

УДК 338.1

DOI 10.52452/18115942_2024_1_30

КЛАССИФИКАЦИЯ И ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ПРОГНОЗИРОВАНИЯ ФИНАНСОВЫХ ПОКАЗАТЕЛЕЙ

© 2024 г.

А.И. Полазнов, Д.А. Корнилов

Полазнов Алексей Игоревич,
магистр Нижегородского государственного технического университета им. Р.Е. Алексева;
техник отдела научно-технического развития ОАО «Атомэнергопроект»
witlex21@gmail.com

Корнилов Дмитрий Анатольевич, д.э.н.; проф.;
профессор кафедры управления инновационной деятельностью
Нижегородского государственного технического университета им. Р.Е. Алексева
dkornilov2020@yandex.ru

*Статья поступила в редакцию 23.12.2023**Статья принята к публикации 29.01.2024*

Данная статья представляет собой исследование методов построения нейронных сетей для прогнозирования финансовых показателей, в частности для прогнозирования курсов валют. Актуальность данного исследования обусловлена необходимостью разработки эффективных моделей прогнозирования курсов валют, акций, криптовалют и других финансовых показателей. Цель исследования – сделать обзор нейронных сетей, сформировать их классификацию, выбрать оптимальные методы и предложить модели применения нейронных сетей для прогнозирования финансовых показателей, в том числе курсов валют. В рамках работы были рассмотрены различные архитектуры нейронных сетей, включая простые рекуррентные нейронные сети (RNN), сверточные нейронные сети (CNN), глубокие нейронные сети (DNN) и другие. Более 40 методов было классифицировано и представлено их описание. Для каждого метода проанализированы их преимущества и недостатки с точки зрения прогнозирования курса. Результаты исследования показали, что выбор оптимального метода для прогнозирования курса зависит от характеристик конкретной задачи. Например, RNN показали себя лучше в задачах с временными рядами, в то время как CNN оказались эффективными в задачах с пространственной структурой. Рассмотрены вопросы повышения точности предсказания и недопущения переобучения нейронной сети. Составлена блок-схема SA-алгоритма, которая позволяет наглядно представить последовательность шагов и операций. Эта работа может быть полезным руководством для исследователей и практиков, занимающихся прогнозированием курсов финансовых активов. Результаты работы могут способствовать более точной и надежной разработке моделей прогнозирования, что имеет большое значение для финансового сектора и других областей, где важно предсказание курсов валют и повышение эффективности финансового планирования.

Ключевые слова: искусственный интеллект, нейронные сети, прогнозирование, классификация, процессная модель, прогнозирование, курс валют, курс рубля, методы.

Введение

Исследование методов построения нейросетей для прогнозирования является актуальным по нескольким причинам.

Во-первых, прогнозирование является важной задачей во многих областях, включая финансовые рынки, экономику, климатологию, медицину и т.д. Точные прогнозы позволяют принимать более обоснованные решения и снизить потери.

Во-вторых, нейронные сети являются мощным инструментом для прогнозирования благодаря своей способности обучаться на данных и находить сложные взаимосвязи между входными и выходными переменными. Однако существует множество различных архитектур и методов обучения нейросетей, и выбор наиболее

подходящего метода для конкретной задачи является сложной темой.

Также нейросети часто используются для прогнозирования в условиях нестационарности данных, то есть когда статистические свойства данных изменяются со временем. Нестационарность представляет особую сложность для нейросетей, так как требуется обработка и изменение архитектуры модели по мере изменения статистических свойств данных.

И, наконец, разработка новых методов построения нейросетей для прогнозирования может привести к улучшению качества прогнозов и повышению эффективности работы моделей. Исследователи постоянно ищут новые способы улучшить производительность нейросетей, уменьшить их сложность, ускорить время обучения и так далее [1–4].



Рис. 1. Методы построения нейросетей
Составлено авторами на основе данных [5–9]

Методы построения нейросетей и их классификация

Методы построения нейросетей представлены на рисунке 1, а далее дается их краткое описание с целью дальнейшей классификации и разделения на группы (рис. 2).

1. Метод обратного распространения ошибки (Backpropagation) – метод обучения нейронной сети, основанный на вычислении и распространении ошибки обратно по сети.

2. Градиентный спуск (Gradient Descent) – метод оптимизации, используемый для минимизации функции путем поиска направления наискорейшего убывания.

