МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

им. Н.Э. Баумана

Факультет «Информатика и системы управления»

Кафедра «Систем обработки информации и управления»

ОТЧЕТ

**Лабораторная работа №\_\_3\_\_**

по дисциплине«Разработка нейросетевых систем»

Тема: «Регуляризация и аугментация»

ИСПОЛНИТЕЛЬ: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

ФИО

группа ИУ5-24М \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

подпись

""\_\_\_\_\_\_\_2024 г.

ПРЕПОДАВАТЕЛЬ: \_\_\_Канев А.И.\_\_\_\_\_\_

ФИО

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

подпись

"\_\_"\_\_\_\_\_\_\_\_\_202\_ г.

Москва - 2024

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

**Задание**

1. По заданию выбрать свои классы и обучить сверточную нейронную сеть c использованием техник дропаут и аугментации данных. Сравнить три варианта дропаута и три варианта аугментации данных.
2. Проанализировать результаты обучения в Tensorboard.
3. Проанализируйте результаты обучения вашей модели. Как изменилась точность на обучающей и тестовой выборке по сравнению со сверточной моделью?
4. Попробуйте применить регуляризацию и аугментацию к модели из 2-ой лабораторной. Возникает ли переобучение вашей модели?
5. Измените вашу модель - сравните три варианта параметра дропаута по заданию. Сравните результаты обучения этих 3-ех вариантов
6. Измените 3 варианта штрафа за сложность модели. Проанализируйте результаты обучения.
7. Измените гиперпараметры обучения для повышения точности модели: количество эпох, размер батча, скорость обучения
8. Сравните три варианта аугментации данных: с маленьким коэффициентом, средним и большим. Проанализируйте результаты обучения.
9. Укажите, какие действия помогли повысить точность вашей модели и объясните почему.

**Часть 1. Проанализируйте результаты обучения вашей модели. Как изменилась точность на обучающей и тестовой выборке по сравнению со сверточной моделью?**

Точность модели получилась 0.8067, что ниже свёрточной, которая равна 0.84, за счёт того, что в данная модель недообучена. Результат обучения модели представлен на рис. 1.

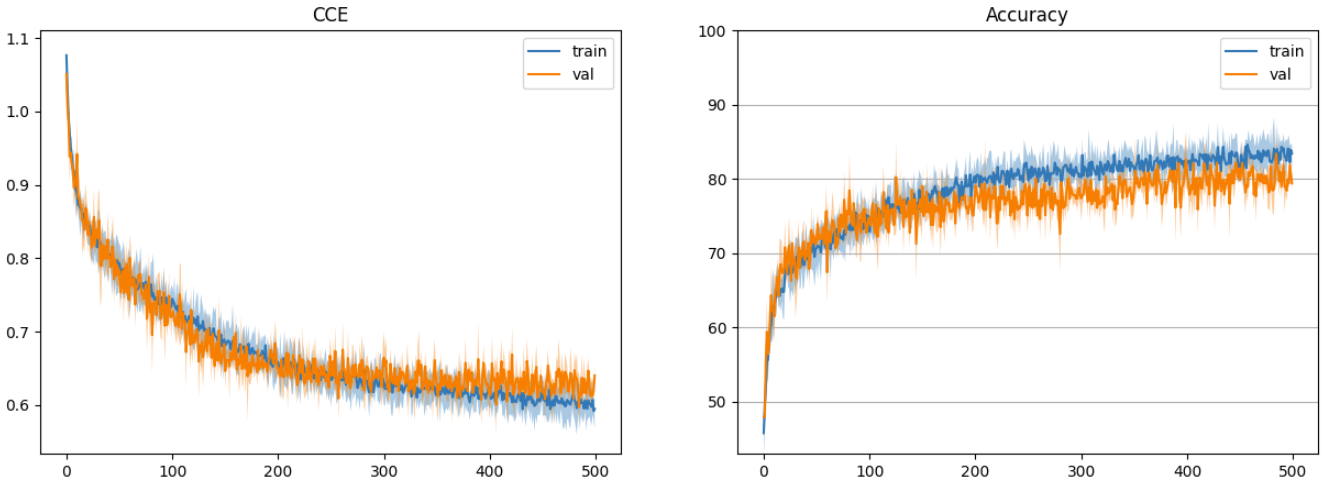


Рисунок 1 – Результат обучения модели

**Часть 2. Попробуйте применить регуляризацию и аугментацию к модели из 2-ой лабораторной. Возникает ли переобучение вашей модели?**

После применения регуляризации и аугментации уменьшилась и составляет 0.8233. Переобучение не возникает, т.к. точность на тестовой выборке и на тренировочной не имеют существенную разницу, однако, возможно точность упала из-за недообучения. Результат представлен на рис. 2.

