МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное   
учреждение высшего образования

**«Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского»   
(ННГУ)**

**Институт информационных технологий, математики и механики**

**Кафедра: алгебры, геометрии и дискретной математики**

Направление подготовки: «Прикладная математика и информатика»

Профиль подготовки: «Прикладная математика и информатика (общий профиль)»

**ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА   
БАКАЛАВРА**

**Тема:**

**«Анализ применения LSTM-сетей и трансферного обучения для диагностики сердечно-сосудистых заболеваний на наборах данных PTB-XL и ICBEB2018»**

Выполнил:

студент группы 381903\_3

\_\_\_\_\_\_\_Манухов В.В.\_\_\_\_\_\_\_

ф.и.о.

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

Подпись

Научный руководитель:

\_\_\_д.ф.-м.н. Золотых Н.Ю.\_\_\_

ф.и.о.

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

Подпись

Нижний Новгород

2023

**Оглавление**

[Аннотация 3](#_Toc138087794)

[1. Введение 4](#_Toc138087795)

[2. Материалы и методы 7](#_Toc138087796)

[2.1 Наборы данных PTB-XL и ICBEB2018 7](#_Toc138087797)

[2.2 Алгоритмы классификации временных рядов 8](#_Toc138087798)

[3. Сравнительный анализ моделей 10](#_Toc138087799)

[3.1. Задачи и метрики 10](#_Toc138087800)

[3.2. Прогнозирование заключений ЭКГ на наборе данных PTB-XL 12](#_Toc138087801)

[3.3. Прогнозирование заключений ЭКГ на наборе данных ICBEB2018. Трансферное обучение 18](#_Toc138087802)

[4. Итоги и заключение 33](#_Toc138087803)

[Список литературы 34](#_Toc138087804)

# Аннотация

Сердечно-сосудистые заболевания (ССЗ) являются одной из ведущих причин высокой смертности во всем мире [18]. Электрокардиография (ЭКГ) – это главный инструмент неинвазивной диагностики подобных патологий. В данной работе исследуется применение методов машинного обучения, в частности глубокого обучения (DL), в анализе сигналов ЭКГ на основе наборов данных PTB-XL [10] и ICBEB2018 [12]. Основное внимание в исследовании уделяется оценке качества работы рекуррентных нейронных сетей (RNN), а именно однонаправленной и двунаправленной LSTM-сетей (S. Hochreiter, J. Schmidhuber) [9], при этом основной метрикой оценки является макроусреднённая площадь под ROC кривой – macro AUC. Основным опорным материалом для данного исследования является статья «Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL» (Nils Strodthoff, Patrick Wagner, Tobias Schaeffter, Wojciech Samek) [1], в которой проведен бенчмаркинг алгоритмов глубокого обучения на основе наборов данных PTB-XL и ICBEB2018. Данная статья описывает актуальные проблемы в диагностике ССЗ и представляет идеи и методы для решения задач классификации временных рядов, представленных в виде сигналов ЭКГ. В исследовании [1] демонстрируется эффективность алгоритмов глубокого обучения, а также рассматривается применение трансферного обучения (transfer learning). В нашей работе мы использовали представленные в статье [1] архитектуры и результаты для оценки эффективности обученных моделей на наборах данных PTB-XL и ICBEB2018. В ходе исследования нами были получены результаты, подтверждающие эффективность применения LSTM-сетей для работы с временными рядами, а также перспективность применения трансферного обучения на малом объеме данных.

# Введение

Сердечно-сосудистые заболевания (ССЗ) являются ведущей причиной высокой смертности во всем мире. Главным инструментом неинвазивной диагностики ССЗ является электрокардиография (ЭКГ). Интерпретация ЭКГ представляет из себя достаточно сложную задачу, особенно в ситуациях, когда требуется немедленное обследование, но отсутствует кардиолог и имеются только врачи общей практики или службы неотложной помощи.

Для проведения данного исследования мы будем опираться на статью «Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL» (Nils Strodthoff, Patrick Wagner, Tobias Schaeffter, Wojciech Samek) [1]. В ней авторы обращают внимание на ограниченный доступ к большим медицинским данным и на отсутствие четко определенной процедуры оценки качества алгоритмов машинного обучения. В статье представлены различные методы для проведения сравнительного анализа на основе крупнейшего на данный момент открытого набора данных ЭКГ сигналов PTB-XL, а также даются результаты работы различных алгоритмов классификации, которые служат ориентиром для будущих исследований. Среди алгоритмов глубокого обучения (deep learning - DL) представлены: сверточные нейронные сети (inception1d, xresnet1d101, resnet1d\_wang, fcn wang), рекуррентные нейронные сети (lstm, lstm\_bidir, gru, gru\_bidir), нейронные сети с предварительным вейвлет преобразованием (wavelet+NN) и ансамбль (ensemble). Нами для исследования были выбраны рекуррентные нейронные сети (recurrent neural networks - RNN), а именно однонаправленные и двунаправленные LSTM-сети [9, 20]. Сравнительный анализ результатов этих алгоритмов глубокого обучения с использованием набора данных PTB-XL позволяет оценить их эффективность и точность в контексте анализа ЭКГ.

Также, авторы статьи [1] предлагают рассмотреть применение подхода трансферного обучения (transfer learning). Этот подход позволяет использовать популярные модели глубокого обучения без сильных требований со стороны вычислительных мощностей и эффективно работать с ограниченным объемом данных, доступных для обучения.

**Актуальность исследования** определяется рядом факторов**:**

1. Использование алгоритмов на основе глубоких нейронных сетей может существенно облегчить работу врачей, предоставляя им системы поддержки и диагностики для автоматической интерпретации ЭКГ и состояний пациентов. Такая помощь может заключаться, например, в скорости выполнения задачи и высокой точности результата.
2. Развитие алгоритмов для автоматической интерпретации ЭКГ является важным направлением, поскольку это может помочь улучшить скорость и точность диагностики ССЗ.
3. Существует необходимость в создании четких процедур тестирования алгоритмов автоматической интерпретации для оценки эффективности и надежности разработанных алгоритмов и их применимости в клинической практике.
4. Анализ эффективности трансферного обучения представляет интерес в решении проблем, связанных с недостатком данных или вычислительных ресурсов для обучения моделей.

**Объектом исследования** являются методы машинного обучения в задаче анализа медицинских сигналов.

**Предметом исследования** являются рекуррентные нейронные сети в задаче анализа сигналов ЭКГ и диагностики ССЗ.

**Целями** в данном исследовании являются следующие пункты**:**

* Оценка качества работы однонаправленных и двунаправленных LSTM-сетей на наборах данных PTB-XL и ICBEB2018.
* Проведение сравнительного анализ моделей на основе полученных оценок.
* Измерение эффективности применения трансферного обучения (transfer learning) на небольших наборах данных.

**Апробация работы.** Часть положений данной работы излагалась на защите итогового проекта на НЕЙМАРК ML-интенсивe 14 апреля 2023 года.

**Структура работы** состоит из введения, глав про материалы и методы, сравнительный анализ моделей, итогов исследования и списка литературы.

Во **введении** представляется проблемы рассматриваемой области, обосновывается актуальность исследования, обозначаются объект и предмет, а также цели.

В главе про **материалы и методы** описываютсяданные PTB-XL и ICBEB2018 с которыми нам предстоит работать и приводится описание архитектур моделей для классификации временных рядов.

Глава со **сравнительным анализом моделей** содержит в себе постановку основных задач при проведении экспериментов, а также приводятся метрики, с помощью которых будет происходит сравнение работы различных моделей. В этой главе мы даем результаты проведения экспериментов и заключения, основанные на полученных данных.

В последней содержательной главе представлены **итоги и заключение** исследования. В этой части работы подчеркивается значимость полученных результатов, а также указываются возможные аспекты, которые можно улучшить при проведении дальнейших исследований.

# Материалы и методы

## Наборы данных PTB-XL и ICBEB2018

Для достижения высокой надежности и точности в автоматической интерпретации ЭКГ необходим доступ к обширным наборам данных, охватывающим разнообразные типы сердечных патологий. Для решения данной проблемы предлагаются два клинических набора данных ЭКГ: PTB-XL [10] и ICBEB2018 [12]. Благодаря доступу к PTB-XL, размещенному на ресурсе PhysioNet [11], проблема нехватки данных для исследований в области медицинских сигналов стала решаемой для широкой аудитории исследователей. Представленный набор данных содержит информацию по 12 отведениям, что позволяет получить более полное представление о сердечной активности и повысить точность диагностики. Отметим, что с точки зрения математико-статистических методов анализа, сигнал ЭКГ рассматривается как временной ряд.

PTB-XL – это публичный набор данных, который является основным для наших дальнейших экспериментов. PTB-XL – включает в себя 21837 клинических записей по 12-отведениям ЭКГ длительностью 10 секунд от 18885 пациентов, где 52% составляют мужчины и 48% женщины. Всего предоставлена 71 расшифровка ЭКГ, которые соответствуют стандарту SCP-ECG [19]. Они были отнесены к трем не взаимоисключающим категориям:

* 44 диагностических расшифровки (далее - diag), например, «передний инфаркт миокарда». Диагностические расшифровки имеют иерархическую организацию на 5 суперклассов (далее - superdiag) и 24 подкласса (далее - subdiag).
* 19 (4 используются в диагностических расшифровках) изменений определенных сегментов ЭКГ (далее - form), например, «аномальный комплекс QRS».
* 12 изменений ритма (далее - rhythm), например, «Фибрилляция предсердий».

Благодаря тому, что PTB-XL является большим набором данных, который помимо временных рядов сопровождается богатым набором аннотаций ЭКГ и дополнительными метаданными, он является уникальным материалом для обучения и оценки работы моделей машинного обучения. В данной работе мы используем данные с частотой дискретизации 100 Гц, а также рекомендованное разделение данных на обучающую и тестовую выборки [10].