3. Стохастический градиентный спуск (Stochastic Gradient Descent) – модификация градиентного спуска, при которой обновление весов происходит для каждого примера обучающего набора.

4. Метод адаптивного градиентного спуска (Adaptive Gradient Descent) – метод оптимизации, который адаптирует скорость обучения для каждого параметра в зависимости от его исторического градиента.

5. Методы оптимизации второго порядка (Second-order optimization methods) – методы оптимизации, которые учитывают информацию о вторых производных функции.

6. Методы регуляризации (Regularization methods) – методы, используемые для предотвращения переобучения путем внесения штрафа на сложность модели.

7. Конволюционные нейронные сети (Convolutional Neural Networks) – архитектура нейронных сетей, специально разработанная для обработки и анализа данных с пространственной или временной структурой, таких как изображения или звуковые файлы.

8. Рекуррентные нейронные сети (Recurrent Neural Networks) – архитектура нейронных сетей, способная моделировать последовательности данных и сохранять информацию о предыдущих состояниях.

9. Глубокое обучение (Deep Learning) – подраздел машинного обучения, связанный с обучением и использованием глубоких нейронных сетей.

10. Автоэнкодеры (Autoencoders) – нейронные сети, которые используются для обучения



Рис. 2. Группировка методов по признакам
Составлено авторами на основе данных [10–14]

компактного представления данных через обучение без учителя.

11. Метод свертки (Pooling) – метод, используемый в сверточных нейронных сетях для уменьшения размерности выходного показателя на основе весовых коэффициентов ядра свертки, которые определяются в процессе обучения.

12. Прямоугольные сверточные сети (Rectangular Convolutional Networks) – реализация сверточной нейронной сети, которая позволяет использовать прямоугольные фильтры для анализа изображений или данных в двухмерном пространстве.

13. Сети с операциями депуляции (Sparse Networks) – нейронные сети, в которых используются разреженные связи между нейронами, что позволяет снизить вычислительную сложность и улучшить обобщающую способность модели.

14. Методы обучения без учителя (Unsupervised Learning methods) – методы машинного обучения, где модель обучается на размеченных данных без явно заданной цели.

15. Методы обучения с подкреплением (Reinforcement Learning methods) – методы машинного обучения, где модель обучается взаимодействуя с окружающей средой и получая обратную связь в виде награды или штрафа.

16. Генеративные нейронные сети (Generative Neural Networks) – нейронные сети, которые ис-

пользуются для генерации новых данных, таких как изображения, звуки или тексты, на основе имеющегося обучающего набора.

17. Сети с долгой краткосрочной памятью (Long Short-Term Memory Networks) – вид рекуррентных нейронных сетей, которые способны обрабатывать последовательности данных с долгосрочной зависимостью между элементами последовательности.

18. Сети Гамма (Gamma Networks) – нейронные сети, которые используются для моделирования и анализа гамма-распределений, которые часто встречаются в статистике и вероятностном моделировании.

19. Бинарные нейронные сети (Binary Neural Networks) – нейронные сети, в которых используются бинарные значения (–1 и 1) вместо стандартных вещественных чисел, что позволяет снизить требования к вычислительным ресурсам и ускорить обучение.

20. Метод Гиббса (Gibbs Sampling) – метод, используемый для сэмплирования из многомерных вероятностных распределений, основанный на итеративном обновлении переменных с использованием условных распределений. Часто используется в байесовском машинном обучении.

21. Рекурсивные нейронные сети (Recursive Neural Networks) – архитектуры нейронных сетей, в которых нейроны могут иметь связи с

предыдущими входными данными. Они позволяют обрабатывать данные, имеющие структуру дерева или графа, такие как естественный язык или изображения.

22. Сети с модульными функциями активации (Networks with modular activation functions) – нейронные сети, в которых для каждого слоя или модуля используется отдельная функция активации. Это позволяет настраивать каждый модуль независимо и улучшает обобщающие способности сети.

23. Генетические алгоритмы (Genetic algorithms) – эволюционные алгоритмы, основанные на принципах естественного отбора и генетики. Они используют случайный поиск и последующий выбор лучших вариантов для решения оптимизационных задач.

24. Кооперативные нейронные сети (Cooperative neural networks) – нейросети, работающие в группах или командах для выполнения сложных задач совместно. Каждая нейронная сеть может иметь свою собственную функцию активации, и они обмениваются информацией и сотрудничают друг с другом для достижения общей цели.

