

ИССЛЕДОВАНИЕ И ПРИМЕНЕНИЕ АРХИТЕКТУР ГЛУБОКИХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ КЛАССИФИКАЦИИ НА МНОГОМЕРНЫХ ВРЕМЕННЫХ РЯДАХ

© 2022 г. А. С. Есенков^{a,*}, Е. М. Захарова^{b,**}, М. Д. Ковалева^{b,***},
Д. Е. Константинов^{b,****}, И. С. Макаров^{a,*****}, Е. А. Панковец^{b,*****}

^a ФИЦ ИУ РАН, Москва, Россия

^b МФТИ (национальный исследовательский ун-т), Долгопрудный, Моск. обл., Россия

*e-mail: esenkov@mail.ru

**e-mail: zakharova.em@mipt.ru

***e-mail: kovaleva.md@phystech.edu

****e-mail: konstantinov.de@phystech.edu

*****e-mail: i.s.m.mipt@yandex.r

*****e-mail: pankovets.ea@phystech.edu

Поступила в редакцию 07.03.2022 г.

После доработки 25.03.2022 г.

Принята к публикации 28.03.2022 г.

Приведены результаты исследования существующих архитектур глубоких нейронных сетей, предназначенных для решения задач классификации. В результате формируются атрибуты для эффективной автоматизации принятия решений. В качестве данных используются многомерные временные ряды финансовых рынков. Рассмотрены задачи бинарной и множественной классификации. Проанализированы полносвязные, рекуррентные (long short-term memory) и гибридные комбинированные архитектуры нейронных сетей. Изучаемый многомерный временной ряд получен путем объединения одномерных временных рядов стоимости актива, объема торговли, технических индикаторов и других параметров.

DOI: 10.31857/S0002338822040072

Введение. Основной тенденцией проведения исследований финансовых рынков, оценки и прогнозирования изменений на нем является использование технологии искусственного интеллекта. Одна из задач такого анализа — задача классификации данных. В статье описаны методы и алгоритмы, предназначенные для реализации системы поддержки принятия решений в области торговли акциями. Приведены основные этапы подготовки данных, выраженных в виде многомерных временных рядов финансовых рынков, для обучения и результаты исследования существующих архитектур глубоких нейронных сетей, предназначенных для решения задач классификации.

1. Постановка задачи. Изучению и прогнозированию финансовых рынков с использованием методов искусственного интеллекта посвящено много работ. В [1] представлен ансамбль независимых и параллельных нейронных сетей с долгой кратковременной памятью (long short-term memory, LSTM) для прогнозирования движения цены акций. Было показано, что LSTM особенно подходит для данных временных рядов из-за их способности включать прошлую информацию, в то время как было обнаружено, что ансамбли нейронных сетей уменьшают изменчивость результатов и улучшают обобщение. В [1] исследуется именно бинарная классификация.

Работа [2] представляет собой интеграцию современных методов обнаружения объектов и кодирования временных рядов GAF (gramian angular field) в задачах свечных паттернов. Предлагаемая модель, основанная на глубоких нейронных сетях и уникальном архитектурном дизайне, хорошо работает в классификации свечей и распознавании местоположения.

В [3] было проведено исследование применения методов компьютерного зрения к финансовым временным рядам с целью уменьшения воздействия шума и создания правильных меток. Результаты показывают, что созданные таким образом метки с шумоподавлением улучшают

производительность алгоритма обучения нисходящего потока как для небольших, так и для больших наборов данных.

Данная работа направлена на исследование гибридных архитектур нейронных сетей для классификации многомерных рядов. Для решения поставленной задачи были использованы глубокая нейронная сеть, состоящая из рекуррентных слоев LSTM-типа, полносвязная глубокая нейронная сеть и гибридные модели со сложной (разветвленной) топологией. При составлении выборок в дополнение к существующим данным были рассчитаны индикаторы и осцилляторы технического анализа.

Основная задача статьи — доказать, что подобранные в процессе исследования архитектуры позволяют получить высокую точность прогнозирования изменений на финансовых рынках и их можно применить для автоматизации принятия решения.

2. Подготовка и выбор данных обучающих выборок. Для обучения берутся данные, построенные на базе свечных графиков в дневном таймфрейме [4]. Свечой называется технический индикатор, применяемый главным образом для отображения изменений биржевых котировок акций, цен на сырье и т.д. Длительность временных рядов составляет от 1 до 20 лет, шаг равен одним суткам, т.е. всего было использовано 20×365 свечей. Графики в меньших таймфреймах не рассматриваются по причине сильной зашумленности данных. Прогноз строится на период от 1 до 5 дней от текущего момента. Были исследованы и обучены пять моделей, из которых была выбрана наилучшая.