Рассмотрим набор данных ICBEB2018. Он содержит 6877 записей по 12 отведениям, длительностью от 6 до 60 секунд. Каждая запись ЭКГ аннотируется до трех расшифровок, взятых из набора девяти классов (класс здоровый (NORM) и 8 классов патологий). Авторы статьи [1] используют объединение меток, создавая набор данных с несколькими метками. Также, исходный набор разделяется на 10 выборок, аналогично набору данных PTB-XL, сохраняя общее распределение меток [10]. В свою очередь, набор данных PTB-XL рассматривается нами как основной для проведения предварительного обучения классификатора в задаче трансферного обучения, где ICBEB2018 выступает в роли малочисленного ресурса для обучения.

Таким образом, наличие набора данных PTB-XL открывает возможности для решения различных задач в сфере медицинской диагностики, а наличие набора данных ICBEB2018 позволяет расширить работу и рассмотреть область перспективного трансферного обучения.

Для обучения моделей будут предоставлены исходные наборы данных без возможной их предобработки, которая свойственна для временных рядов, такой как избавление от шума или борьба с дисбалансом классов, за исключением того, что мы провели стандартизацию данных, а также избавились от пустых значений.

## Алгоритмы классификации временных рядов

В данном исследовании нами рассматривается задача многозначной классификации (multi-label classification) [3;139-168], где классификатор в соответствие образцу выдает вектор двоичных меток. Такая задача имеет место при анализе сигналов ЭКГ, когда мы хотим по ним установить всевозможные заболевания у пациентов. Основываясь на статье [1], мы приведем архитектуры нейронных сетей, которые будут работать с необработанными многомерными данными временных рядов.

Авторы статьи используют в своих архитектурах слой объединения «concat-pooling» [1; 1521], который агрегирует результаты слоев подвыборки «GlobalMaxPooling1D» и «GlobalAveragePooling1D» [16]. Мы будем использовать рекуррентные нейронные сети, а именно однонаправленные (unidirectional) и двунаправленные (bidirectional) LSTM-сети (S. Hochreiter, J. Schmidhuber) [9] с двумя скрытыми слоями LSTM (long short-term memory – долгая краткосрочная память), содержащими по 256 нейрона в каждом, выходы которых объединяются с помощью упомянутого слоя «concat-pooling». Выходным слоем будет являться полносвязный слой с числом нейронов равных количеству классов, где в качестве функции активации будет функция сигмоиды [3; 207-208, 367-430]. В дальнейшем тексте, для удобства, мы будем использовать обозначения «lstm» и «lstm\_bidir» для однонаправленных и двунаправленных LSTM-сетей соответственно.

Функцией потерь была выбрана бинарная перекрестная энтропия (binary сross-entropy), которая может быть применена в задачах многозначной классификации [3; 367-430, 679-731]. Во время обучения модели мы используем метод управления скоростью обучения «1-cycle» [2; 447-465] и оптимизатор AdamW [16, 17].

При реализации подхода трансферного обучения мы применяем метод тонкой настройки (finetuning). Этот подход включает в себя использование предварительно обученной модели (pretrained model), после чего мы либо удаляем, либо переобучаем последние слои модели. Таким образом, мы стараемся усилить внимание модели к деталям новых данных, на которых она будет обучаться. Мы используем постепенное размораживание и дифференцированное обучение (differential learning) [1; 1521].

Для разработки программной части нашего исследования мы используем язык программирования Python [21] и ключевые библиотеки в области машинного обучения: Scikit-learn [14] и Keras [16].

# Сравнительный анализ моделей

## Задачи и метрики

Для достижения поставленных целей данного исследования нам необходимо провести 3 серии экспериментов:

1. Первая серия содержит в себе эксперименты по работе моделей на наборе данных PTB-XL.
2. Вторая серия содержит в себе эксперименты по работе моделей на наборе данных ICBEB2018.
3. Третья серия содержит в себе применение трансферного обучения с использованием обоих наборов данных.

В рамках каждого эксперимента будут выполнены следующие действия:

1. Загрузка данных с указанием категории (например, diag).
2. Обучение модели (например, lstm) на тренировочном наборе данных.
3. Оценка качества работы модели на тестовых данных.

Сравнительный анализ моделей lstm и lstm\_bidir будет выполнен внутри каждой серии, а после мы сможем сделать выводы и дать общую оценку качества их работы, основываясь на результатах всех экспериментов данного исследования. Далее приводятся задачи и подсчет количества экспериментов внутри каждой серии:

1. Провести загрузку наборов данных PTB-XL и ICBEB2018 и подготовить их для обучения моделей. Для загрузки используются материалы статьи [1] и ресурсы [11] и [12].
2. Используя материалы статьи [1], реализовать архитектуры однонаправленных и двунаправленных LSTM-сетей.
3. Провести обучение моделей lstm и lstm\_bidir на наборе данных PTB-XL и выполнить сравнительный анализ полученных результатов на тестовых данных. В случае PTB-XL мы имеем 6 категорий (all, diag, superdiag, subdiag, rhythm, form). Итого, мы должны будем провести для PTB-XL 12 экспериментов (2 модели и 6 категорий).
4. Провести обучение моделей lstm и lstm\_bidir на наборе данных ICBEB2018 и выполнить сравнительный анализ полученных результатов на тестовых данных. В случае ICBEB2018 мы имеем 5 категорий (all, diag, superdiag, subdiag, form). Итого, мы должны будем провести для ICBEB2018 - 10 экспериментов (2 модели и 5 категорий).
5. Оценить перспективность применения трансферного обучения на малом объеме данных в анализе сигналов ЭКГ. Для этого мы используем предобученную на всех расшифровках ЭКГ (категория all) набора данных PTB-XL модель, а в качестве дополнительного набора данных будем использовать ICBEB2018, причем, ICBEB2018 будет разделен на 8 выборок: ICBEB2018, ICBEB2018, ICBEB2018, ICBEB2018, ICBEB2018, ICBEB2018, ICBEB2018 (итого 7 экспериментов).
6. Провести сравнительный анализ с результатами статьи [1].

В качестве основной метрики производительности моделей машинного обучения для всех экспериментов классификации нами используется макроусреднённая площадь под ROC-кривой (receiver operating characteristic – рабочая характеристика приемника) – macro AUC (area under curve – площадь под кривой) [3; 153-157], которая получается путем усреднения значений AUC по всем классам [14]. Для удобства «macro AUC» в дальнейшем будем называть просто «AUC». Мы используем макроусреднение AUC, поскольку ожидаем дисбаланс классов и не хотим, чтобы в оценке производительности модели доминировали крупные классы. Также, мы не применяем пороговые значения, чтобы решить, к какому классу отнести каждый элемент, потому что это позволяет получить более полную картину дискриминационной способности алгоритма классификации, то есть способность модели отличать разные классы данных. Отказываясь от порога, мы исключаем вопрос его оптимизации, который должен корректироваться в соответствии с клиническими требованиями, а не для оптимизации какой-либо глобальной целевой метрики.

Для ICBEB2018 приводится две дополнительные метрики производительности, а именно макроусреднение и макроусреднение . Приведем формулы для данных метрик:

где величины TP, FN, FP берутся из матрицы ошибок (confusion matrix) [3; 145-171], а параметр принимает любые вещественные значения. Значение > 1 позволяет придать больше значения полноте (recall) в ущерб точности (precision) [3; 145-152], что является предпочтительным, если нашей целью является максимизация правильно определенных больных пациентов. Однако такой подход может иметь противоположный эффект на класс здоровых пациентов, где требуется более высокий показатель точности. Поэтому выбор оптимального значения зависит от конкретной задачи и требуемого баланса между полнотой и точностью [1].

Все результаты оцениваются соответственно статье [1]. Полученные в ней результаты служат для проведения сравнительного анализа, например, качества работы моделей классического машинного обучения и глубокого обучения.

## Прогнозирование заключений ЭКГ на наборе данных PTB-XL

Мы приводим таблицу 1, которая содержит результаты всех шести экспериментов, а также графики ошибки и метрики AUC при обучении. Значения метрик оценивают качество моделей-классификаторов lstm и lstm\_bidir при работе с набором данных PTB-XL.

Таблица 1.

Результаты работы моделей lstm и lstm\_bidir для набора данных PTB-XL

|  |  |
| --- | --- |
| Модель, категория | AUC |
| lstm, all | 0.905817 |
| lstm\_bidir, all | 0.904496 |
| lstm, diag | 0.925139 |
| lstm\_bidir, diag | 0.926086 |
| lstm, superdiag | 0.8982 |
| lstm\_bidir, superdiag | 0.904164 |
| lstm, subdiag | 0.922938 |
| lstm\_bidir, subdiag | 0.924261 |
| lstm, rhythm | 0.966073 |
| lstm\_bidir, rhythm | 0.980625 |
| lstm, form | 0.873378 |
| lstm\_bidir, form | 0.887075 |

Глядя на таблицу, можно отметить, что оценка варьируется от 0.873378 до 0.980625, если сравнивать категории form и rhythm. Но стоит отметить, что обучающее множество для категории form много меньше по сравнению с другими категориями, поскольку около 12 тысяч ЭКГ не содержат аннотаций для неё. Поэтому в этой категории модели lstm и lstm\_bidir показывают низкие оценки: 0.873378 и 0.887075 соответственно. На наиболее больших категориях, таких как all или diag, мы можем наблюдать достаточно высокие оценки в диапазоне от 0.904496 и до 0.926086. Полученные нами результаты очень близки к «эталонным», представленным в статье [1]. Особенно заметен значительный прирост производительности для моделей lstm и lstm\_bidir в категории rhythm, где мы достигли оценок 0.966073 и 0.980625 соответственно. Сопоставляя наши результаты с лучшей моделью статьи [1], а именно xresnet1d101 можно заключить, что рекуррентные нейронные сети имеют большой потенциал в решении задач многозначной классификации временных рядов.

Сравнивая модели lstm\_bidir и lstm отдельно в категориях rhythm и form, можно заметить, что результат модели lstm\_bidir почти на 0.02 пункта выше, чем у модели lstm.

Общее сравнение по всем категориям показывает, что модели lstm и lstm\_bidir на данном наборе данных демонстрируют одинаково высокие значения AUC. Это говорит о их высокой эффективность в классификации сигналов ЭКГ, представленных в виде временных рядов. Также, это указывает на то, что использование двунаправленной LSTM-сети не всегда приводит к существенному преимуществу по сравнению с однонаправленной LSTM-сетью.