25. Сети с отражательными функциями активации (Networks with reflective activation functions) – нейронные сети, в которых используются отражательные функции активации, такие как ReLU (Rectified Linear Unit). Они помогают устранить проблему затухающего градиента и позволяют сети обучаться быстрее.

26. Сети с функциями активации типа Штивелле (Networks with Stivelle-type activation functions) – нейронные сети, в которых используется функция активации типа Штивелле, которая имеет более широкий диапазон аргументов, чем сигмоидная функция активации. Это позволяет сети представлять более сложные зависимости в данных.

27. Ансамбль нейронных сетей (Ensemble of neural networks) – метод, в котором несколько нейронных сетей объединяются вместе для решения задачи. Каждая сеть вносит свой вклад в решение, и их предсказания объединяются для получения более точного результата.

28. Алгоритмы оптимизации на основе искусственных пчел (Optimization algorithms based on artificial bees) – методы оптимизации, вдохновленные поведением пчел в природе. Они используют идеи поиска пищи и коммуникации в пчелиных колониях для эффективного поиска оптимального решения.

29. Нейровизуализация (Neuroimaging) – процесс визуализации результатов обработки данных нейронной сетью. Он позволяет визуально представить и интерпретировать резуль-

таты работы сети, например показать активации нейронов или веса связей между нейронами.

30. Методы переноса обучения (Learning transfer methods) – методы, позволяющие использовать знания и навыки, полученные в ходе обучения на одной задаче, для решения другой, связанной задачи. Это позволяет сокращать время и объем данных, необходимых для обучения новых моделей.

31. Автоматическое определение архитектуры нейронной сети (Automatic determination of neural network architecture) – методы, которые позволяют автоматически находить оптимальную архитектуру нейронной сети для решения конкретной задачи. Эти методы основаны на использовании алгоритмов оптимизации и поиска, таких как генетические алгоритмы или алгоритмы глубокого обучения, которые изменяют архитектуру сети, добавляя или удаляя слои или модифицируя размеры слоев, чтобы достичь лучших результатов.

32. Методы масштабирования данных (Data scaling methods) – методы, которые используются для изменения масштаба значений данных перед их подачей на вход нейронной сети. Популярными методами масштабирования данных являются нормализация, стандартизация и масштабирование в диапазоне. Эти методы позволяют улучшить производительность и сходимость нейронной сети, так как они помогают обеспечить более однородное распределение значений входных данных.

33. Сети глубокого сжатия (Deep compression networks) – методы, которые позволяют уменьшить размер и объем нейронной сети, не сильно снижая ее производительность. Они используются для сжатия и упрощения архитектуры сети, чтобы она занимала меньше памяти и быстрее выполнялась. Одним из популярных методов глубокого сжатия является метод прунинга, который удаляет ненужные или слабые связи между нейронами в сети.

34. Методы улучшения обобщающей способности нейронных сетей (Methods of improvement of generalization ability of neural networks) – методы, которые позволяют увеличить способность нейронной сети обобщать и делать точные прогнозы на новых данных. К таким методам относятся регуляризация (например, L1- или L2-регуляризация), добавление шума во время обучения, использование аугментации данных, признаковое извлечение и др.

35. Сети с функциями активации типа Гаусса (Networks with Gauss-type activation functions) – сети, которые используют функции активации, основанные на гауссовом распределении, например радиальные базисные функции (RBF).

Функции активации типа Гаусса позволяют нейронной сети моделировать нелинейные зависимости с помощью гладких кривых и гауссовых пиков.

36. Методы объединения множества моделей (Methods of combining multiple models) – методы, которые используют несколько моделей для решения задачи и объединяют их прогнозы или веса для достижения лучших результатов. Такие методы могут включать использование ансамблей нейронных сетей, стекинг, бэггинг и другие подходы к комбинированию моделей.

37. Сети прямого распространения (Direct propagation networks) – самый простой и стандартный тип нейронной сети, где сигнал передается только в одном направлении – от входных слоев к выходным. Они состоят из последовательного соединения нескольких слоев, где каждый слой состоит из нейронов, связанных взвешенными связями. Сети прямого распространения широко используются для решения различных задач, таких как классификация, регрессия и обработка текста.

38. Иерархические нейронные сети (Hierarchical neural networks) – сети, которые организованы в иерархическую структуру с несколькими уровнями. Каждый уровень нейронной сети обрабатывает данные на разных уровнях абстракции или разрешения, что позволяет сети выявлять сложные зависимости в данных. Иерархические нейронные сети широко используются в обработке изображений и распознавании образов.

