

МИНОБРНАУКИ РОССИИ  
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра физики открытых систем

**Решение задачи классификации крыс предрасположенных к  
эпилепсии**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студента 4 курса 4041 группы  
направления 09.03.02 «Информационные системы и технологии»  
Института Физики  
Саможён Сергея Сергеевича

Научный руководитель  
доцент кафедры физики  
открытых систем, к.ф. –м.н.

\_\_\_\_\_

дата, подпись

М.О. Журавлев

Заведующий кафедрой  
физики открытых систем  
д.ф. –м.н., профессор

\_\_\_\_\_

дата, подпись

А.А. Короновский

Саратов 2023 год

## Введение

Тема данного исследования — обучение нейросетей для бинарной классификации по данным электрокортикографии крыс. Электрокортикография (ЭКоГ) - это инвазивный метод, используемый для регистрации активности мозга непосредственно с поверхности коры головного мозга. Цель данной работы - провести обучение двух нейросетей с разной архитектурой на данных ЭКоГ для классификации эпилепсии и сравнить их эффективность. Для достижения цели были поставлены следующие задачи:

1. Изучить различные архитектуры нейросетей и применение их в анализе ЭКоГ.
2. Провести предварительную обработку данных ЭКоГ и разделить каждую запись на окна с фиксированной длиной.
3. Обучить на подготовленных данных две нейросети, одна имеет архитектуру перцептрон, другая сверточную архитектуру.
4. Проанализировать полученные результаты и выявить эффективную архитектуру.

Содержание работы:

### 1 ВВЕДЕНИЕ

### 2 Машинное обучение

#### 2.1 Искусственный интеллект

#### 2.2 Нейросети

### 3 Архитектура нейросетей

#### 3.1 Перцептрон

#### 3.2 Свёрточные нейросети

#### 3.3 Рекуррентные нейросети

#### 3.4 Слои нейронных сетей

##### 3.4.1 Слой Dense

##### 3.4.2 Слой Conv1D

##### 3.4.3 Слой BatchNormalization

##### 3.4.4 Слой Flatten

##### 3.4.5 Слой Dropout

#### 3.5 Функции активации

##### 3.5.1 Sigmoid

##### 3.5.2 ReLU

### 3.5.3 Tanh

## 4 Практическая реализация

### 4.1 Подготовка данных

### 4.2 Обучение перцептрона

### 4.3 Подготовка данных для сверточной нейронной сети

### 4.4 Обучение сверточной нейронной сети

## Основное содержание работы

**Описание методологии** В данном исследовании было обучено две нейросети имеющие следующую архитектуру: первая нейросеть является пятислойным перцептроном, вторая нейросеть является сверточной нейросетью с двумя сверточными слоями и тремя полносвязными слоями. Первым этапом работы была предварительная подготовка данных ЭКоГ. Данные были переведены из формата edf в формат csv. Были рассчитаны дельты с разной дистанцией для каждой записи. Затем данные были разделены на окна разной длины, в зависимости от архитектуры нейросети. Для обработки данных использовались библиотеки Pandas[1] и Numpy[2] на языке Python. Для визуализации полученных результатов были использованы графики, построенные в библиотеке Matplotlib[3].

**Подготовка данных** Данные, предоставленные научным руководителем, для этой задачи выглядят так:

Сигналы:

Time	1	2	3
0.000000000	-0.218562	-0.131125	0.022312
0.002500000	-0.207063	-0.126875	0.025875
0.005000000	-0.199313	-0.092438	0.007188
0.007500000	-0.170500	-0.086125	0.001375
0.010000000	-0.157875	-0.078625	0.021063
0.012500000	-0.132250	-0.014063	0.008188
0.015000000	-0.070187	0.050000	0.021125
0.017500000	-0.056688	0.043188	0.033500
0.020000000	-0.018562	0.038875	0.078750
0.022500000	0.031625	0.052562	0.113375
0.025000000	0.060625	0.078063	0.091875
0.027500000	0.034813	0.069875	0.057813
0.030000000	0.006125	0.066625	0.045187

Сегменты:

NN	SWD start, s	SWD end, s	SWD, s
1	104	108	3.27
2	157	169	11.91
3	233	237	3.41
4	1256	1262	5.95
5	2138	2144	6.30
6	2211	2218	7.06
7	3876	3878	2.05
8	4815	4825	10.29
9	4840	4848	7.61
10	4870	4877	7.56
11	4902	4914	12.68
12	4952	4961	8.97
13	5003	5007	4.83

Теперь необходимо соединить эти данные. Исходя из них, можно увидеть, что частота сигнальных данных 400 раз в секунду, а эпилепсия замечена с точностью до миллисекунд. Следовательно, программа подготавливающая данные, должна не только сопоставлять участки, но и уменьшать

точность до миллисекунд. Для такой задачи, была разработана программа и ее алгоритм будет описан далее.

1. Считываются размеченные сегменты с файла.
2. Считываются сигналы с файла и сразу же уменьшается их точность в 40 раз.
3. Сигналы и сегменты объединяются образуя конечную таблицу.
4. Шаги 1-3 повторяются для каждого файла и после все данные объединяются в один файл.

Итак, данные приведены к единой точности и формату. Теперь необходимо преобразовать их так, чтобы сделать их понятными для нейросети.

Сначала будет обучен перцептрон на данных с 1 животного. Для этого нужно следовать следующему алгоритму: Импортировать необходимые библиотеки. Импортировать созданный датасет. Его визуализация выглядит так:

Index	time	FrL	FrR	OcR	ep
0	0.00	-0.218562	-0.131125	0.022312	0
1	0.01	-0.157875	-0.078625	0.021063	0
2	0.02	-0.018562	0.038875	0.078750	0
3	0.03	0.006125	0.066625	0.045187	0
4	0.04	-0.046000	-0.001875	0.043563	0

А размерность этих данных таковая:

$$|(838880, 5)|$$

После применение модуля к данным, можно увидеть что все значения колеблются от 0 до 2:

FrL	0.0
FrR	0.0
OcR	0.0
dtype:	float64

FrL	2.0
FrR	2.0
OcR	2.0
dtype:	float64

Результат применения модуля:

Index	FrL	FrR	OcR
0	0.218562	0.131125	0.022312
1	0.157875	0.078625	0.021063
2	0.018562	0.038875	0.078750
3	0.006125	0.066625	0.045187
4	0.046000	0.001875	0.043563

Необходимо создать больше данных, таким образом нейросеть будет лучше обучаться. Для этого можно добавить дельту между соседними ячейками. Визуализация полученного:

Index	FrL	FrR	OcR	FrL <sub>d1</sub>	FrR <sub>d1</sub>	OcR <sub>d1</sub>
3	0.006125	0.066625	0.045187	-0.012437	0.027750	-0.033563
4	0.046000	0.001875	0.043563	0.039875	-0.064750	-0.001624
5	0.045687	0.033625	0.018375	-0.000313	0.031750	-0.025188
6	0.045687	0.093625	0.059875	0.000000	0.060000	0.041500
7	0.061875	0.063813	0.052187	0.016188	-0.029812	-0.007688

Данные подготовлены, теперь следует создать итоговый набор данных для обучения. Для этого данные преобразовываются в окна размером 30 ячеек. После данные стоит разделить на тестовые и тренировочные. Получившиеся размеры данных составляют:

Тренировочных:

$$\left| \begin{array}{l} (196142, 360) \\ (196142, 1) \end{array} \right|$$

Тестовых:

$$\left| \begin{array}{l} (84362, 360) \\ (84362, 1) \end{array} \right|$$

**Обучение перцептрона** Теперь можно обучать нейронную сеть. Нейросеть будет состоять из 5 полносвязных слоев.

Обучение будет состоять из 50 эпох. Результаты тренировки перцептрона:

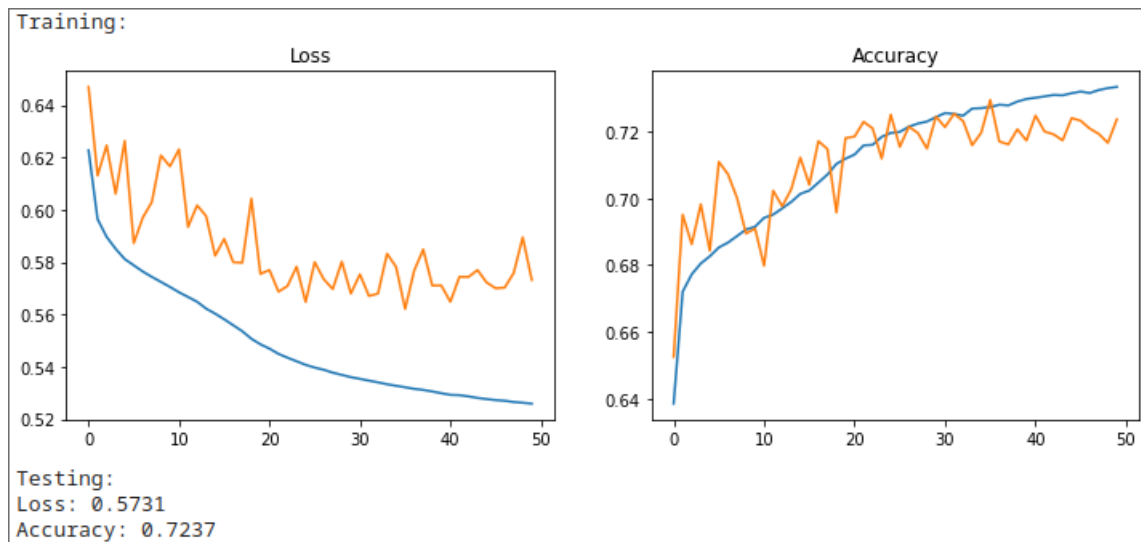


Рисунок 1 – Результаты обучения

Точность нейросети составила 72%.

Теперь можно построить матрицу ошибок, чтобы лучше оценить эффективность нейронной сети.

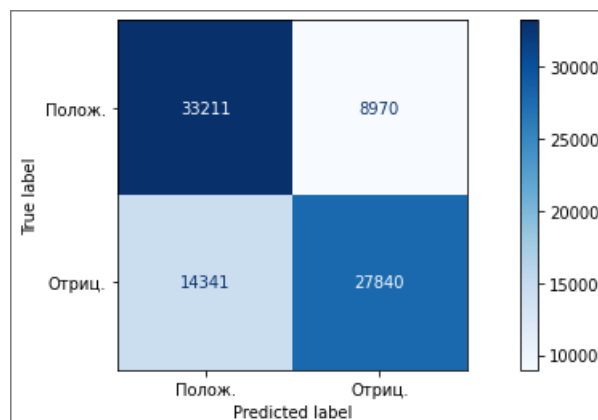


Рисунок 2 – Матрица ошибок

**Подготовка данных для сверточной нейронной сети** В отличие от перцептрона, для сверточной нейронной сети нужно подготовить больше данных, таким образом будет объединены записи ЭКоГ 10 животных.

Таким образом, объем данных таков:

$$\left| (8164160, 5) \right|$$

В целом, подготовка данных следует той же концепции что и для перцептрона, за единственным отличием, что вместо 3 дельт для каждого сигнала будет создано 5 дельт.

После добавления дельт, размерность данных представляет собой:

$$\begin{vmatrix} (8164155, 18) \\ (8164155,) \end{vmatrix}$$

Теперь данные также как и в случае с перцептронам объединяются в окна. Далее данные делятся на тестовые и тренировочные. Получившиеся размеры данных:

Тренировочные:

$$\begin{vmatrix} (135726, 1152) \\ (135726, 1) \end{vmatrix}$$

Тестовые:

$$\begin{vmatrix} (40226, 1152) \\ (40226, 1) \end{vmatrix}$$

**Обучение сверточной нейронной сети** Теперь все готово для обучения. Для начала необходимо определить модель. Ее архитектура представляет собой 2 сверточных одномерных слоя и 3 полносвязных слоя.