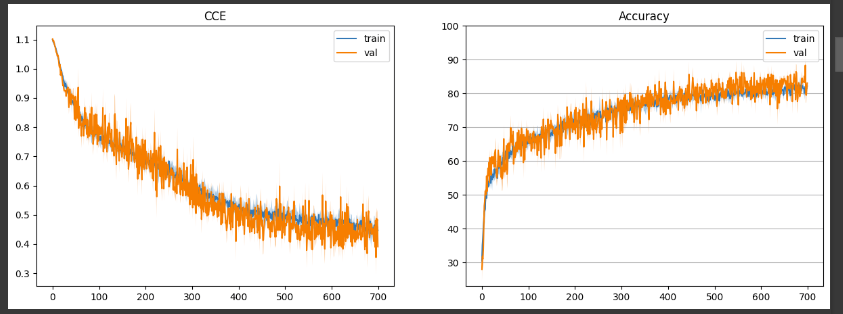


Рисунок 2 – результаты обучения модели из лаб.2 с применение регуляризации и аугментации

**Часть 3. Измените вашу модель - сравните три варианта параметра дропаута по заданию. Сравните результаты обучения этих 3-ех вариантов**

Для второго дропаута изменим на значение 0.4; представлен на рис. 3 и точность составляет 0,80.

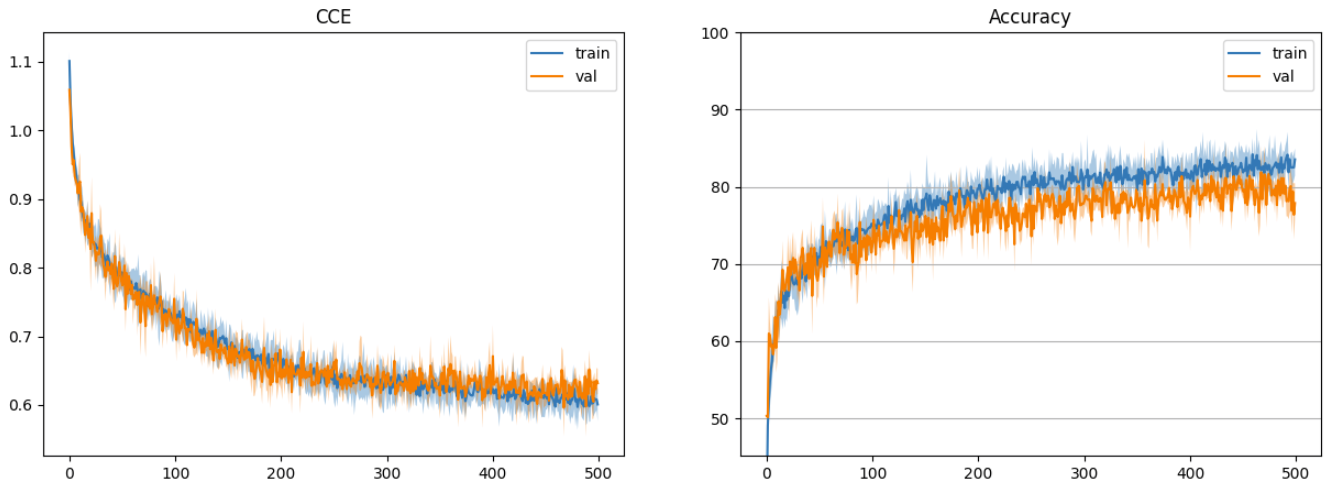


Рисунок 3 – Обучение модели, при изменении второго дропаута на 0.4

Точность немного уменьшилась, т.к. такое большое значение дропаута может привести к потере слишком большого количества информации в процессе обучения.

Для первого дропаута изменим на значение 0.1; результат представлен на рис. 4 и точность составляет 0,8033.

