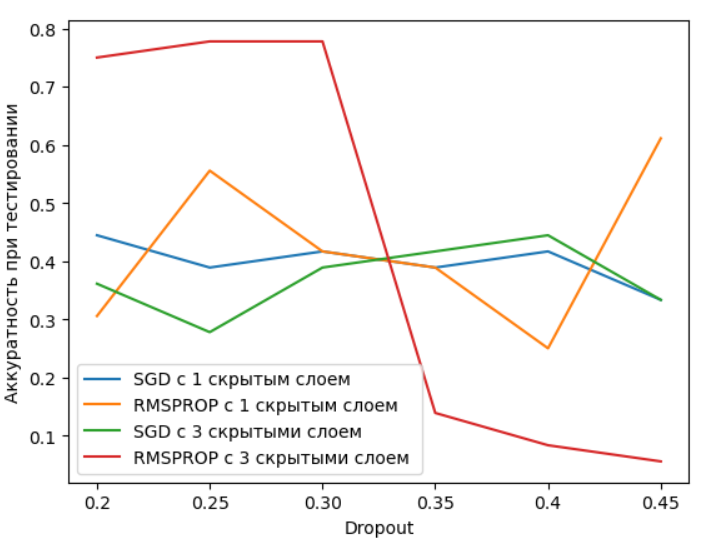


Рисунок 22 – Метрика accuracy от Dropout при классификации

Сравнительный анализ лучших параметров двух моделей многослойного персептрона

Выводы

По результатам исследований можно сделать вывод, что многослойный персептрон хорошо подходит для задач бинарной классификации и классификации с малым количеством классов. Для правильного выбора параметров сети важно проводить исследования и выбирать их оптимальные соотношения. Самым эффективным алгоритмом обучения на выбранной структуре сети является RMSPROP, оптимальным разделением является разделение: 80% на обучение, 20% на проверку. Результаты исследований приведены в таблице 1.

Оптимальными значениями Dropout является 0,35.

Таблица 1 – Результаты исследований двух моделей многослойного персептрона

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Алгоритм обучения 1 - SGD | | | | | |
|  | Оптимальный объём обучающей выборки | Оптимальное число эпох обучения | Максимальное значение accuracy при обучении | Максимальное значение accuracy при тестировании | Минимальное значение функции ошибки при обучении | Минимальное значение функции ошибки при тестировании |
| Модель 1  1 слой | 20% | 5 | 1 | 0.97 | 0.01 | 0.01 |
| Модель 2  3 слоя | 55% | 45 | 1 | 0.97 | 0.11 | 0.012 |
|  | Алгоритм обучения 2 - RMSPROP | | | | | |
|  | Оптимальный объём обучающей выборки | Оптимальное число эпох обучения | Максимальное значение accuracy при обучении | Максимальное значение accuracy при тестировании | Минимальное значение функции ошибки при обучении | Минимальное значение функции ошибки при тестировании |
| Модель 1  1 слой | 20% | 10 | 1 | 1 | 0.0001 | 0.0001 |
| Модель 2  3 слоя | 20% | 10 | 1 | 1 | 0.00001 | 0.00001 |