Далее представлены графики измерений (рис. 1-3) во время обучения, очередность которых соответствует результатам в таблице. На графике значения потерь отложены по оси ординат на левой половине, а значения метрики AUC отложены по оси ординат на правой половине. По оси абсцисс отображается количество эпох, на протяжении которых происходили измерения.

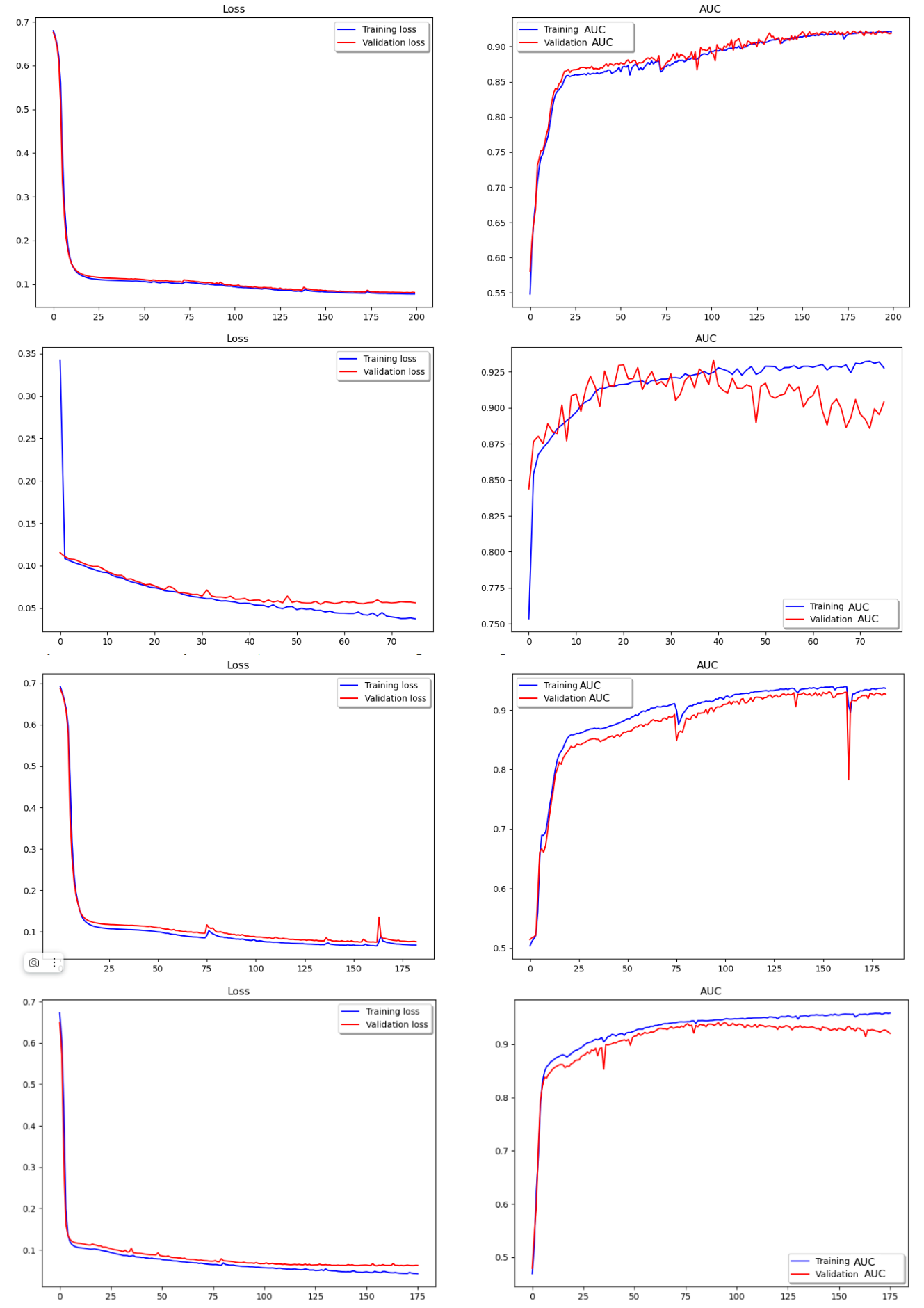


Рис. 1. Набор данных PTB-XL. Категории all - diag.

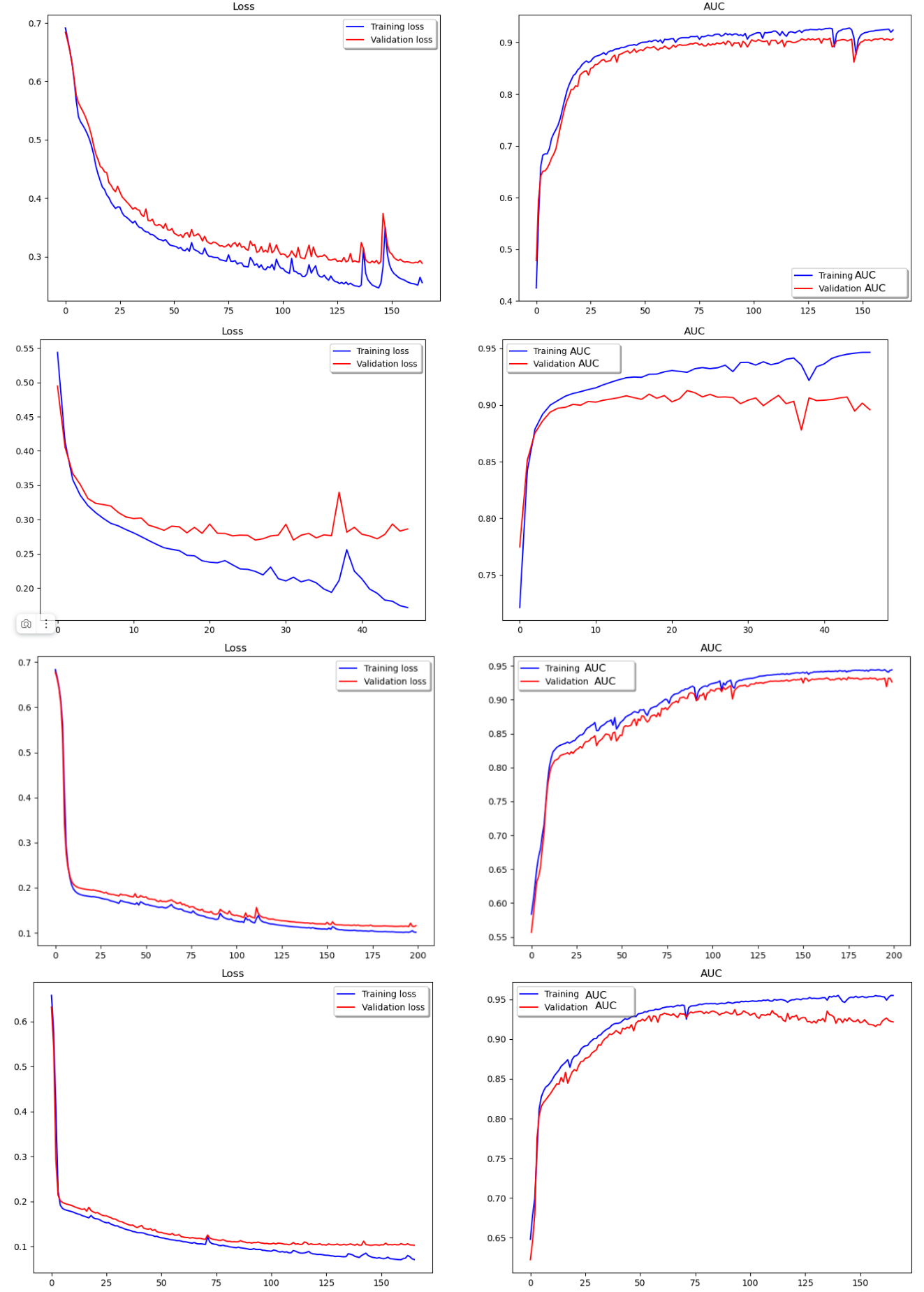


Рис. 2. Набор данных PTB-XL. Категории superdiag - subdiag.