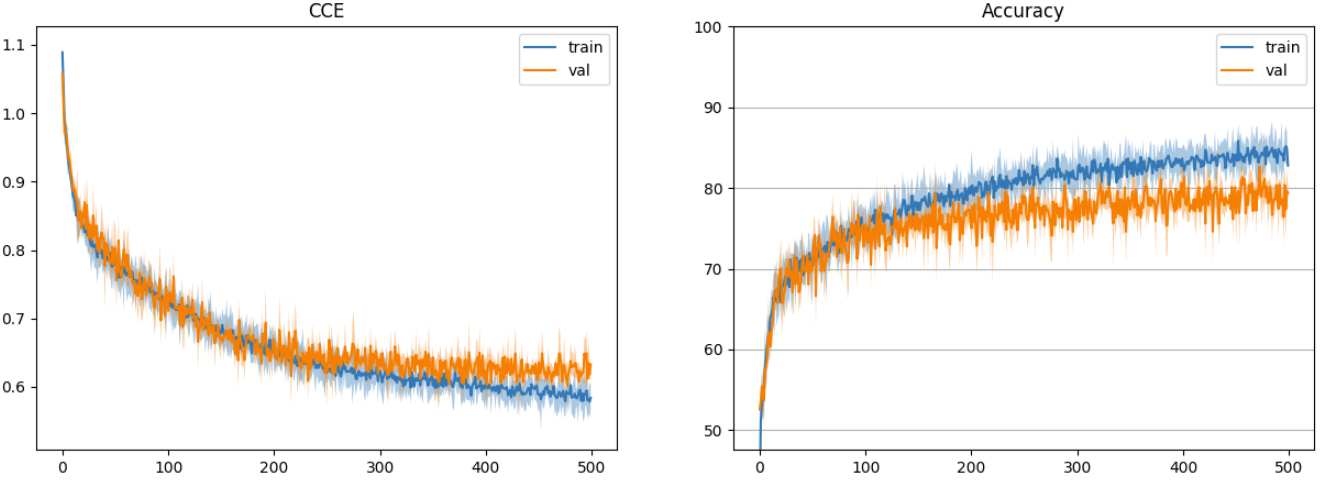


Рисунок 4 - Обучение модели, при изменении первого дропаута на 0.1

Точность незначительно уменьшилась, т.к. возникло переобучение.

Для первого дропаута изменим на значение 0.1, а для второго на 0.4; результат представлен на рис. 5 и точность составляет 0.8067.

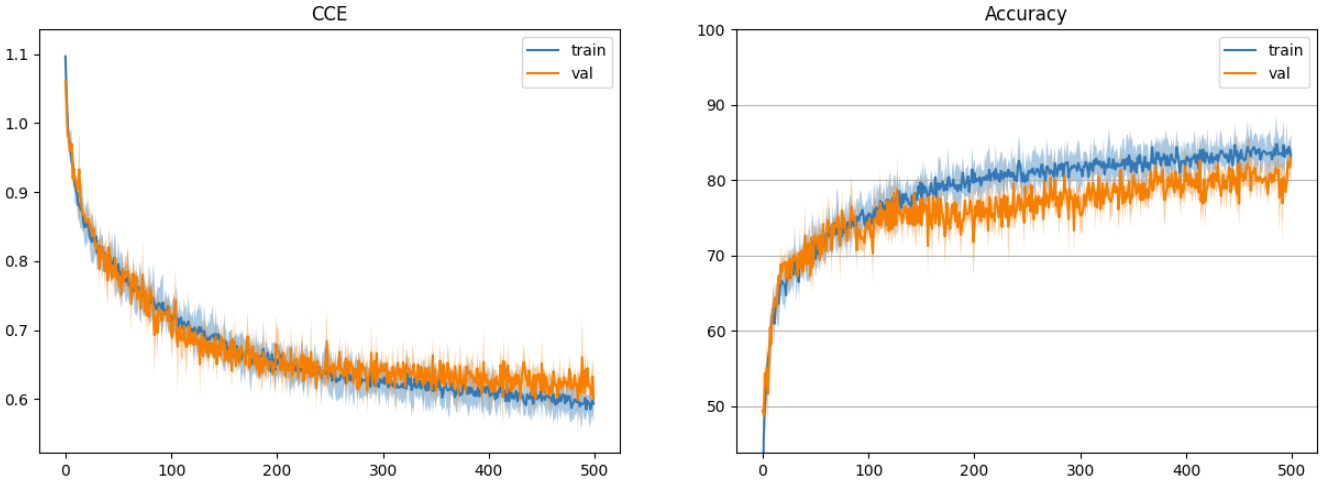


Рисунок 5 - Обучение модели, при изменении первого дропаута на 0.1 и второго на 0.4

Точность не изменилась, однако по графику видно, что в среднем стало хуже, т.к. выбрали два самый крайних случая, которые оба делают модель менее точной.