Исходные данные временного ряда найдены из баз акционерного общества “Инвестиционный холдинг Финам” [5] и содержат в себе следующую информацию: дата и время, максимальная и минимальная цены, цены закрытия и открытия и объем торгов соответствующей свечи. Для получения данных с сервера было создано собственное программное обеспечение.

В малой обучающей выборке применялись свечные графики 37 компаний, входящих в индекс Московской биржи [6]. Всего туда входит 45 акций. Три компании, ЛКОН (публичное акционерное общество “Лукойл”), GAZP (публичное акционерное общество “Газпром”) и RSTI (публичное акционерное общество “Россети”), были исключены из обучающей выборки по причине их использования для тестирования. Причины их выбора для тестирования следующие.

На рис. 1 представлены акции ЛКОН, содержащие глобальный растущий тренд, где по горизонтали — время, по вертикали — стоимость. График приведен с масштабом по месяцам для наглядности. Столбчатая диаграмма — объемы торговли для каждой свечи. GAZP содержит флэт, т.е. цена на их акции колеблется в определенном диапазоне без четко определенного направления. RSTI содержит нисходящую длительную тенденцию. Активы данных компаний позволяют проводить тестирование на различных по характеру данных. Тестовая выборка представлена за полгода (с 1 января 2021 г. по 29 июля 2021 г.) и не пересекается по времени с обучающей, чтобы не было наложения временных рядов.

Была подготовлена большая обучающая выборка, включающая в себя акции компаний, которые входят как в индекс Московской биржи, так и в американский индекс S&P-500 [7].

В некоторых акциях присутствуют достаточно сильные смены направлений. Подобные сильные колебания наблюдаются преимущественно в начале временного ряда, поэтому была произведена обрезка начальных фрагментов всех временных рядов, содержащих сильные колебания.

3. Подготовка данных. Для проведения исследования полученные исходные данные свечных графиков были подвергнуты ряду преобразований. Был исключен год из данных свечи, так как этот элемент данных не несет полезной информации для прогнозирования временного ряда в будущем.

Для номера месяца были рассмотрены два варианта нормировки. Суть первого способа заключается в делении на 12, второго — в использовании one-hot encoding (ОНЕ). Для нормировки номера дня применялось деление на 31, в результате которого значения были равномерно распределены в диапазоне (0,1). По данным полной даты для каждого дня был вычислен номер дня недели, который был закодирован с помощью ОНЕ.

Использование дня месяца и дня недели оправдано тем, что в течение месяца и в течение недели наблюдаются определенные тенденции, например сокращение позиций в пятницу или в конце месяца. В частности, было проведено исследование поведения акций компаний, входящих в индекс Московской биржи в пятницу по всем имеющимся данным.

В табл. 1 представлены результаты анализа, которые показывают, что акции компаний банковского сектора имеют тенденцию к снижению по пятницам, а акции компаний нефтегазового сектора наоборот — тенденцию к росту.



Рис. 1. Глобально растущий тренд акций ПАО «Лукойл»

Данные о часе, минуте и секунде свечи были полностью исключены, так как минимальное разбиение свечей составляет одни сутки. Включение в обучающую выборку одновременно данных о месяце, дне и дне недели имеет негативный эффект для обучения модели, так как по ним достаточно просто идентифицировать конкретный фрагмент временного ряда. Модель обучается распознавать не паттерны других компонент временного ряда, а просто запоминает дату.

Информация о месячных сезонных зависимостях также не была включена в итоговую обучающую выборку, так как в системе поддержки принятия решений основной акцент ставится на удержание позиции в течение короткого периода времени (1–10 дней), при этом была включена информация о дне недели в ONE-формате.