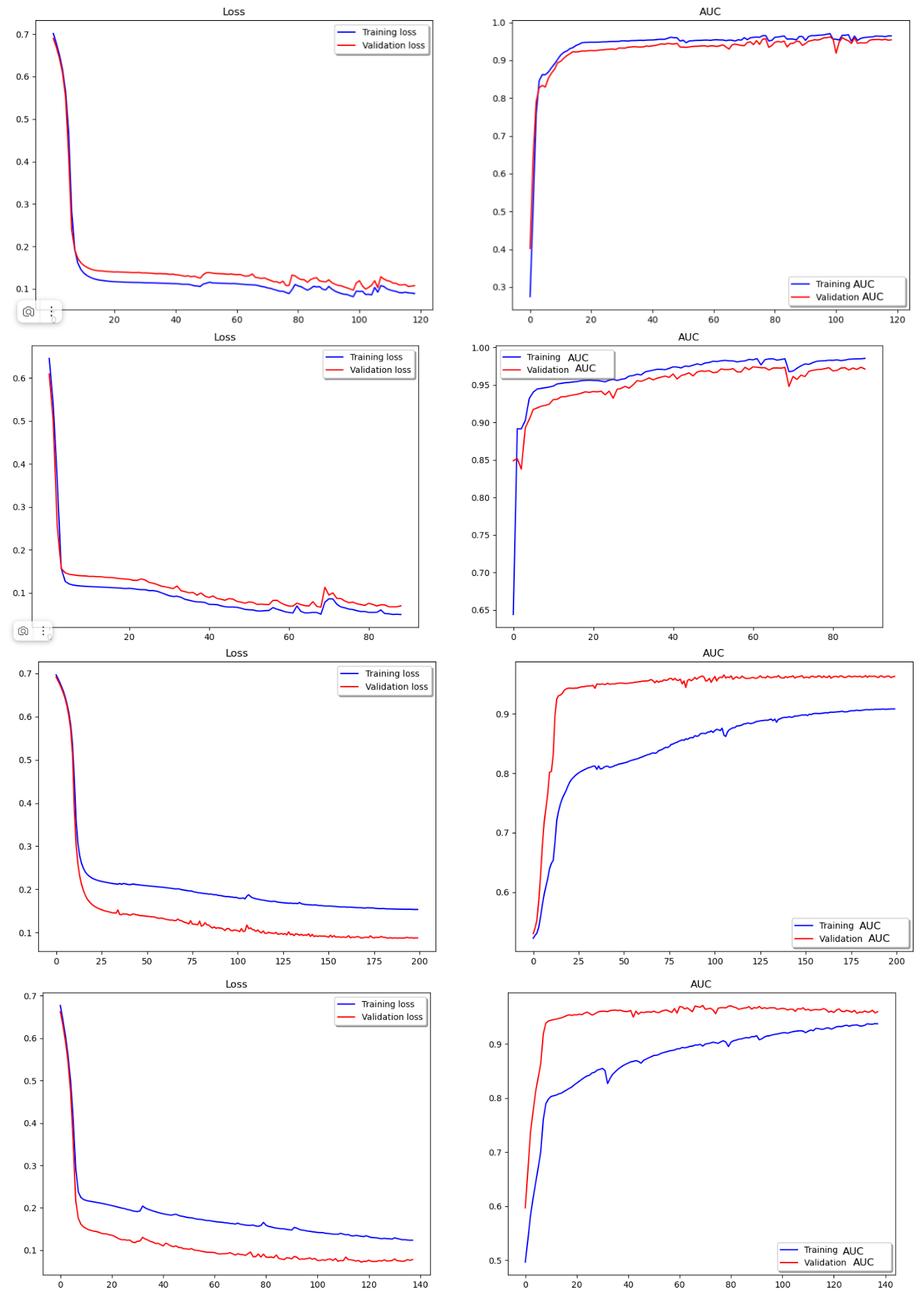


Рис. 3. Набор данных PTB-XL. Категории rhythm - form.

## Прогнозирование заключений ЭКГ на наборе данных ICBEB2018. Трансферное обучение

Здесь мы приводим таблицу 2 и графики с результатами оценки качества моделей-классификаторов lstm и lstm\_bidir при работе с набором данных ICBEB2018, а также таблицу 3 с результатами оценки качества моделей, обученных по методу трансферного обучения. В качестве базовой модели для трансферного обучения была взята модель lstm, обученная на категории all набора данных PTB-XL.

Таблица 2.

Результаты работы моделей lstm и lstm\_bidir для набора данных ICBEB2018.

|  |  |  |  |
| --- | --- | --- | --- |
| Модель, категория | AUC |  |  |
| lstm, all | 0.927058 | 0.639206 | 0.397033 |
| lstm\_bidir, all | 0.948024 | 0.72525 | 0.498124 |
| lstm, diag | 0.951087 | 0.772821 | 0.577438 |
| lstm\_bidir, diag | 0.978353 | 0.848132 | 0.721983 |
| lstm, superdiag | 0.976456 | 0.886053 | 0.8125 |
| lstm\_bidir, superdiag | 0.981656 | 0.920987 | 0.843602 |
| lstm, subdiag | 0.9778 | 0.875125 | 0.729437 |
| lstm\_bidir, subdiag | 0.979296 | 0.936381 | 0.819444 |
| lstm, form | 0.928469 | 0.779212 | 0.578723 |
| lstm\_bidir, form | 0.946614 | 0.798212 | 0.594828 |

Из таблицы видно, что модели lstm и lstm\_bidir достигают высоких результатов во всех категориях. При сравнении этих моделей можно сделать вывод, что двунаправленная рекуррентная нейронная сеть (lstm\_bidir) имеет небольшое преимущество в метрике AUC, но разница составляет не более 0.03 пункта. Отсюда мы можем сделать вывод, что двунаправленная LSTM-сеть демонстрирует более высокую производительность по сравнению с однонаправленной LSTM-сетью на относительно небольшом наборе данных (ICBEB2018 практически в 3 раза меньше, чем набор данных PTB-XL). Наибольшие различия в значениях метрики AUC между lstm и lstm\_bidir наблюдаются в категориях all и diag: 0.948024 против 0.927058 и 0.978353 против 0.951087 соответственно.

Наибольшие значения метрик и обычно соответствуют наибольшему значению метрики AUC. В большинстве категорий модель lstm\_bidir также показывает лучшие результаты по этим метрикам по сравнению с моделью lstm. Это указывает на то, что двунаправленная LSTM-сеть обеспечивает более точное и сбалансированное предсказание классов.

Далее представлены графики измерений (рис. 4-6) во время обучения, очередность которых соответствует результатам в таблице. На графике значения потерь отложены по оси ординат на левой половине, а значения метрики AUC отложены по оси ординат на правой половине. По оси абсцисс отображается количество эпох, на протяжении которых происходили измерения.

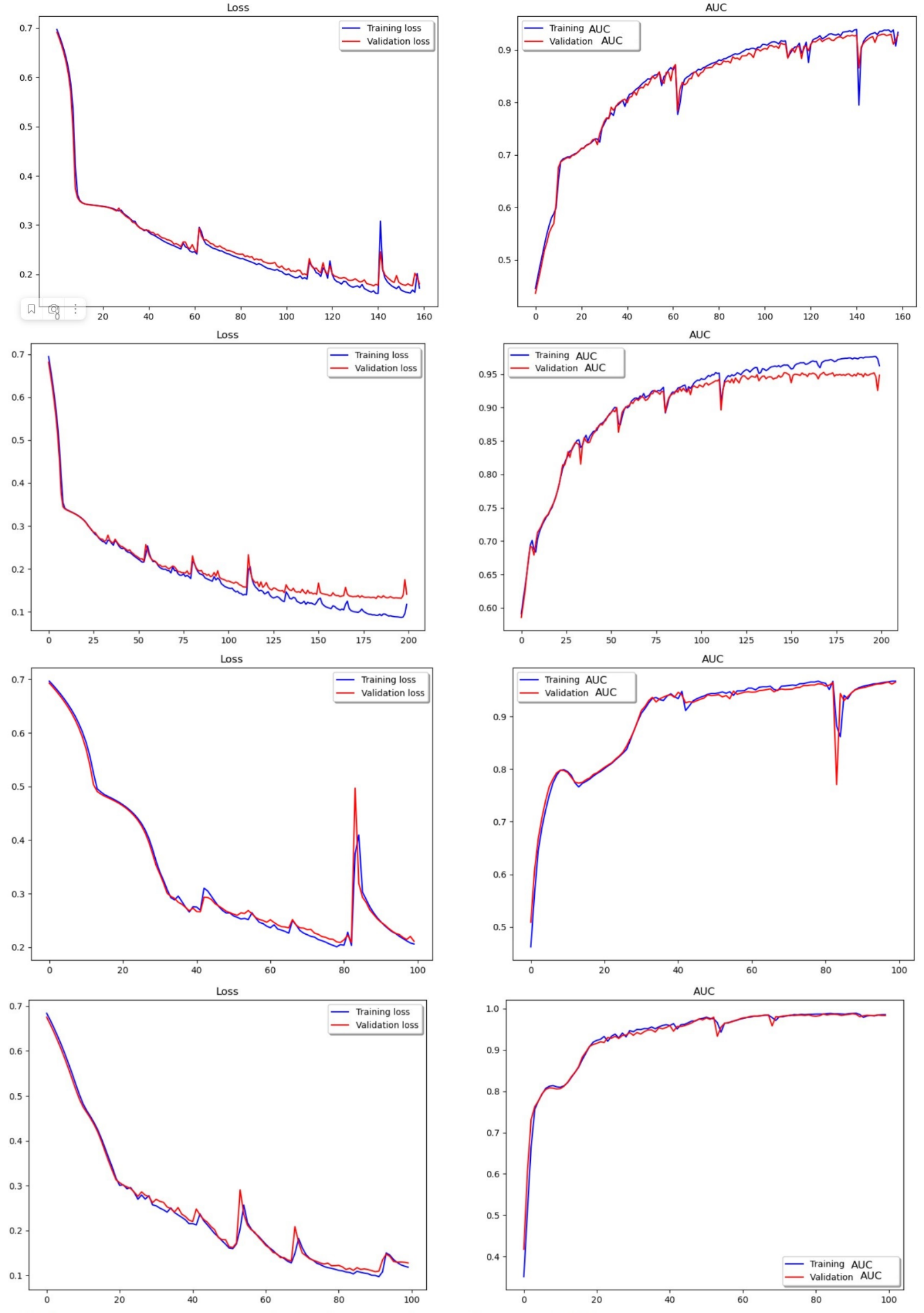


Рис. 4. Набор данных ICBEB2018. Категории all - diag.