**Часть 4. Измените 3 варианта штрафа за сложность модели. Проанализируйте результаты обучения.**

Возьмём 3 значения штрафа: 1e-3, 1e-1, 1e-7.

Для значения 1e-3 результат представлен на рис. 6 и точность составляет 0.79.

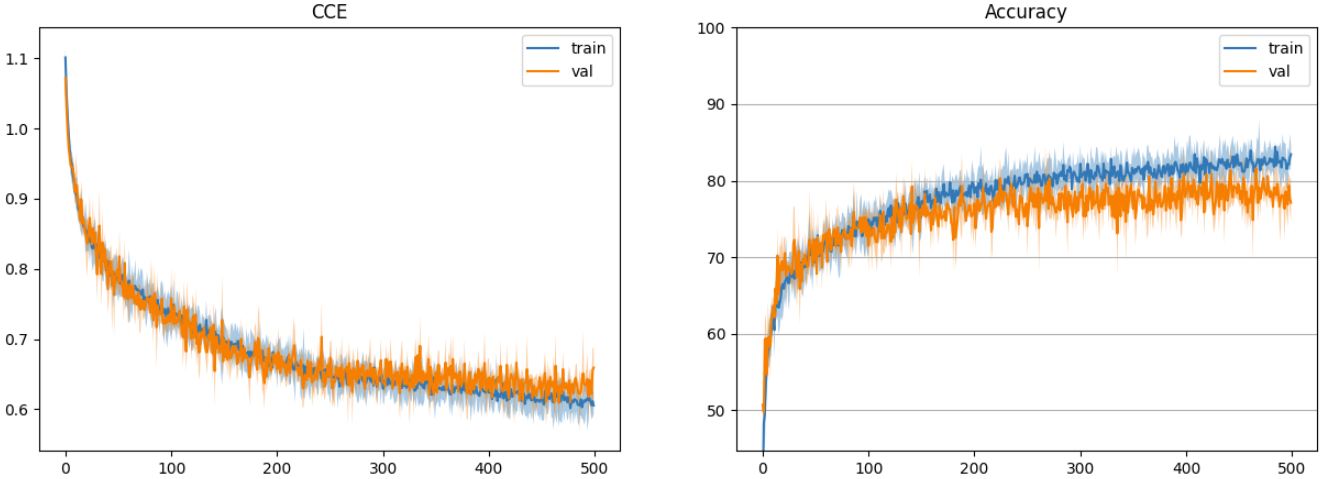


Рисунок 6 - Обучение модели, при штрафе 1е-3

Точность уменьшилась, т.к. возникло переобучение.

Для значения 1e-1 результат представлен на рис. 7 и точность составляет 0.65.