Таблица 1. Акции компаний банковского и нефтегазового секторов

Компания	Количество пятниц, в которые наблюдался рост	Количество пятниц, в которые наблюдалось падение
Лукойл	517	475
Норникель	503	473
Сбербанк	308	363
Газпром	379	392
Роснефть	371	376
Банк ВТБ	301	377

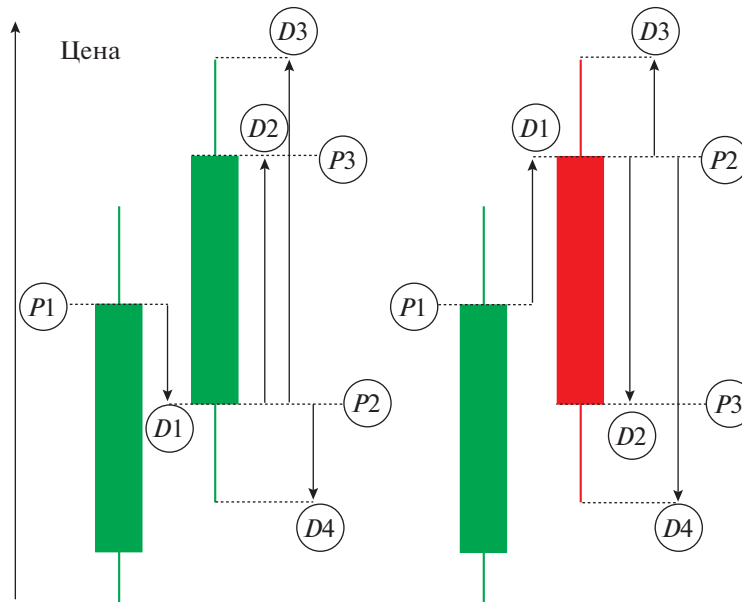


Рис. 2. Процентные изменения актива относительно некоторого значения цен

Использование необработанных значений стоимости актива в обучающей выборке не представляется возможным, так как различные активы имеют разную стоимость, поэтому было выполнено преобразование сырых значений стоимости актива в процентное изменение.

На рис. 2 представлены процентные изменения актива относительно некоторого значения цены. Зеленый цвет обозначает подъем цены, красный — ее падение. Стрелками обозначены величины процентных изменений, которые вычисляются как:

$$D_1 = \frac{P_2 - P_1}{P_1}, \quad D_2 = \frac{P_3 - P_2}{P_2}, \quad D_3 = \frac{P_{\max} - P_2}{P_2}, \quad D_4 = \frac{P_2 - P_{\min}}{P_{\min}}, \quad (3.1)$$

где P_1 — цена закрытия предыдущего дня, P_2 — цена открытия текущего дня, P_3 — цена закрытия текущего дня, P_{\min} и P_{\max} — минимальная и максимальная цены текущего дня соответственно, D_1 , D_2 , D_3 , D_4 — величины процентных изменений. При этом D_3 и D_4 больше нуля, D , D_1 , D_2 — любые по знаку.

Из формул (3.1) следует что D — процентное изменение актива, вычисляется по следующей формуле:

$$D = \frac{P_3 - P_1}{P_1}. \quad (3.2)$$

Было проведено исследование процентных изменений стоимости актива. На рис. 3 представлены результаты анализа, который показал, что значение D по данным всех активов имеет вид нормального распределения, матожидание которого составляет 0.

Пусть σ представляет собой стандартное отклонение величины, а μ — ее матожидание. Тогда по правилу трех сигм [8] 3σ -диапазон представляет собой интервал $[\mu - 3\sigma; \mu + 3\sigma]$. В работе значения 3σ -диапазона лежат в диапазоне $[-0.1; 0.1]$, т.е. процентное изменение стоимости актива в день составляет не более 10%.

Значения D для последних N свечей были включены в обучающую выборку. При этом каждое значение было умножено на 10. Таким образом было достигнуто нормальное распределение в диапазоне $[-1; 1]$, а значения вне этого диапазона были обрезаны до -1 и $+1$ соответственно.

Помимо информации о стоимости актива также использовалась информация об объеме торговли $V_{\text{deviation}}$. Данные по объемам распределялись нормально, и при составлении векторов обучающей выборки применялись процентные изменения объемов. Характерное процентное изменение объема составляло 50%, поэтому исходное нормальное распределение не подвергалось