39. Сети с функциями активации типа сигмоида (Networks with sigmoid-type activation functions) – сети, которые используют сигмоидальные функции активации, такие как логистическая функция или гиперболический тангенс. Эти функции активации преобразуют взвешенные суммы сигналов в ограниченный диапазон значений (обычно от 0 до 1 или от –1 до 1), что позволяет нейронной сети моделировать нелинейные зависимости.

40. Методы построения ансамблей нейронных сетей (Methods of building ensembles of neural networks) – методы, которые используют несколько нейронных сетей для решения задачи и комбинируют их прогнозы для достижения лучших результатов. Ансамбли нейронных сетей могут быть построены путем обучения независимых сетей на разных наборах данных, случайного исключения некоторых нейронов или слоев во время обучения или комбинирования прогнозов отдельных нейронных сетей с помощью голосования или взвешивания.

Рассмотрим подробнее несколько самых распространенных методов построения нейронных сетей [11, 15–18].

1. *Метод обратного распространения ошибки (Backpropagation)*. Самый популярный метод обучения нейронных сетей, который основывается на вычислении градиентов ошибки для каждого параметра модели и их последующем обновлении в направлении, противоположном градиенту. Он состоит из двух фаз: прямого прохода для вычисления выходных значений и обратного прохода для расчета градиента и обновления весов.

2. *Сверточные нейронные сети (Convolutional Neural Networks, CNN)*. Особый тип нейронных сетей, наиболее широко применяемый для обработки изображений. Они используют сверточные слои для поиска локальных шаблонов и пулинговые слои для уменьшения размерности изображения, что позволяет создавать специализированные фильтры для изображений.

3. *Рекуррентные нейронные сети (Recurrent Neural Networks, RNN)*. Тип нейронных сетей, используемый для обработки последовательных данных, таких как тексты, речь или временные ряды. Они имеют обратные связи между своими нейронами, что позволяет сохранять информацию о предыдущих состояниях и использовать ее для принятия решений.

4. *Генетические алгоритмы*. Это эволюционный метод, использующий принципы естественного отбора и мутации для оптимизации параметров нейронных сетей. Генетические алгоритмы создают случайные популяции моделей, затем выбирают лучшие и изменяют их параметры, чтобы найти наилучшее решение.

5. *Обучение без учителя*. Некоторые методы обучения нейронных сетей не требуют аннотированных данных или явных меток классов. Эти методы, такие как автоэнкодеры или генеративные состязательные сети (GAN), позволяют моделям самим извлекать иерархические признаки из неаннотированных данных или генерировать новые примеры на основе заданного распределения.

6. *Трансферное обучение*. Метод, который позволяет использовать знания, полученные при решении одной задачи, для улучшения работы на другой задаче. Этот метод может быть полезен, если у нас есть ограниченное количество размеченных данных для тренировки, но при этом есть доступ к большему количеству неразмеченных данных или данных из другой задачи.

Существует множество других методов и алгоритмов для построения нейронных сетей, и выбор оптимального метода зависит от конкретной задачи и доступных данных.

Методы, подходящие для прогнозирования курса

Для прогнозирования курсов валют часто используются методы временных рядов и регрессионные модели [10, 16, 17, 19]. Некоторые из наиболее эффективных методов в этой области:

1. *Стохастический градиентный спуск* (3) – метод оптимизации, часто применяемый в регрессионных задачах;

2. *Методы оптимизации второго порядка* (5) – используются для оптимизации параметров модели в регрессии;

3. *Методы регуляризации* (6) – помогают предотвратить переобучение, что важно при прогнозировании курса валют;

4. *Глубокое обучение* (9) – в некоторых случаях нейронные сети глубокого обучения могут успешно прогнозировать временные ряды, включая курсы валют;

5. *Рекуррентные нейронные сети (RNN)* (8) – эффективны для работы с последовательными данными и могут быть применены к временным рядам. RNN позволяют учитывать зависимости в прошлых значениях входных данных при прогнозировании будущих значений. Например, можно использовать LSTM (Long Short-Term Memory) нейронные сети для учета долгосрочных трендов;

6. *Методы обучения без учителя* (14) – возможно использование для обнаружения паттернов в данных курсов валют;

7. *Глубокие рекуррентные нейронные сети (Deep RNN)* – эти сети являются комбинацией рекуррентных слоев и обычных слоев нейронных сетей. Они могут применяться для обработки большого объема данных и создания более сложных временных моделей. Это особенно полезно при прогнозировании курса рубля, так как он может быть подвержен сезонным или более сложным паттернам;

8. *Глубокие автокодировщики (Deep Autoencoders)* – это нейронные сети, обученные реконструировать входные данные. Они могут быть использованы для сжатия и анализа временных рядов. После обучения они могут использоваться для прогнозирования будущих значений путем кодирования текущего состояния ряда и декодирования его в будущих точках времени.