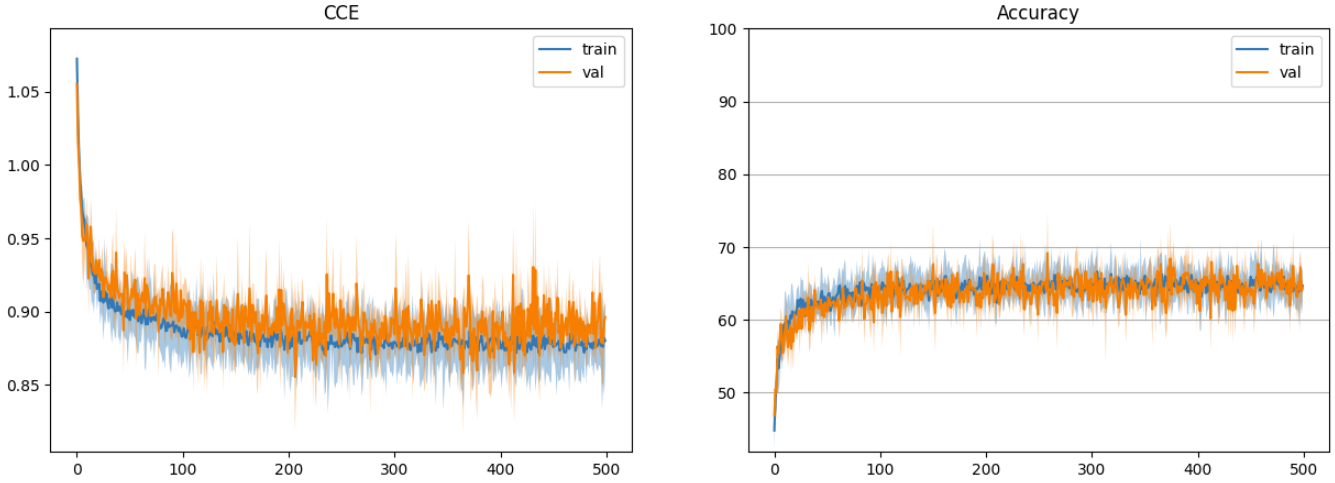


Рисунок 7 - Обучение модели, при штрафе 1е-1

Точность уменьшилась, т.к. возникло недобучение.

Для значения 1e-7 результат представлен на рис. 8 и точность составляет 0.7967.

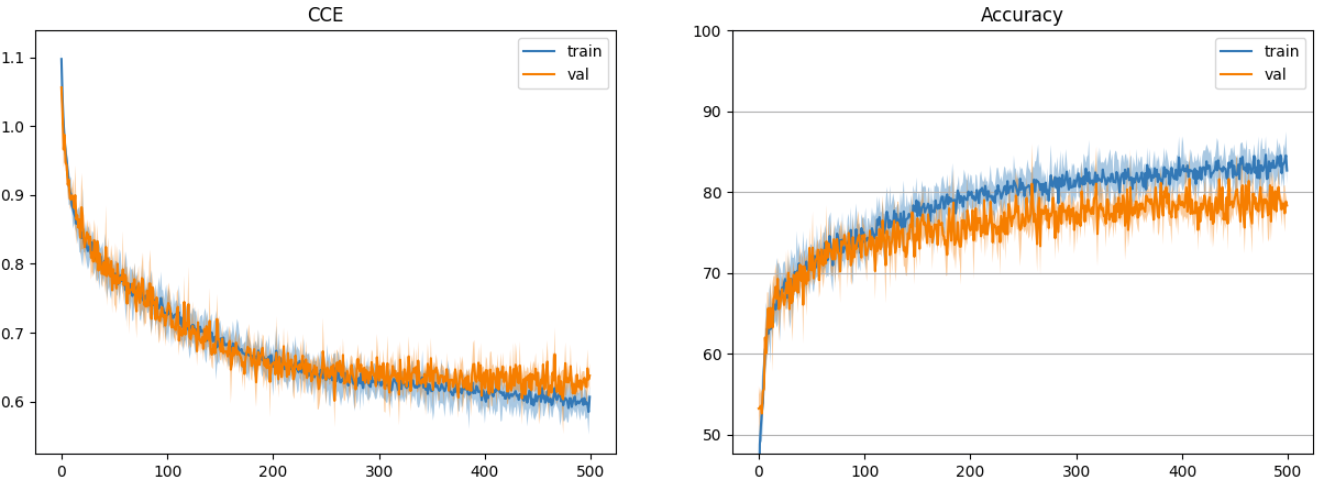


Рисунок 8- Обучение модели, при штрафе 1е-7

Точность уменьшилась, т.к. из-за уменьшения штрафа уменьшается обобщающая способность модели.

**Часть 5. Измените гиперпараметры обучения для повышения точности модели: количество эпох, размер батча, скорость обучения**

После изменения гиперпараметров был выявлен наиболее лучший вариант с точностью 0.86, что представлено на рис. 9.