Обучение будет состоять из 4 эпох. Результаты обучения:

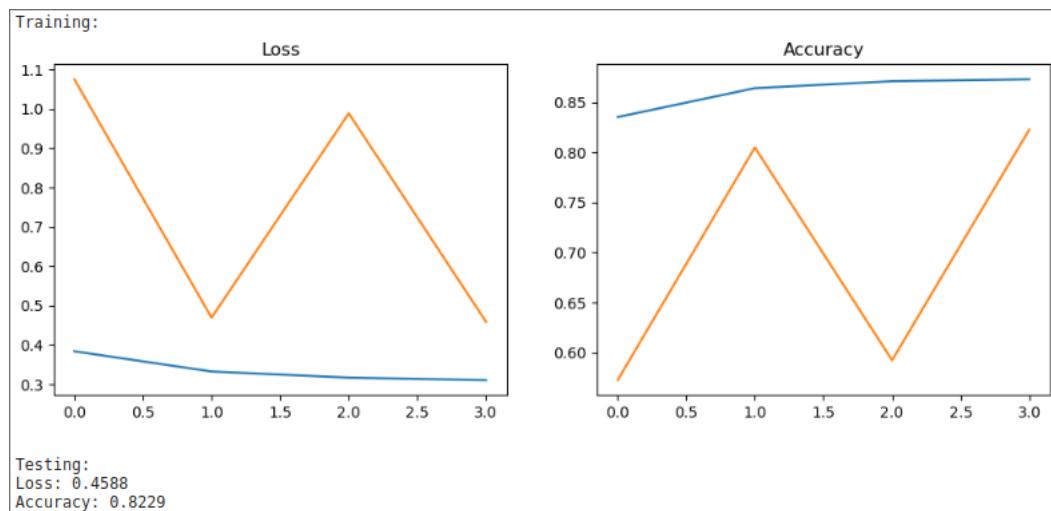


Рисунок 3 – Результаты обучения

Также получена матрица ошибок:



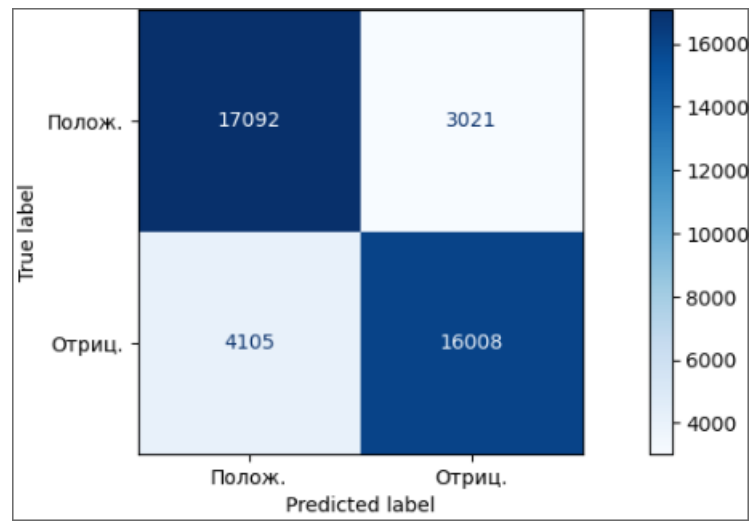


Рисунок 4 – Матрица ошибок

## Заключение

В ходе данной работы была затронута история появления нейросетей, а также их становление как части современной жизни. Рассмотрены архитектуры нейросетей, различные слои и функции активации. Решена задача классификации двумя разными нейронными сетями. Лучший показатель точности показала нейронная сеть с сверточной архитектурой: 82%. Результаты этого исследования могут быть полезны для дальнейшего изучения эпилепсии и других патологий с помощью электрокортикографии.

## СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 *У.МакКинни*,. pandas: фундаментальная библиотека python для анализа данных и статистики / У.МакКинни // *Python для производительных и научных вычислений*. — 2011.
- 2 *У.МакКинни*,. Python для анализа данных: Обработка данных с Pandas, NumPy, и IPython / У.МакКинни. — O'Reilly, 2012.
- 3 *С.Тоси*,. Matplotlib для Python разработчиков / С.Тоси. — Packt, 2009.