Выбор конкретного метода зависит от характеристик и объема доступных данных, требуемого уровня точности прогнозирования и других факторов. Рекомендуется провести тщательный анализ данных и эксперименты, чтобы выбрать наиболее подходящий метод для вашей задачи прогнозирования курса.

Методы, которые не подходят для прогнозирования курса

Для прогнозирования курса рубля не рекомендуется использовать некоторые методы построения нейронных сетей, которые хорошо работают для других задач [18, 20]. Вот некоторые из таких методов.

1. *Полносвязные нейронные сети*. Анализ курса рубля обычно требует работы с временными рядами, которые имеют сложную зависимость между значениями в разные моменты времени. Полносвязные сети не всегда хорошо моделируют такую зависимость, поэтому они могут показать плохие результаты.

2. *Простейшие рекуррентные нейронные сети (RNN)*. Ранее указывалось, что RNN эффективны для прогнозирования, но есть ряд ограничений:

- *Проблема затухающего/взрывающегося градиента*. При обучении RNN данные могут передаваться на несколько временных шагов, и при этом градиенты могут уменьшаться или увеличиваться экспоненциально. Это может привести к проблемам с обучением модели и ограничить ее способность к долговременной памяти;

- *Ограниченная память*. RNN имеют ограниченную память и могут терять информацию из начала или середины последовательности, особенно если последовательность очень длинная. Это может затруднить точное прогнозирование в отдаленном будущем или на большом временном интервале;

- *Вычислительная сложность*. Обучение и использование RNN может быть вычислительно сложным, особенно если модель имеет много слоев и временных шагов. Это может ограничить их применимость для прогнозирования в реальном времени или на устройствах с ограниченными вычислительными ресурсами.

Несмотря на эти ограничения, RNN все еще являются полезным инструментом для прогнозирования и могут быть эффективно применены во многих задачах. Кроме того, существуют различные улучшения RNN, такие как LSTM (*долгая краткосрочная память*) и GRU (*обновляемые единицы*), которые были разработаны для решения некоторых из указанных проблем.

3. *Сверточные нейронные сети (CNN)*. Хотя CNN широко используются для обработки изображений, они, вероятно, не дадут хороших результатов при работе с временными рядами, такими как курс рубля, поскольку CNN не учитывают временные зависимости и могут проигрывать по сравнению с другими архитектурами.

4. *Метод обратного распространения ошибки* (1). Устаревший метод, который может иметь ограниченную эффективность при прогнозировании валютных курсов.

5. *Метод Гиббса* (20). Устаревший метод, преимущественно использовавшийся в марковских случайных полях, что не является типичным для задач прогнозирования валютных курсов.

6. *Алгоритмы оптимизации на основе искусственных пчел* (28). Эти алгоритмы обычно используются для решения задач оптимизации, а не для прогнозирования временных рядов, включая валютные курсы.

7. *Нейровизуализация* (29). Хотя нейровизуализация полезна для понимания работы нейронных сетей, она не предназначена для конкретного прогнозирования валютных курсов.

8. *Автоматическое определение архитектуры нейронной сети* (31). Этот метод скорее связан с выбором архитектуры, чем с прогнозированием конкретных значений временных рядов.

9. *Методы объединения множества моделей* (36). Хотя ансамбли моделей могут быть полезными, но в их использовании для прогнозирования валютных курсов может быть ограниченный смысл, если не учитываются специфические факторы, влияющие на валютные рынки.

Обычно для прогнозирования временных рядов, включая курс рубля, применяются *рекуррентные нейронные сети с долгой краткосрочной памятью (LSTM)* или *сверточные LSTM (ConvLSTM)*, которые способны лучше улавливать долгосрочные временные зависимости.

Результаты

Способы улучшения точности предсказания

Существует несколько способов минимизировать ошибку при прогнозировании нейронными сетями курса рубля [21, 22].