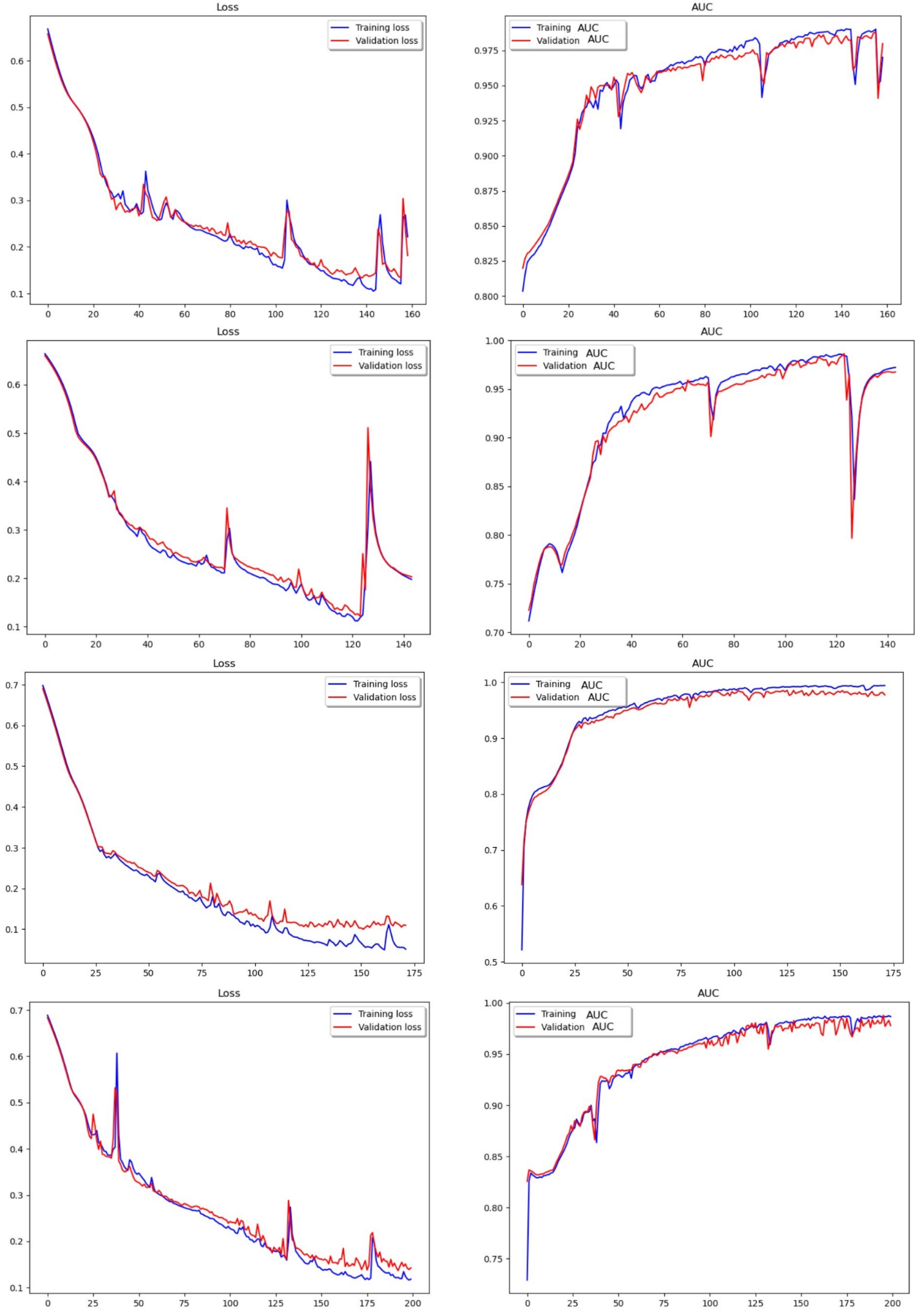


Рис. 5. Набор данных ICBEB2018. Категории superdiag - subdiag.

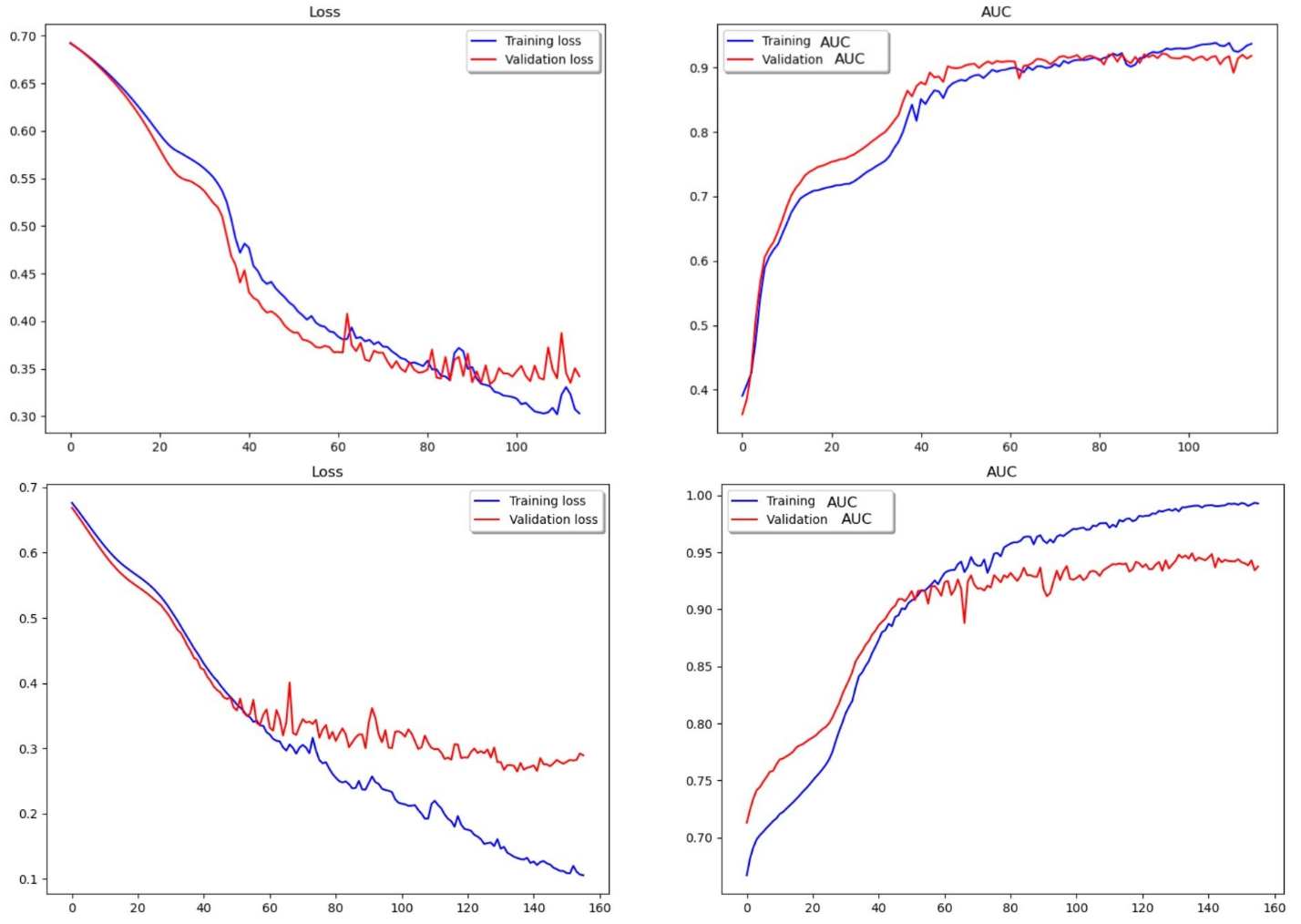


Рис. 6. Набор данных ICBEB2018. Категория form.

Таблица 3.

Результаты применения трансферного обучения для набора данных ICBEB2018. (1) – использовалась не предобученная модель lstm. (2) – использовалась предобученная модель lstm. Количество данных – часть от исходного тренировочного набора

|  |  |  |  |
| --- | --- | --- | --- |
| (1) Количество данных | AUC | (2) Количество данных | AUC |
|  | 0.681826 |  | 0.72637 |
|  | 0.717801 |  | 0.787981 |
|  | 0.774448 |  | 0.814242 |
|  | 0.823121 |  | 0.815335 |
|  | 0.813335 |  | 0.864923 |
|  | 0.836932 |  | 0.910273 |
| 1 | 0.89913 |  | 0.933337 |

Таким образом, используя данные двух наборов PTB-XL и ICBEB2018 было проведено исследование в области трансферного обучения.

Из данной таблицы можно сделать следующие выводы. Во-первых, с увеличением количества данных, значение метрики AUC увеличивается, что указывает на то, что большее количество данных позволяет классификатору делать более точные предсказания. Во-вторых, дообученная модель демонстрирует более высокую производительность при уменьшении размера обучающего набора и превосходит модель, обученную с нуля. Это хорошо заметно на рисунке 14, где показаны столбчатые диаграммы, визуализирующие результаты, указанные в таблице 3. Отметим, что использование обучающей выборки, равной 1/8 от исходного набора данных, приводит к всего лишь 13% снижению значения метрики AUC у дообученного классификатора по сравнению с классификатором, обученным на наборе, который в восемь раз больше. Аналогично, при использовании обучающей выборки, равной 1/64 от исходного набора данных, снижение составляет 23% по сравнению с использованием полного исходного набора данных.

Далее представлены графики (рис. 7-13) измерений во время обучения. Графики рассматриваются блоками по 4 штуки, где два верхних графика соответствуют эксперименту без использования трансферного обучения, а два нижних с применением этого подхода. На графике значения потерь отложены по оси ординат на левой половине, а значения метрики AUC отложены по оси ординат на правой половине. По оси абсцисс отображается количество эпох, на протяжении которых происходили измерения.



