**Лабораторная работа 4**

**Классификатор**

Описание датасета

Данные содержат 830 записей и 6 атрибутов. Датасет содержит информацию для скрининга рака молочной железы.

Attributes

BI-RADS assessment: 1 to 5

Age: patient's age in years

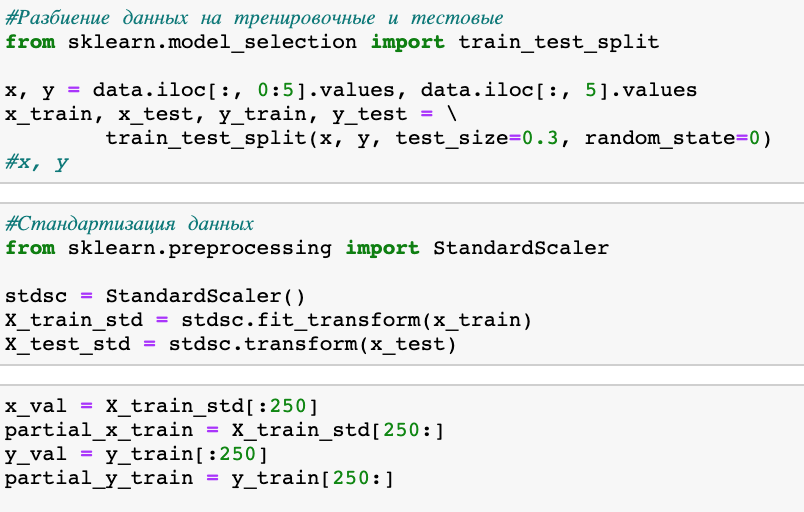
Shape: mass shape: round=1 oval=2 lobular=3 irregular=4

Margin: mass margin: circumscribed=1 microlobulated=2 obscured=3 ill-defined=4 spiculated=5