1. *Увеличить объем и качество данных*. Чем больше и качественнее данные, используемые для обучения нейронной сети, тем точнее будет прогноз. Важно использовать разнообразные и актуальные данные, например исторические данные о курсе рубля, макроэкономические показатели, политические и экономические события.

2. *Настроить архитектуру и параметры нейронной сети*. Разработчикам следует экспериментировать с различными архитектурами нейронных сетей и параметрами обучения, чтобы найти наилучшую модель. Например, можно попробовать различные типы слоев (рекуррентные, сверточные, полносвязные) и количество скрытых нейронов.

3. *Регуляризация и предотвращение переобучения*. Переобучение может привести к низкой обобщающей способности нейронной сети. Для предотвращения переобучения можно использовать методы регуляризации, такие как

L1- или L2-регуляризация, а также можно использовать методы сокращения размерности данных, например *PCA (Principal Component Analysis)* или *t-SNE (t-distributed Stochastic Neighbour Embedding)*.

4. *Внедрение дополнительной информации*. Курс рубля может зависеть от других внешних факторов, таких как политические события, экономические показатели и рыночные тренды. Внедрение таких данных в обучающий набор может улучшить прогноз.

5. *Использование ансамблей моделей*. Ансамблирование моделей позволяет комбинировать прогнозы нескольких нейронных сетей, что может улучшить точность прогнозирования.

6. *Регулярное обновление модели*. Как только новые данные становятся доступными, модель нейронной сети должна быть обновлена и переобучена на этих данных, чтобы сохранять актуальность и точность прогнозов.

Важно отметить, что нейронные сети не всегда дают абсолютно точные прогнозы, так как рынки валют подвержены множеству факторов и сложным взаимосвязям между ними. Однако правильное применение вышеуказанных способов может помочь минимизировать ошибку при прогнозировании.

Подробнее рассмотрим второй и третий пункты, которые можно считать более сложными.

Количество скрытых нейронов в нейронной сети может оказывать влияние на точность прогнозирования курса рубля, на способность сети извлекать сложные или нелинейные зависимости из данных.

Если количество скрытых нейронов слишком мало, то сеть может быть недостаточно гибкой для моделирования сложных паттернов и зависимостей в данных. В таком случае прогнозы сети могут быть не очень точными.

С другой стороны, если количество скрытых нейронов слишком велико, то сеть может начать переобучаться, то есть запоминать шумовые или случайные зависимости из обучающих данных, которые не обобщаются на новые данные. В таком случае точность прогнозов сети также снижается.

Таким образом, необходимо подобрать оптимальное количество скрытых нейронов, чтобы балансировать между гибкостью модели и ее способностью обобщать новые данные. Это может потребовать экспериментирования и оптимизации на основе результатов валидации модели.

Переобучение нейронных сетей при прогнозировании курса рубля может иметь следующие последствия:



Рис. 3. Блок-схема SA-алгоритма
Составлено авторами

1. *Потеря обобщающей способности.* Переобучение может привести к тому, что нейронная сеть будет очень хорошо предсказывать значения курса рубля на обучающей выборке, но плохо справляться с новыми данными. Сеть может запоминать шум или случайные корреляции в данных, что приводит к искаженным прогнозам на новых данных;

2. *Высокая чувствительность к изменениям.* Переобученная нейронная сеть может быть слишком чувствительной к небольшим изменениям в данных, что может привести к непредсказуемым и нестабильным прогнозам курса рубля. Это может быть особенно проблематично в случае значительных изменений в экономике или политике;

3. *Неправильные выводы.* Переобучение может привести к неправильным выводам о том, какие факторы на самом деле влияют на курс рубля. Нейронная сеть может недооценивать или переоценивать важность определенных факторов, что может привести к неверным прогнозам;

4. *Отсутствие способности обучения на новых данных.* Если нейронная сеть переобучена на доступных данных, она может не справиться с новыми данными или не сможет обучиться на них эффективно. Это может привести к постепенному ухудшению качества прогнозов или невозможности адаптироваться к новым условиям.

Для предотвращения переобучения нейронных сетей при прогнозировании курса рубля можно использовать методы регуляризации, такие как L1- и L2-регуляризация, ранняя остановка обучения, а также можно ограничить сложность модели при помощи выбора правильной архитектуры сети и оптимальных гиперпараметров. Кроме того, необходимо обеспечить разнообразие и репрезентативность данных, используемых для обучения, и проводить оценку качества модели на валидационной выборке.