Рис. 7. Набор данных ICBEB2018, результат применения трансферного обучения на 1/64 от исходного тренировочного набора. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

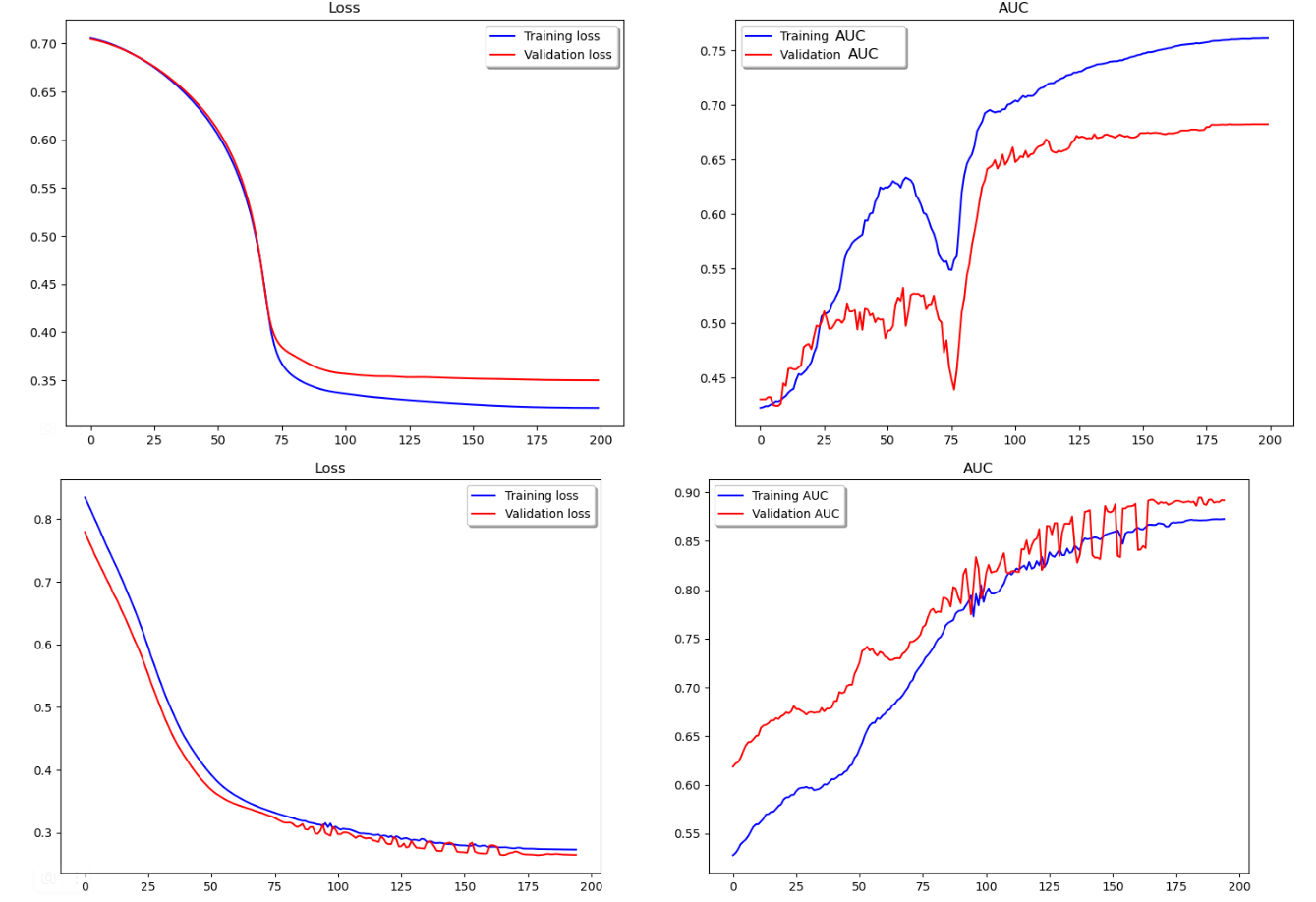


Рис. 8. Набор данных ICBEB2018, результат применения трансферного обучения на 1/32 от исходного тренировочного набора. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

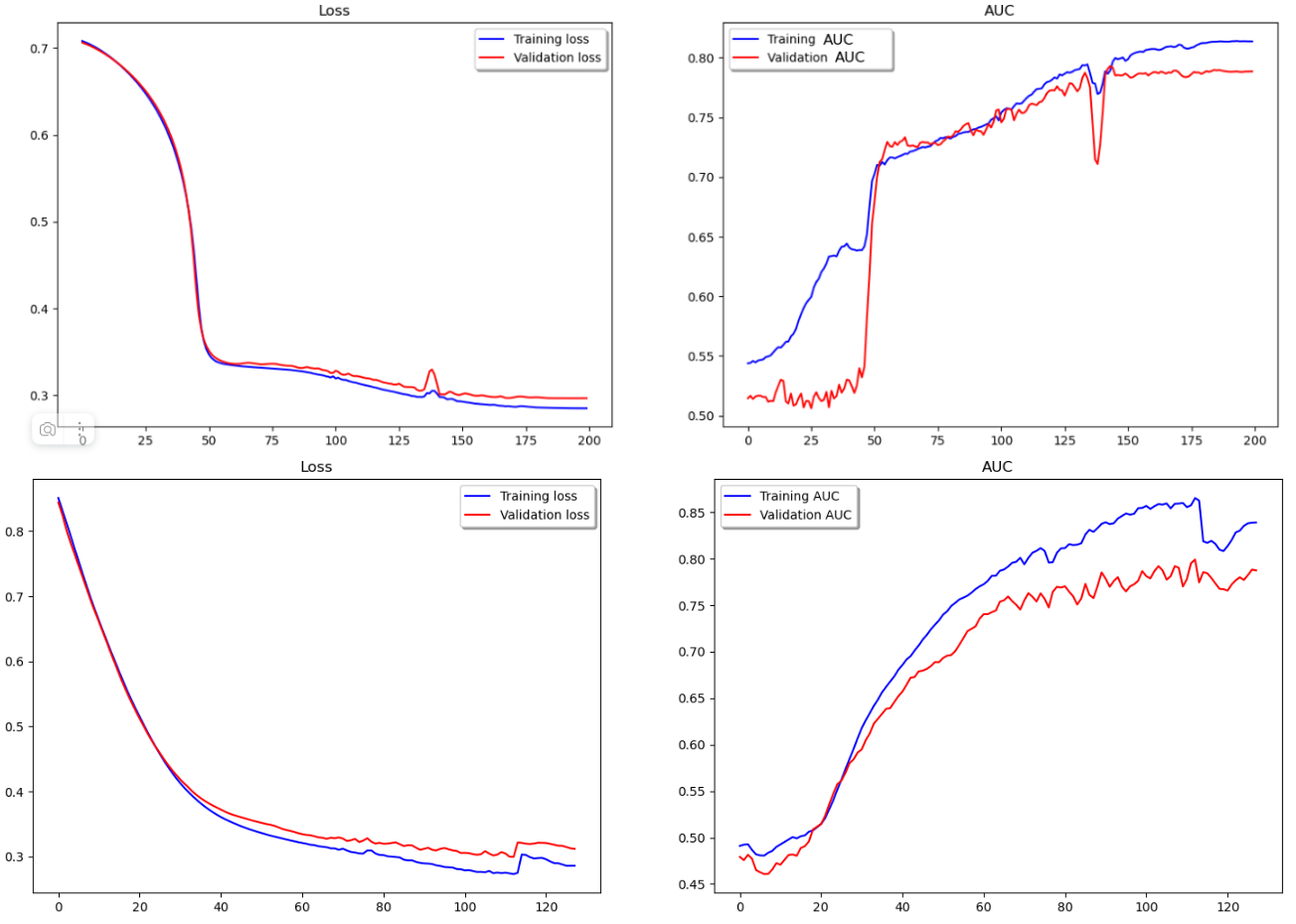


Рис. 9. Набор данных ICBEB2018, результат применения трансферного обучения на 1/16 от исходного тренировочного набора. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

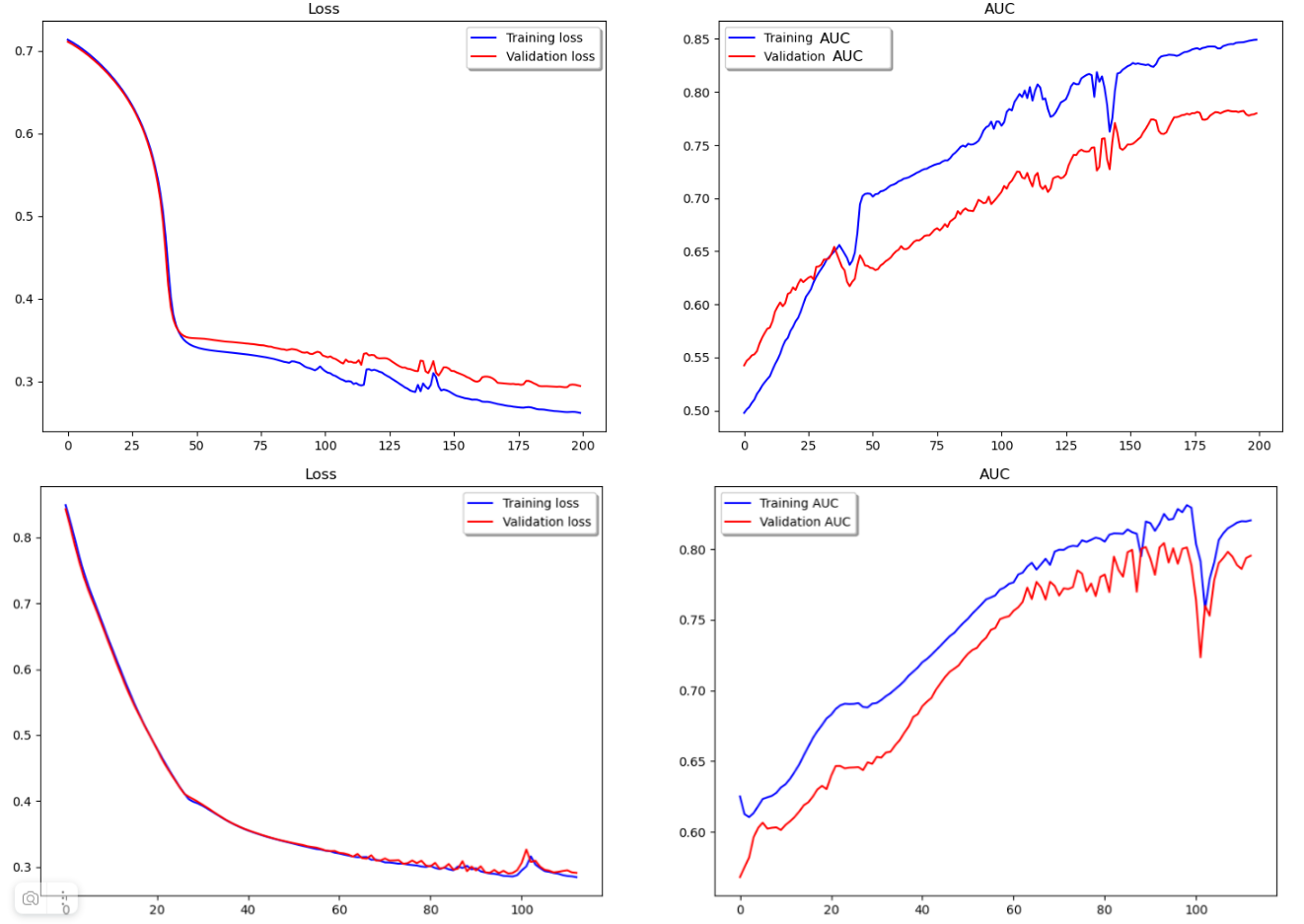


Рис. 10. Набор данных ICBEB2018, результат применения трансферного обучения на 1/8 от исходного тренировочного набора. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