Density: mass density high=1 iso=2 low=3 fat-containing=4

Severity: benign=0 or malignant=1



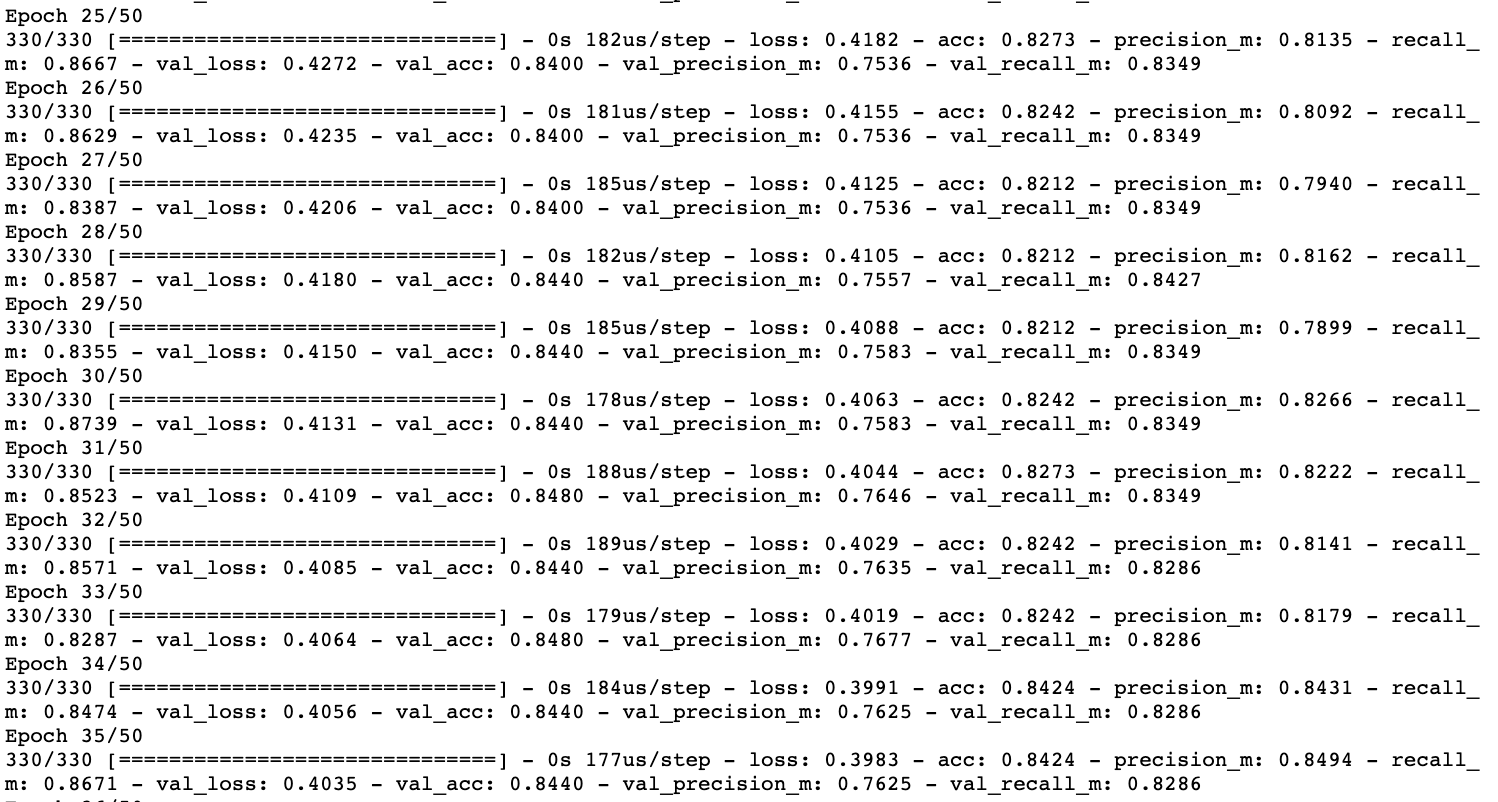


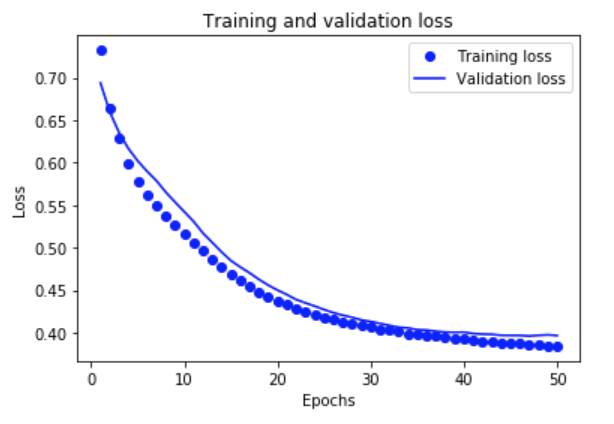
Нужно выбрать функцию потерь и оптимизатор. Так как перед нами стоит задача бинарной классификации и результатом работы сети является вероятность (наша сеть заканчивается одномодульным слоем с сигмоидной функцией активации), предпочтительнее использовать функцию потерь binary\_crossentropy. Но это не единственный приемлемый выбор: можно также задействовать, например, mean\_squared\_error. Однако перекрестная энтропия обычно дает более качественные результаты, когда результатами работы моделей являются вероятности. Перекрестная энтропия (crossentropy) — это термин из области теории информации, обозначающий меру расстояния между распределениями вероятностей, или в данном случае — между фактическими данными и предсказаниями.

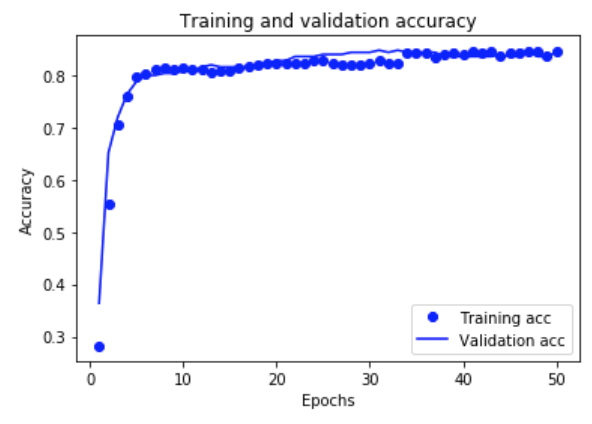
Поэтому проверим и то и другое:

**Модель классификатора binary\_crossentropy**

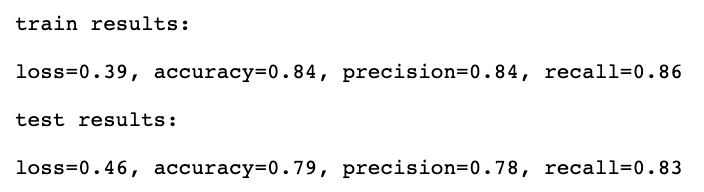
На этом шаге мы настраиваем модель оптимизатором rmsprop и функцией потерь binary\_crossentropy. Мы также задали мониторинг точности и плотности во время обучения.