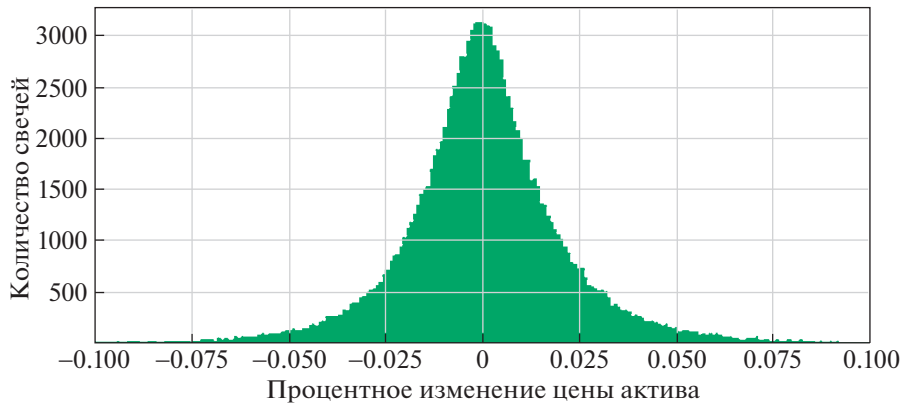


Рис. 3. Распределение процентного изменения актива

масштабированию. Аналогично с преобразованиями процентных изменений актива D выполняется обрезка процентных изменений объема $V_{\text{deviation}}$: значения, выходящие из диапазона от -1 до $+1$, приводились к значениям -1 и $+1$ соответственно.

В процессе анализа и прогнозирования временных рядов финансовых рынков часто используются уровни поддержки и сопротивления. В работе рассматриваются только горизонтальные уровни, определяемые по локальным экстремумам.

На рис. 1 видно, что уровень сопротивления может превратиться в уровень поддержки после преодоления его ценой, т.е. один и тот же уровень может проходить как через локальные минимумы, так и через локальные максимумы. Определение уровней осуществляется с помощью алгоритма детектирования локальных экстремумов посредством сдвига окна-фильтра фиксированной ширины (ширина измеряется в свечах). Каждый уровень характеризуется двумя параметрами. Первый – его время жизни, которое измеряется в свечах и отсчитывается от первого локального экстремума, проходило через данный уровень. Второй – сила уровня, соответствующая числу экстремумов, через которые данный уровень проходит в пределах погрешности. Максимальное время жизни уровня ограничено сверху, уровни, превышающие данное ограничение, считаются устаревшими и не учитываются при анализе текущей ситуации. Нормировка силы уровня S и времени жизни уровня L выполняется посредством деления данных величин на максимальную силу уровня и максимальное время жизни уровня, которые заданы как константы алгоритма нормализации данных:

$$S_{\text{level}}^{(norm)} = \min\left(1.0; \frac{S_{\text{level}}}{S_{\text{max}}}\right), \quad L_{\text{level}}^{(norm)} = \min\left(1.0; \frac{L_{\text{level}}}{L_{\text{max}}}\right), \quad (3.3)$$

где $S_{\text{level}}^{(norm)}$ – нормированное значение силы уровня, $L_{\text{level}}^{(norm)}$ – нормированное значение времени жизни уровня, S_{level} и L_{level} – ненормированные значения силы и времени жизни уровня, S_{max} и L_{max} – их заданные максимальные значения.

Для каждой свечи каждого графика вычисляется ближайший уровень поддержки (снизу) и сопротивления (сверху). Для определенных свечей уровни могут отсутствовать. В таком случае все параметры приравниваются к нулю. Такая ситуация чаще всего возникает для свечей в начале графика, а также для свечей ATH (all time high) и ATL (all time low). Для каждой свечи вычисляются расстояния до ближайшего уровня сопротивления и до ближайшего уровня поддержки, которое измеряется в единицах стоимости актива:

$$D_{\text{level}} = \frac{P_{\text{level}} - P_{\text{close}}}{P_{\text{close}}}, \quad (3.4)$$

где D_{level} – расстояние, P_{close} – цена закрытия свечи, P_{level} – цена уровня.

Полученные процентные отклонения также имеют нормальное распределение, значения 3σ -диапазона которого лежат на отрезке $[-0.1; 0.1]$. Все процентные отклонения берутся по модулю.