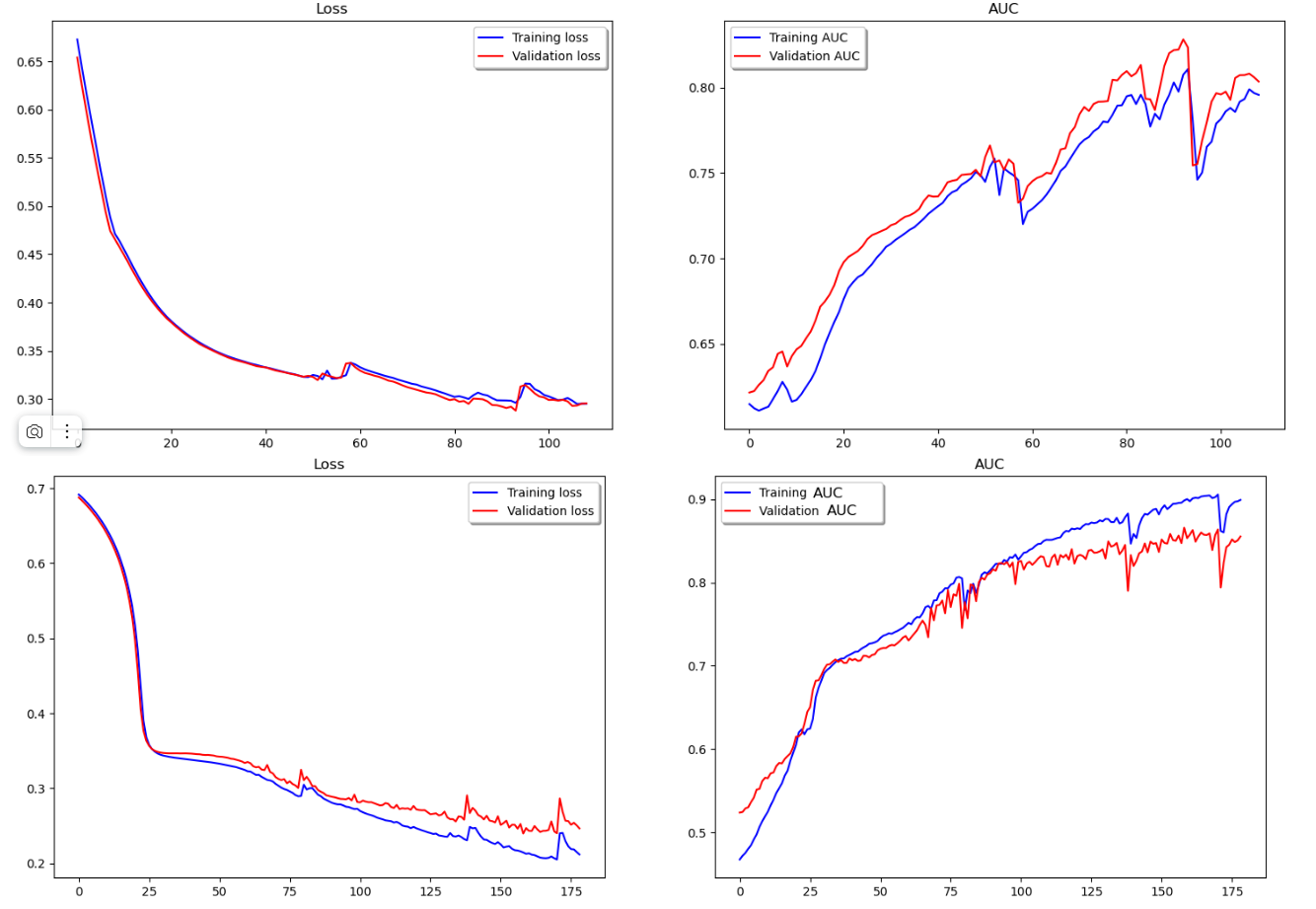


Рис. 11. Набор данных ICBEB2018, результат применения трансферного обучения на 1/4 от исходного тренировочного набора. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

Рис. 12. Набор данных ICBEB2018, результат применения трансферного обучения на 1/2 от исходного тренировочного набора. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

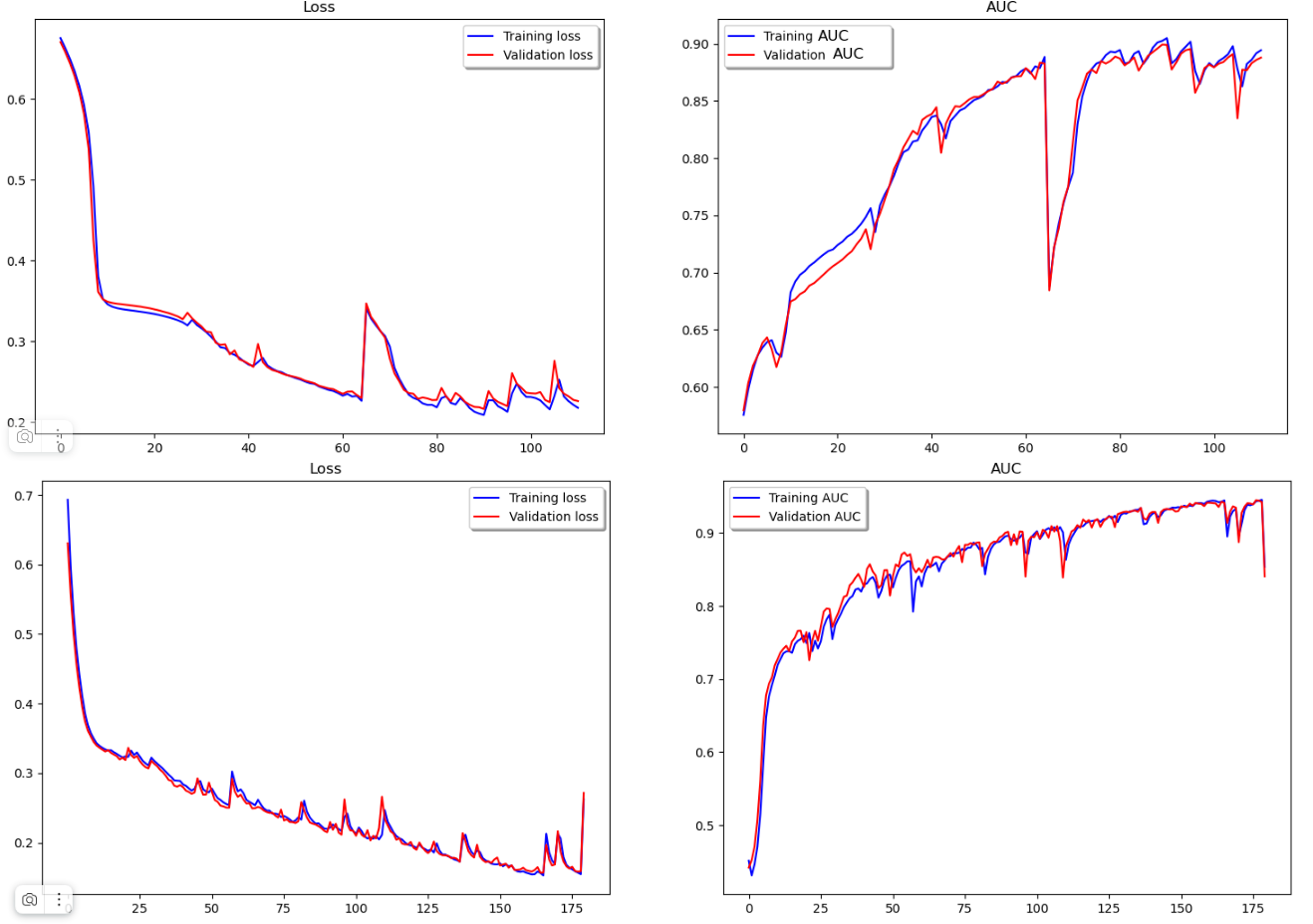
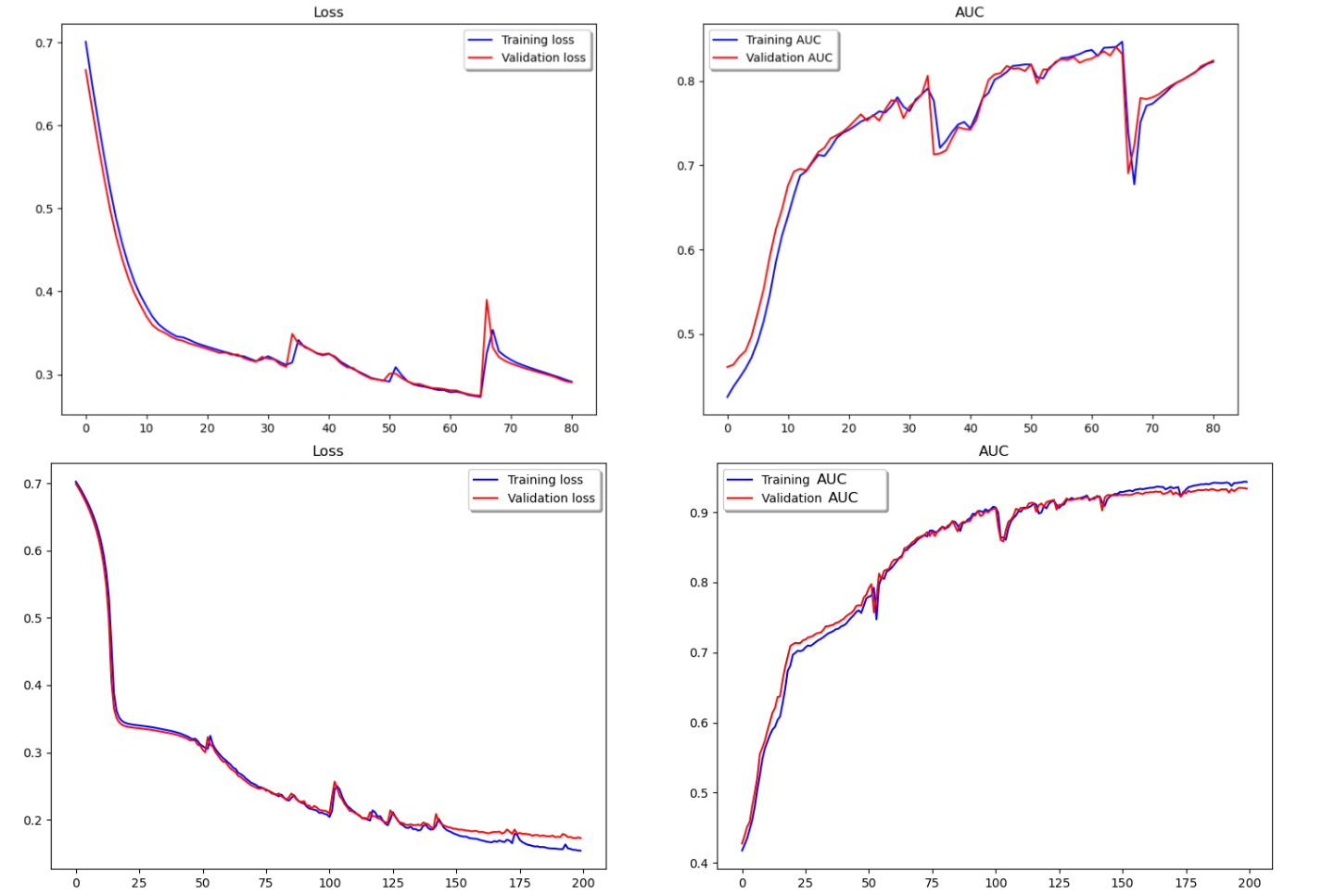