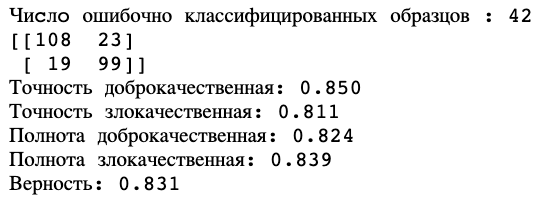




Итог

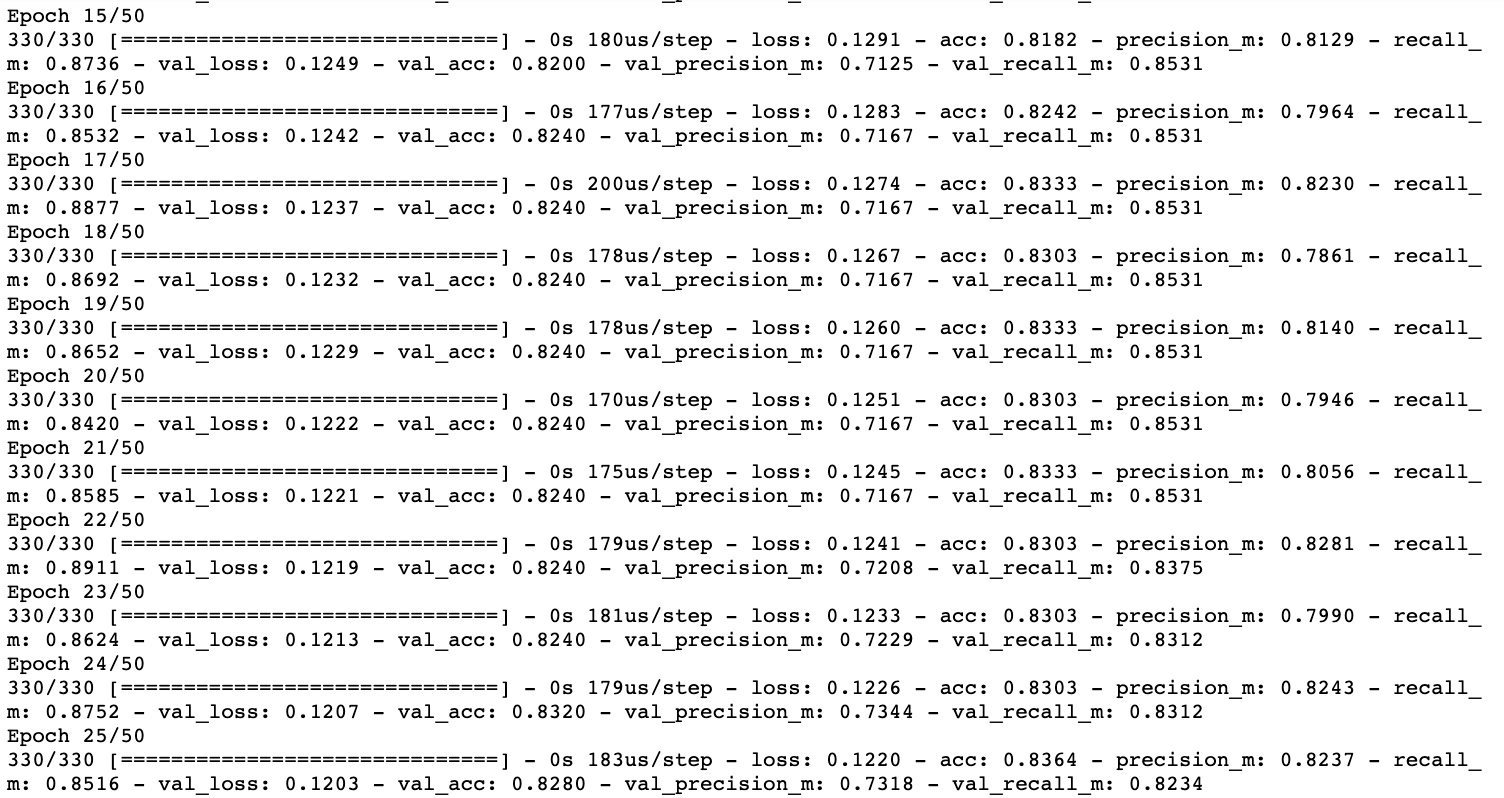


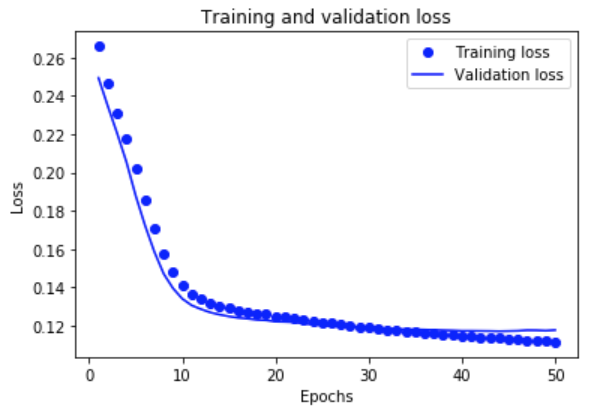
для сравнения дерево решений на таком же датасете из лаб. 3:

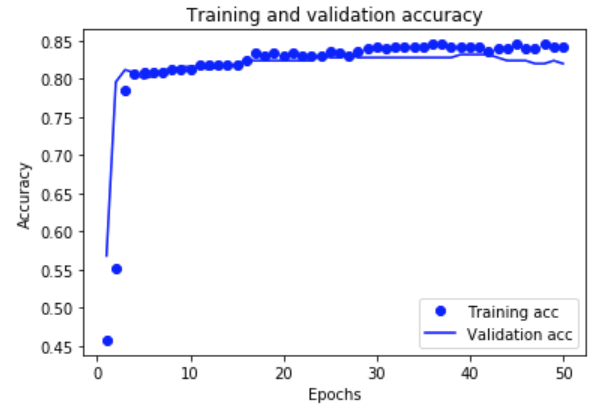


**Модель классификатора MSE**

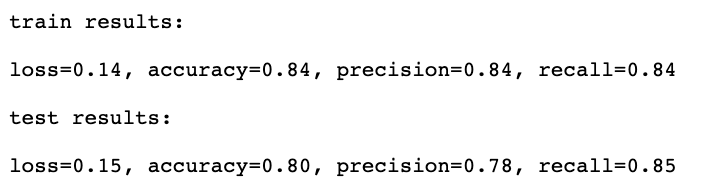
На этом шаге мы настраиваем модель оптимизатором rmsprop и функцией потерь MSE. Мы также задали мониторинг точности и плотности во время обучения.