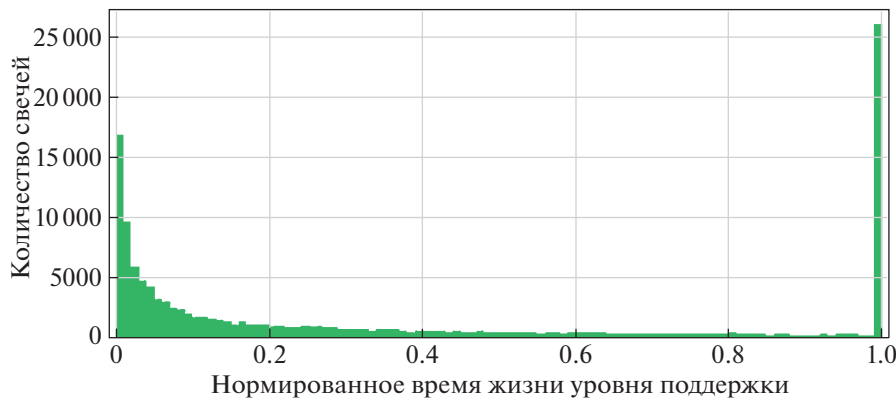


Рис. 4. Экспоненциальное распределение времени жизни уровней для всех свечей всех свечных графиков

В конечном итоге полная информация об уровнях для каждой свечи характеризуется шестью числами: цена, время жизни и сила для двух уровней этой свечи. Однако в ходе экспериментов было установлено, что наилучшая результативность обучения достигается при указании информации об уровнях только последней текущей свечи. Информация о силе уровня является нерепрезентативной, так как существуют сильные уровни, проходящие через один локальный экстремум, и слабые уровни, проходящие через множество локальных экстремумов.

К тому же, как было сказано, определение уровней имеет некоторую погрешность. В статье ее значение составляет 0.25%. Поэтому в итоговом рабочем варианте данных об уровнях была оставлена только информация о наличии или отсутствии любого уровня в окрестности цены закрытия текущей свечи с указанной погрешностью. Это обусловлено тем, что в реальности часто происходят ложные пробои уровней. В таких ситуациях уровень поддержки может трактоваться как уровень сопротивления, что не является верным. Будет правильнее предоставить модели в процессе обучения самостоятельно идентифицировать тип уровня при его наличии.

На рис. 4 представлено экспоненциальное распределение времени жизни уровней для всех свечей всех свечных графиков. Таким образом, описание уровней для текущей свечи состоит только из одного параметра, который равен 0, если уровень отсутствует, или имеет ненулевое значение в противном случае. При этом ненулевое значение параметра экспоненциально распределено в диапазоне $[0, 1]$ и соответствует времени жизни уровня.

Распределение имеет экспоненциальный вид, выброс справа обозначает уровни, время жизни которых превышает установленный лимит (в данной работе — 5 лет). При подготовке обучающей выборки подобные уровни игнорируются. При составлении выборок наравне с перечисленными выше первичными данными были использованы вторичные данные, полученные посредством расчета индикаторов и осцилляторов технического анализа [9]. Применение индикаторов приводит к увеличению точности на 1–2% [10].

Однако в выборку данных для обучения моделей следует включать сырые значения индикаторов и осцилляторов. В процессе обучения модель самостоятельно определит значения, которые могут выступать в качестве полезных сигналов. В дополнение к этому на графике стоимости акций некоторых компаний могут возникать разрывы. Наибольшие разрывы можно встретить на графиках стоимости акций тех компаний, которые выплачивают высокие дивиденды. Все индикаторы и осцилляторы используют предысторию заданной продолжительности цен закрытия или типичных цен. После возникновения (дивидендного) разрыва значения индикаторов и осцилляторов могут претерпевать резкие изменения, поэтому следует исключить из обучающей и тестовых выборок содержащие их сегменты временных рядов. Обученную модель также не рекомендуется применять в течение t дней после возникновения дивидендного разрыва, где t — максимальная длительность приведенной предыстории по всем индикаторам и осцилляторам.

4. Построение векторов обучающей выборки. В работе представлено два варианта структуры входного вектора обучающей и тестовых выборок.

В табл. 2 описан вектор первого типа размерности $10 + N + M$. Вектор второго типа имеет размерность $k \times N$, где $k = 30$ и состоит из 12 индикаторов + 16 осцилляторов + цены закрытия и объемы торговли; N и M приведены в табл. 2.