Рис. 13. Набор данных ICBEB2018, результат применения трансферного обучения на исходном тренировочном наборе. Вверху – до применения трансферного обучения, внизу – после применения трансферного обучения.

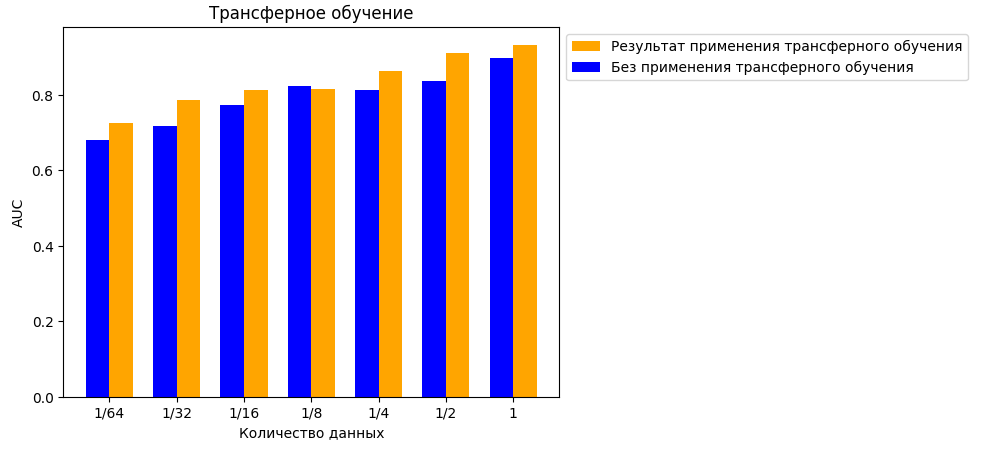


Рис. 14. Столбчатые диаграммы демонстрируют прирост значений метрики AUC от применения трансферного обучения на наборе данных ICBEB2018 при использовании модели lstm

Таким образом, результаты исследования показывают, что использование предварительно обученной на всем наборе данных PTB-XL модели может быть полезным при работе с относительно небольшими наборами данных, улучшая производительность моделей классификации.

Подводя итоги на основе полученных результатов, можно утверждать, что рекуррентные нейронные сети способны качественно решать поставленную задачу многозначной классификации временных рядов. Сравнивая модели lstm и lstm\_bidir по показателю AUC, делаем вывод, что они очень близки по своему качеству решения задачи, но чаще модель lstm\_bidir показывала результат лучше.

# Итоги и заключение

Результаты наших экспериментов подтверждают значительный потенциал алгоритмов автоматической интерпретации ЭКГ. Мы основываемся на идеях статьи [1] и наши материалы могут быть использованы для будущих исследований. Благодаря наличию набора данных PTB-XL, мы сегодня имеем возможность внести свой вклад в развитие алгоритмов автоматической интерпретации медицинских сигналов, в частности ЭКГ. Это особенно важно, так как существует проблема ограниченного доступа к большому объему медицинской информации. Таким образом, наше исследование подтверждает, что набор данных PTB-XL играет важную роль в диагностике ССЗ.

Стоит отметить, что рекуррентные нейронные сети, в частности LSTM-сети, показывают высокую эффективность в решении задач классификации. Из наших экспериментов следует, что модель lstm\_bidir, то есть двунаправленная LSTM-сеть, демонстрирует наилучшие результаты. Также, мы видим, что обучение на небольшом количестве данных приводит к плохим результатам. Поэтому мы также оценили перспективы применения трансферного обучения, которое показывает себя особенно полезным в ситуации, когда данных мало или присутствует нехватка вычислительных мощностей.

Кроме того, наши результаты указывают на потенциал для дальнейшего улучшения самих моделей. В частности, проведение подбора гиперпараметров, экспериментирование с количеством слоев, их архитектурой, функциями активации, параметром регуляризации и оптимизатором.

Также, одним из потенциальных улучшений является предварительная обработка данных в форме временных рядов. В этом контексте важными факторами являются фильтрация шумов, для которой можно применить прямое и обратное преобразование Фурье. Также важным аспектом является борьба с дисбалансом классов, например, с использованием методов аугментации данных.

# Список литературы

1. Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL. Nils Strodthoff, Patrick Wagner, T. Schaeffter, W. Samek. DOI: 10.1109/JBHI.2020.3022989 Corpus ID: 216562803, Published 28 April 2020, IEEE Journal of Biomedical and Health Informatics.
2. Aurélien Géron. Published by O’Reilly Media, Inc. 2019. – 856 p.
3. Жерон, Орельен. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем, 2-е изд.: Пер. с англ. — СПб.: ООО “Диалектика”, 2020. — 1040 с.: ил. — Парал. тит. англ.
4. Andreas C. Mueller and Sarah Guido. Introduction to Machine Learning with Python. Published by O’Reilly Media, Inc. 2016 – 393 p.
5. Шолле Франсуа.Глубокое обучение на Python. — СПб.: Питер, 2018. — 400 с.: ил. — (Серия «Библиотека программиста»).
6. Липчак Д. А. Обзор методов автоматической диагностики сердечной аритмии для принятия решений о необходимости проведения дефибрилляции / Д. А. Липчак, А. А. Чупов // Ural Radio Engineering Journal. — 2021. — Vol. 5, No. 4. — P. 380–409.
7. A Deep Siamese Convolution Neural Network for Multi-Class Classification of Alzheimer Disease. Mehmood, A.; Maqsood, M.; Bashir, M.; Shuyuan, Y. Brain Sci. **2020**, 10, 84. <https://doi.org/10.3390/brainsci10020084>.
8. Detection of Atrial Fibrillation Using 1D Convolutional Neural Network. Hsieh, C.-H.; Li, Y.-S.; Hwang, B.-J.; Hsiao, C.-H. Sensors **2020**, 20, 2136. <https://doi.org/10.3390/s20072136>.
9. S. Hochreiter and J. Schmidhuber. “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.
10. PTB-XL, a large publicly available electrocardiography dataset / Wagner, P., Strodthoff, N., Bousseljot, RD. et al. Sci Data 7, 154 (2020). <https://doi.org/10.1038/s41597-020-0495-6>.
11. Официальный сайт с набором данных PTB-XL. Режим доступа: <https://physionet.org/content/ptb-xl/1.0.3/>.
12. Официальный сайт с набором данных ICBEB2018. Режим доступа: <http://2018.icbeb.org/Challenge.html>.
13. An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection. Feifei Liu, Chengyu Liu, Lina Zhao, Xiangyu Zhang, Xiaoling Wu, Xiaoyan Xu, Yulin Liu, Caiyun Ma, Shoushui Wei, Zhiqiang He, Jianqing Li, and Eddie Ng Yin Kwee. Journal of Medical Imaging and Health Informatics Vol. 8, 1368–1373, 2018.
14. Официальный сайт с документацией к открытой библиотеке «Scikit-learn». Режим доступа: <https://scikit-learn.org/stable/index.html>.
15. Официальный сайт с документацией к открытой библиотеке «TensorFlow». Режим доступа: <https://www.tensorflow.org/>.
16. Официальный сайт с документацией к открытой библиотеке keras. Режим доступа: <https://keras.io/>.
17. I. Loshchilov and F. Hutter. “Fixing weight decay regularization in ADAM,” in Proc. Int. Conf. Learn. Representations, 2019.
18. Официальный сайт ВОЗ, где приведена статистика по основным причинам смерти в мире на 2019 год. Режим доступа: <https://www.who.int/ru/news-room/fact-sheets/detail/the-top-10-causes-of-death>.
19. Ilias G. Maglogiannis, Kostas Karpouzis and Manolis Wallace. Image and Signal Processing for Networked E-Health Applications. 2006 by Morgan & Claypool, pp. 41-44.
20. K. Cho *et al.*, “Learning phrase representations using RNN encoder–decoder for statistical machine translation,” in *Proc.Conf.EmpiricalMethods Natural Lang. Process.*, Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734.
21. Официальный сайт языка программирования Python. Режим доступа: <https://www.python.org/>.