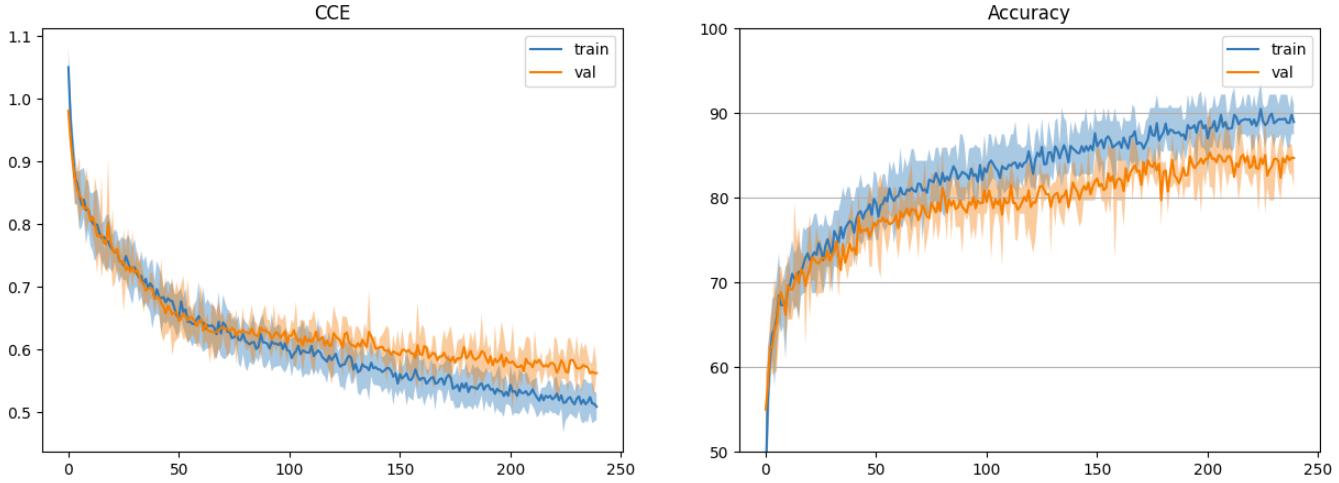


Рисунок 9 – Обучение модели с оптимальными параметрами

**Часть 6. Сравните три варианта аугментации данных: с маленьким коэффициентом, средним и большим. Проанализируйте результаты обучения.**

Возьмём 3 значения аугментации: 0.3, 0.1, 0.7.

Для значения 0.3 результат представлен на рис. 10 и точность составляет 0.85.

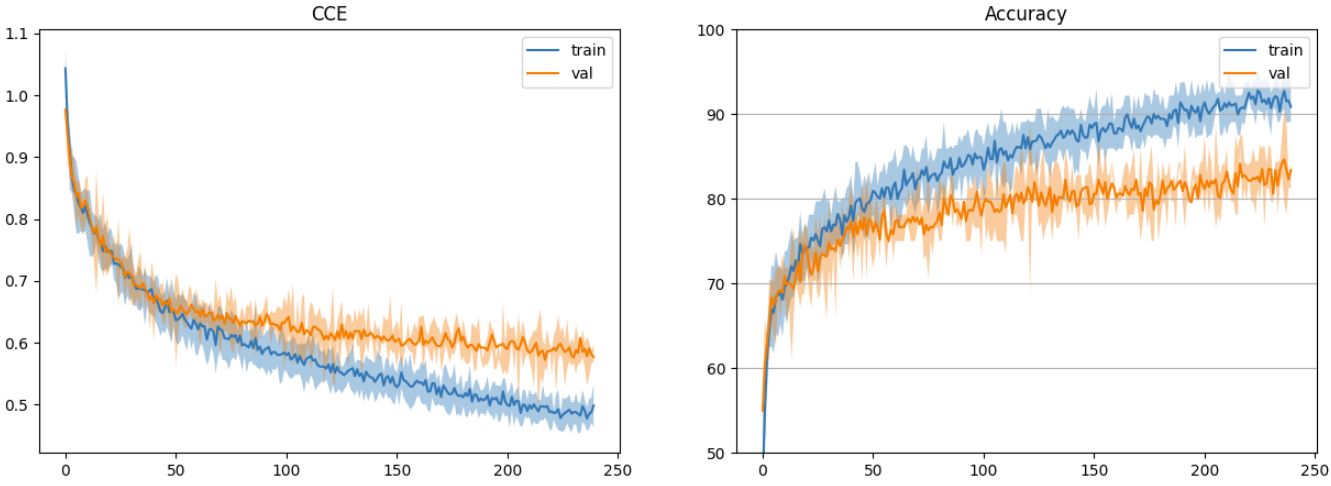


Рисунок 10 - Обучение модели, при p=0.3

Точность уменьшилась, т.к. уменьшилось количество поступающих данных для обучения.