Таблица 2. Описание вектора 1 размерности $10 + N + M$

Количество параметров	Формат параметров	Описание
5	ОНЕ	День недели
1	Экспоненциально распределенное число в диапазоне $[0.0; 1.0]$	Время жизни уровня последней свечи (0.0 означает, что уровня нет, либо он был проигнорирован из-за превышения лимита времени жизни)
M	Нормально распределенных чисел в диапазоне $[-1.0; 1.0]$	Процентное изменение объема для M последних свечей ($V_{Deviation}$)
N	Нормально распределенных чисел в диапазоне $[-1.0; 1.0]$	Процентное изменение объема для N последних свечей (PD)
1	Нормально распределенное число $[0.0; 1.0]$	Отклонение D_3 последней свечи
1	Нормально распределенное число $[0.0; 1.0]$	Отклонение D_4 последней свечи
1	Нормально распределенное число $[-1.0; 1.0]$	Отклонение D_1 последней свечи
1	Нормально распределенное число $[-1.0; 1.0]$	Отклонение D_2 последней свечи

В ходе экспериментов было установлено, что наиболее оптимальное значение $N = 60$, что соответствует анализируемой предыстории за 3 месяца (принято, что в неделе 5 рабочих дней, месяц состоит из 4 недель). Оптимальное значение $M = 10$, что соответствует 2 неделям.

5. Разметка. Было подготовлено несколько типов разметки для решения задач классификации разного рода. Первый вариант простейшей разметки соответствует задаче прогнозирования движения цены в последующие N_{cur} дней от текущего момента, N_{cur} лежит в диапазоне $[1, 5]$. Вторым вариантом разметки был получен с помощью алгоритма детектирования тренда с плавающей правой границей окна-фильтра. В ходе экспериментов было установлено, что значение точности лучше на разметке, найденной вторым способом.

Далее в работе приводятся результаты обучения и тестирования моделей по второй разметке, потому что точность на разметке первого типа составляет 54–58%, на разметке второго типа – 65–70%.

Исходный вариант разметки вторым способом подразумевает три класса: L , C , S . Пометка ставится для каждой свечи. Пометка C означает, что на момент закрытия текущей свечи рекомендуется иметь нулевую позицию, L – рекомендуется иметь длинную позицию, S – рекомендуется иметь короткую позицию. Метка C соответствует моментам, когда на рынке отсутствует определенное движение. В такие моменты можно либо продолжать удерживать имеющуюся позицию до появления более четкого сигнала, либо установить нулевую позицию.

В табл. 3 представлены используемые в данной работе метки и их распределение.

Как видно из табл. 3 метки типа C составляют около 10% от общего количества меток всех типов. При этом в ходе экспериментов было установлено, что модели практически не способны корректно распознавать данную метку после обучения. Это является следствием несбалансированности состава выборки. Поэтому данная трехклассовая разметка была упрощена до двухклассовой, а все векторы, помеченные меткой C , были удалены из обучающей и тестовых выборок.

6. Модели нейронных сетей. В процессе исследования были изучены модели трех типов архитектур. Обучение моделей первого и второго типа архитектур проводилось на малой обучающей выборке, содержащей данные по 37 акциям компаний, входящих в индекс Московской биржи [6]. В качестве результатов тестирования модели приводится значение точности предсказания меток на размеченных тестовых данных акций Лукойла. Для других тестовых данных отклонения значений ассигасу лежит в диапазоне $+/-2\%$.

Первый тип – полносвязные глубокие нейронные сети. В табл. 4 представлена наилучшая по результатам анализа архитектура. Модель состоит из четырех полносвязных слоев с выпадением с коэффициентом $dropout = 0.25$ и функцией активации типа $relu$. Последний пятый слой содержит один нейрон с функцией активации типа сигмоид. В качестве входного вектора используется вектор первого типа (табл. 3).

Таблица 3. Таблица меток

Распределение меток	Количество меток	Содержание меток, %
Распределение меток тестовой выборки, предложенных одной из обученных моделей	C: 67	C: 0.53
	L: 7147	L: 56.31
	S: 5478	S: 43.16
Распределение меток в тестовой выборке, предоставленных алгоритмом разметки	C: 1350	C: 10.64
	L: 5960	L: 46.96
	S: 5382	S: 42.40
Распределение меток в малой обучающей выборке, предоставленных алгоритмом разметки	C: 12561	C: 10
	L: 56811	L: 46
	S: 51569	S: 42
Для отдельных активов		
Название компании на бирже, таймфрейм, количество свечей	Количество меток	Содержание меток, %
Объединённая компания “РУСАЛ”, дневной, 1390	C: 131	C: 9
	L: 628	L: 45
	S: 631	S: 45
Группа компаний “Петропавловск”, дневной, 74	C: 17	C: 22
	L: 15	L: 20
	S: 42	S: 56
X5 Retail Group, дневной, 677	C: 99	C: 14
	L: 301	L: 44
	S: 277	S: 40
Магнитогорский металлургический комбинат, дневной, 3686	C: 342	C: 9
	L: 1719	L: 46
	S: 1625	S: 44