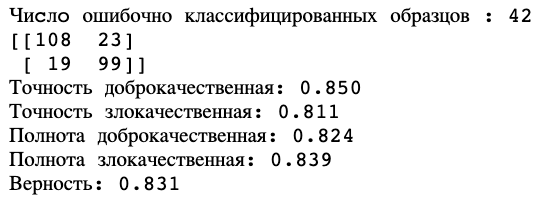




Итог

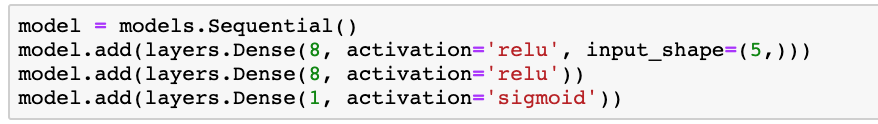


для сравнения дерево решений на таком же датасете из лаб. 3:

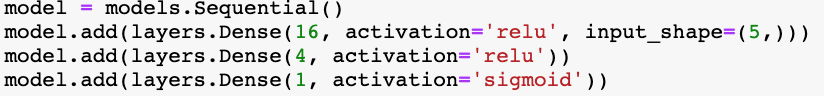


**Вывод для классификатора**

При сравнении с деревом решений из лаб. 3, который показал себя лучше всего модель с функцией потерь binary\_crossentropy показала себя хуже, она использовала два слоя по 8 нейронов.