Для значения 0.1 результат представлен на рис. 11 и точность составляет 0.8233.

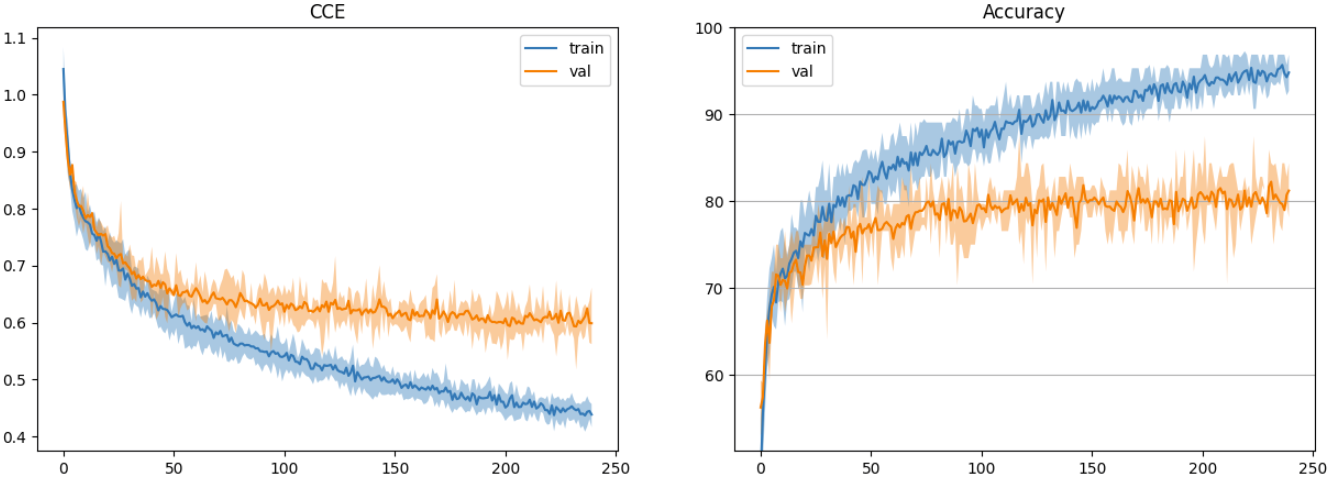


Рисунок 11 - Обучение модели, при p=0.1

Точность уменьшилась, т.к. уменьшилось количество поступающих данных для обучения.

Для значения 0.7 результат представлен на рис. 12 и точность составляет 0.8367.

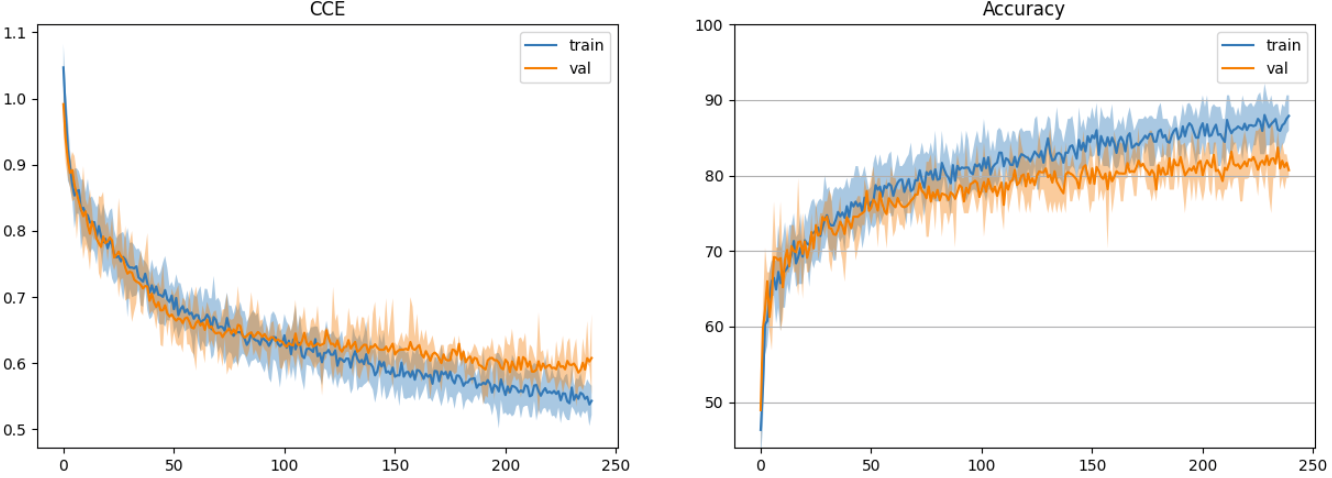


Рисунок 12- Обучение модели, при p=0.7

Точность уменьшилась, т.к. увеличилась сложность модели.