Таблица 4. Архитектура полносвязной глубокой нейронной сети

Слой	Тип	Структура выходного тензора	Количество параметров
Dense_1	Dense	(None, 80)	6480
Dropout_1	Dropout	(None, 80)	0
Dense_2	Dense	(None, 320)	25920
Dropout_2	Dropout	(None, 320)	0
Dense_3	Dense	(None, 320)	102720
Dropout_3	Dropout	(None, 320)	0
Dense_4	Dense	(None, 80)	25680
Dense_5	Dense	(None, 2)	162

Наилучшие результаты были получены с помощью полносвязных моделей, количество нейронов в слоях которых в 2 или 4 раза превышает количество параметров входного вектора. Общее количество параметров модели равно 160962.

На рис. 5 рассмотрены графики изменения точности от эпохи обучения для обучающей и тестовых выборок. Заметим, что с определенной эпохи начинается переобучение. Точность для ЛКОН за все время составила 0.6818, а за последние полгода – 0.6759.

Второй тип моделей – гибридные модели со сложной (разветвленной) топологией. Модели данного типа также используют в качестве входного вектора первого типа (табл. 3), но разбитого на три части. Первая часть включает в себя процентные изменения цен закрытия (60 чисел,

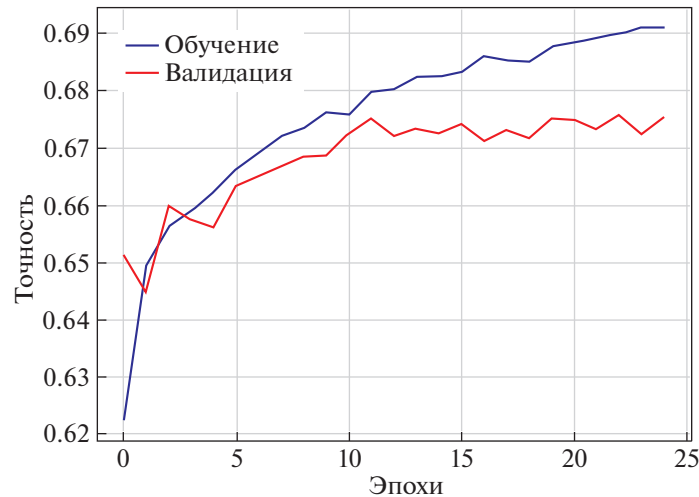


Рис. 5. График зависимости точности от эпохи для лучшей dense-модели

образующих временной ряд), вторая – процентные изменения объемов торговли (10 чисел, образующих временной ряд), третья часть – оставшиеся параметры входного вектора (10 чисел). Первая и вторая части подаются на вход в соответствующие рекуррентные слои типа LSTM, третья часть – на вход полносвязного слоя. Внутри архитектуры выполняется конкатенация трех разделенных слоев. Последний слой содержит один нейрон с функцией активации типа сигмоид. Для всех нейронов, кроме последнего, установлена функция активации типа relu, а для компонент LSTM-слоев выбран вариант по умолчанию из реализации keras.

По результатам анализа наилучшая архитектура из них представлена в табл. 5.

Общее количество параметров модели равно 113841. На рис. 6 рассмотрены графики изменения точности от эпохи обучения для обучающей и тестовых выборок. Заметим, что с эпохи x начинается переобучение. Assurance для ЛКОН за все время составило 0.6651, а за последние полгода – 0.6759.

Таблица 5. Архитектура гибридной модели со сложной (разветвленной) топологией

Слой	Структура выходного тензора	Количество параметров	Входной слой для текущего
Input_V (InputLayer)	(None, 10, 1)	0	
Lstm_25 (LSTM)	(None, 10, 50)	10400	Input_V[0][0]
Input_P (InputLayer)	(None, 60, 1)	0	
Lstm_26 (LSTM)	(None, 10, 50)	20200	Lstm_25[0][0]
Lstm_29 (LSTM)	(None, 60, 40)	6720	Input_P[0][0]
Lstm_27 (LSTM)	(None, 10, 50)	20200	Lstm_26[0][0]
Lstm_30 (LSTM)	(None, 60, 40)	12960	Lstm_29[0][0]
Input_D (InputLayer)	(None, 10)	0	
Lstm_28 (LSTM)	(None, 50)	20200	Lstm_27[0][0]
Lstm_31 (LSTM)	(None, 40)	12960	Lstm_30[0][0]
Concatenate_4 (Concatenate)	(None, 100)	0s	Input_D[0][0] Lstm_28[0][0] Lstm_31[0][0]
Dense_4 (Dense)	(None, 100)	10100	Concatenate_4[0][0]
Output (Dense)	(None, 1)	101	Dense_4[0][0]