модель же с функцией потерь MSE показала себя лучше, она использовала два слоя по 16 и 4 нейронов.



**Регрессор**

Описание датасета

Эти данные являются характеристиками красного вина. Набор данных загружается из репозитория машинного обучения UCI. В нём 1599 записей и 12 атрибутов.

Attribute Information:



Для отбора признаков которые будем использовать в модели посмотрим что наиболее коррелирует с атрибутом «качество»

Другим распространенным типом задач машинного обучения является регрессия, которая заключается в предсказании не дискретной метки, а значения на непрерывной числовой прямой

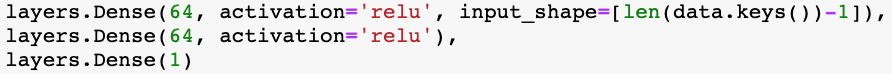
**Модель 1**

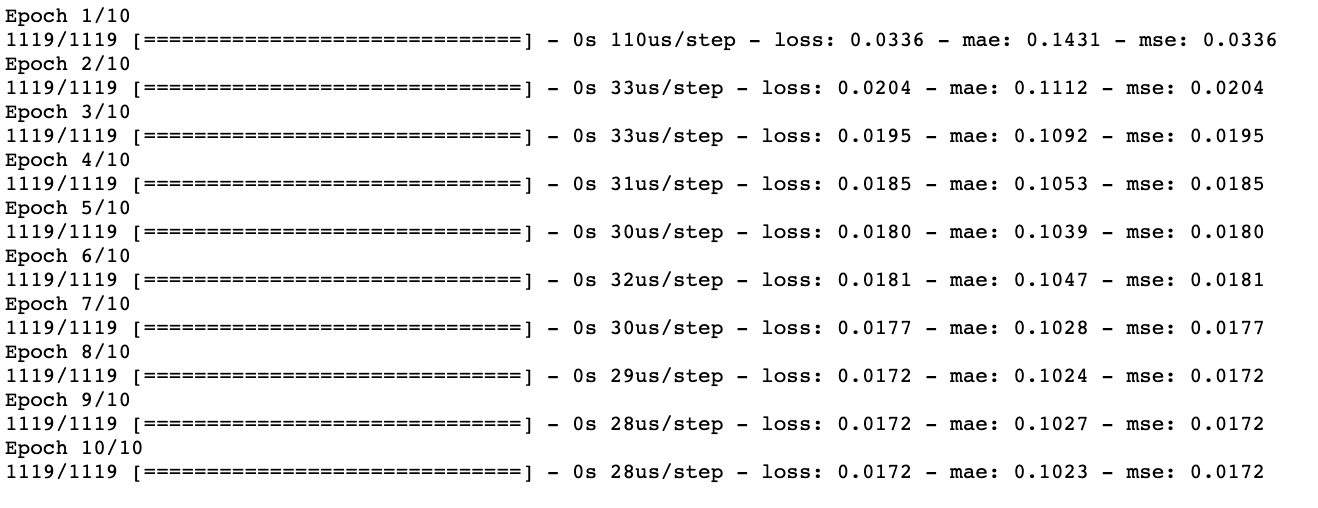
Из-за небольшого количества образцов мы будем использовать очень маленькую сеть с двумя четырехмерными промежуточными слоями. Вообще говоря, чем меньше обучающих данных, тем скорее наступит переобучение, а использование маленькой сети — один из способов борьбы с ним.

Сеть заканчивается одномерным слоем, не имеющим функции активации (это линейный слой). Это типичная конфигурация для скалярной регрессии (целью которой является предсказание одного значения на непрерывной числовой прямой). Применение функции активации могло бы ограничить диапазон выходных значений: например, если в последнем слое применить функцию активации sigmoid, сеть обучилась бы предсказывать только значения из диапазона между 0 и 1. В данном случае, с линейным последним слоем, сеть способна предсказывать значения из любого диапазона.

mse — mean squared error (среднеквадратичная ошибка), вычисляющей квадрат разности между предсказанными и целевыми значениями.

mae — mean absolute error (средняя абсолютная ошибка) - это абсолютное значение разности между предсказанными и целевыми значениями.