**Часть 7. Укажите, какие действия помогли повысить точность вашей модели и объясните почему.**

Повысить точность позволили следующие действия:

* Изменение гиперпараметров
* Увеличение аугментаци

**Итоговая таблица с результатами для всех вариантов обучения**

Таблица 1 - Итоговая таблица с результатами для всех вариантов обучения

|  |  |  |  |
| --- | --- | --- | --- |
| Конфигурация нейросети | Гиперпараметры | Точность | Комментарий |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-5  p=0.5 | train = 88,07%  test= 80,67% | Базовый вариант |
| CL(32, 5, stride=1, padding=2),  MaxPool(4),  Dropout(0.2),  CL(64, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  CL(64, 3, stride=1, padding=1),  FC(256) | lr=0.001  batch\_size = 256  epoch = 700  weight\_decay=1e-5  p=0.5 | train = 85,67%  test= 82,33% | Применение регуляризации и аугментации к модели из лаб.2 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.4),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-5  p=0.5 | train = 87,90%  test= 80% | Изменение значение второго дропаута на 0.4 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.1),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-5  p=0.5 | train = 88,67%  test= 80,33% | Изменение значение первого дропаута на 0.1 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.1),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.4),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-5  p=0.5 | train = 88,60%  test= 80,67% | Изменение значение дропаута на 0.1 и 0.4 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-3  p=0.5 | train = 87,47%  test= 79,0% | Изменение значение штрафа за сложность на 1e-3 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-1  p=0.5 | train = 66,93%  test= 65,00% | Изменение значение штрафа за сложность на 1e-1 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.003  batch\_size = 128  epoch = 500  weight\_decay=1e-7  p=0.5 | train = 87,07%  test= 79,67% | Изменение значение штрафа за сложность на 1e-7 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.006  batch\_size = 64  epoch = 250  weight\_decay=1e-5  p=0.5 | train = 96,53%  test= 86,00% | Изменение гиперпараметров |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.006  batch\_size = 64  epoch = 250  weight\_decay=1e-5  p=0.3 | train = 97,67%  test= 84,00% | Коэффициент аугментации p=0.3 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.006  batch\_size = 64  epoch = 250  weight\_decay=1e-5  p=0.1 | train = 99,2%  test= 82,33% | Коэффициент аугментации p=0.1 |
| CL(64, 3, stride=4, padding=0),  Dropout(0.2),  CL(128, 3, stride=1, padding=1),  AvgPool(4),  Dropout(0.3),  FC(512) | lr=0.006  batch\_size = 64  epoch = 250  weight\_decay=1e-5  p=0.7 | train = 94,40%  test= 83,67% | Коэффициент аугментации p=0.7 |

**Вывод:**

P(вероятность) лежит в диапазонах от 0.1 до 0.4, при этом у последующих слоёв большие значения p. При дропаут 0.2 маленький точность будет улучшаться, так как при больших значениях точность падает из-за потери слишком большого количества информации в процессе обучения, а при маленьких из-за вероятности возникновения переобучения. В текущей лабораторной работе теория подтверждается.

В теории для значений штрафа за сложность модели точность будет улучшаться в том случае, если значение является средним, то есть не слишком большим или не слишком маленьким. При больших значениях точность падает из-за вероятности возникновения недообучения, а при маленьких из-за вероятности возникновения переобучения. В текущей лабораторной работе теория подтверждается.

В теории при значениях коэффициента аугментации при увеличении до 1 точность будет улучшаться, так как чем больше данных для обучения, тем лучше обучение, но это может заниматься больше места или больше вычислительных ресурсов. В текущей лабораторной работе теория подтверждается, кроме случая при значении 0.7, что можно объяснить случайностью при выборе.