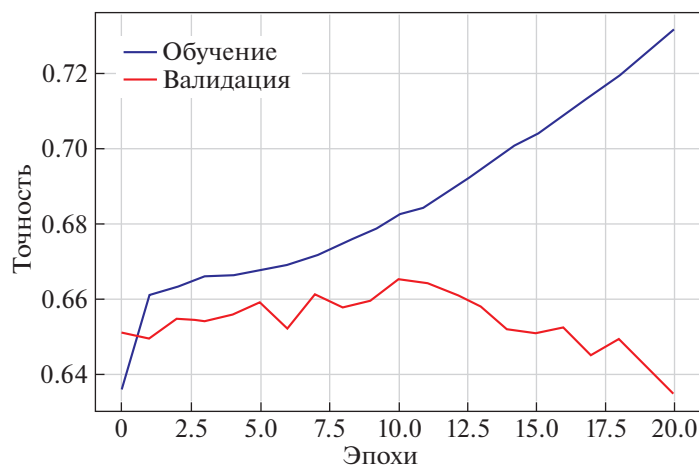


Рис. 6. Графики изменения точности от эпохи обучения для лучшей гибридной модели

Третий тип моделей – глубокая нейронная сеть, состоящая из рекуррентных слоев LSTM-типа. Модели данного типа используют в качестве входного вектора второго типа (табл. 2). Модели анализируемой архитектуры показывают наилучшие результаты среди всех исследуемых в работе моделей, поэтому точная конфигурация архитектуры не приводится. Отметим, что архитектура включает в себя последовательность LSTM-слоев и полносвязных слоев с выпадением. После получения предварительных результатов на малой обучающей выборке модель была переобучена на данных большой обучающей выборки. Точность для ЛКОН за все время составила 0.6914, а за последние полгода – 0.7269. На отдельных активах тест за последние полгода показывал еще более высокие результаты, например, на акции GAZP ассигасу составила 0.7830. Графики зависимости ассигасы от эпохи для train и test не приводятся, так как наилучшие значения были достигнуты уже после первой эпохи по причине большого размера обучающей выборки.

Заключение. Описаны результаты подготовки и разметки данных, а также исследования различных архитектур глубоких нейронных сетей, предназначенных для задач классификации на размеченных данных многомерного временного ряда, включающего в себя исходные данные свечных графиков и производные данные, вычисленные на основе исходных. Наилучшие результаты были достигнуты с помощью многослойной нейронной сети, состоящей из последовательности рекуррентных слоев LSTM-типа и полносвязных слоев с выпадением. Полученная модель может быть применена в системах автоматической торговли, функционирующих по среднесрочной стратегии.

СПИСОК ЛИТЕРАТУРЫ

1. *Fjellström C.* Long Short-Term Memory Neural Network for Financial Time Series // arXiv preprint arXiv:2201.08218. 2022.
2. *Chen J.H., Tsai Y.C.* Dynamic Deep Convolutional Candlestick Learner // arXiv preprint arXiv:2201.08669. 2022.
3. *Ma Y., Ventre C., Polukarov M.* Denoised Labels for Financial Time-Series Data via Self-Supervised Learning // arXiv preprint arXiv:2112.10139. 2021.
4. *Нисон С.* Японские свечи. Графический анализ финансовых рынков. М.: Альпина Паблишер, 2017.
5. <https://www.finam.ru/>
6. <https://www.moex.com/>
7. <https://www.spglobal.com/ratings/en/>
8. *Пискунов Н.С.* Дифференциальное и интегральное исчисления: для вузов. М.: Наука, 1985.
9. *Мэрфи Д.Д.* Технический анализ фьючерсных рынков: теория и практика. М.: Альпина Паблишер, 1996.
10. *Авдеев А.В.* Сравнение видов индикаторов технического анализа и выбор оптимальной категории индикаторов для дальнейшего использования в алгоритмах системы торговли на бирже. Изв. Тульск. гос. ун-та. Экономические и юридические науки. 2019. №. 2. С. 12–18.