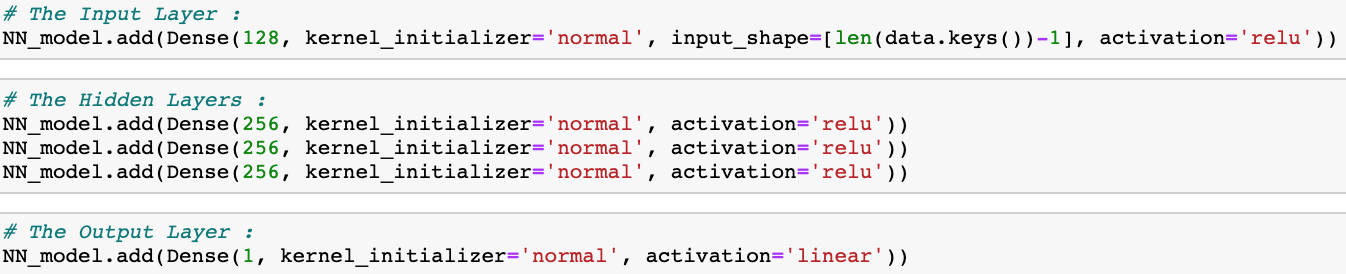


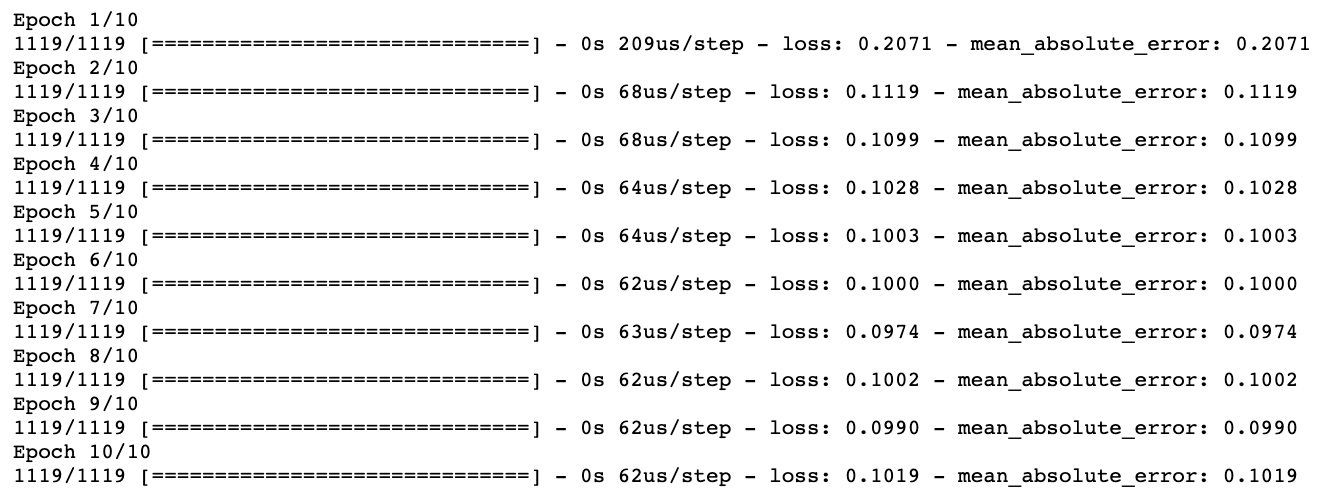
Итог





**Модель 2**









**Вывод для регрессии**

Обе модели показали себя примерно одинаково для задачи регрессии, но для модели 2 требовалось гораздо больше вычислительной мощности, поэтому лучше использовать модель 1.

Регрессия выполняется с применением иных функций потерь, нежели классификация. Для регрессии часто используется функция потерь, вычисляющая среднеквадратичную ошибку (Mean Squared Error, MSE).

При небольшом объеме обучающих данных предпочтительнее использовать маленькие сети с небольшим количеством промежуточных слоев (обычно с одним или двумя), чтобы избежать серьезного переобучения.

В процессе обучения нейронных сетей в некоторый момент наступает эффект переобучения, из-за чего падает качество результатов оценки сети на данных, которые она прежде не видела.

В том случае, когда данные делятся на большое число категорий, у вас может возникнуть узкое место для информации, если вы слишком сильно ограничите размерность промежуточных слоев.