Прогнозирование курса рубля – сложная задача, и эффективность методов может зависеть от различных факторов, таких как экономические условия, политическая обстановка и т.д.

Приведем методы, кроме вышеописанных нейронных сетей, которые могут быть полезны для прогнозирования курса рубля.

Временные ряды и статистические методы:

1. *ARIMA (авторегрессия с интегрированным скользящим средним).* Хорошо подходит для моделирования временных рядов и учета трендов;

2. *GARCH (общий авторегрессивный условный дисперсионный процесс).* Подходит для моделирования волатильности во временных рядах курсов валют.

Регрессионные модели:

1. *Линейная регрессия* может быть использована для учета влияния различных факторов (например, экономических показателей) на курс рубля;

2. *Методы машинного обучения (например, градиентный бустинг, случайные леса)* могут обрабатывать нелинейные зависимости и улучшать прогнозы при наличии большого объема данных.

Эконометрические модели:

1. *Модель множественной регрессии с использованием эконометрических переменных* позволяет учитывать фундаментальные экономические показатели, которые могут влиять на курс рубля.

Важно отметить, что прогнозирование валютных курсов – это сложная задача, и ни один метод не гарантирует абсолютную точность. Самый эффективный подход – это комбинирование различных методов и учет экспертных оценок. Все это поможет улучшить результаты.

Примерная блок-схема SA-алгоритма (см. рис. 3)

1. Подготовка данных:

- Загрузка исторических данных о курсах валют и соответствующих факторов;
- Очистка данных от выбросов и пропущенных значений.

2. Инженерия признаков:

- Создание временных окон, лагов, скользящих средних и других признаков для учета долгосрочных и краткосрочных зависимостей;



Рис. 4. Блок-схема SA-алгоритма для оптимизации гиперпараметров
Составлено авторами

- Нормализация данных для обеспечения стабильности обучения.

3. Выбор архитектуры НС:

- Выбор между RNN, LSTM и другими архитектурами в зависимости от особенностей данных;

- Настройка гиперпараметров, таких как количество слоев, нейронов, скорость обучения.

4. Разделение данных на обучающий и тестовый наборы

5. Обучение НС:

- Процесс обучения с использованием обучающего набора;

- Мониторинг функции потерь и метрик производительности.

6. Оценка производительности:

- Тестирование НС на тестовом наборе;
- Вычисление и анализ метрик производительности.

7. Прогнозирование будущих значений:

- Использование обученной НС для прогнозирования будущих курсов валют.

8. Обновление модели (по мере необходимости):

- Периодическое обновление модели с использованием новых данных [22, 23].

Пример SA-алгоритма (Simulated Annealing) для оптимизации гиперпараметров (рис. 4)

1. Инициализация:

- Задание начальных значений гиперпараметров.

2. Повторение:

2.1. Выбор начальной точки:

- Изменение случайного гиперпараметра или комбинации гиперпараметров.

2.2. Оценка старой и новой модели:

- Обучение модели с текущими гиперпараметрами и оценка ее производительности.

2.3. Принятие/отклонение изменения:

- Принятие изменения с вероятностью, зависящей от разницы в оценках производительности;

- Если новая модель лучше, она становится текущей, в противном случае остается старая.

2.4. Понижение ошибки:

- Уменьшение вероятности принятия худшего решения по мере продвижения по процессу оптимизации.

2.5. Проверка условия завершения:

- Повторение процесса до достижения определенного критерия остановки (например, заданного числа итераций или достижения определенной производительности);

- Если нет, возвращение к шагу 1.

3. Завершение:

- Возврат оптимальных гиперпараметров или лучшей модели.

SA-алгоритм используется для эффективного поиска оптимальных гиперпараметров, обеспечивая баланс между исследованием пространства гиперпараметров и эксплуатацией наилучших находок.

Обзор применения нейронных сетей для прогнозирования валютных курсов

1. Подготовка данных:

- Исторические данные: сбор данных о валютных курсах за определенный период времени;

- Финансовые показатели: включение дополнительных факторов, таких как данные о ВВП, инфляции, процентных ставках, курсе нефти и других макроэкономических показателях.

2. Выбор архитектуры нейронной сети:

- Рекуррентные нейронные сети (RNN) или LSTM: учет временных зависимостей в последовательных данных курсов валют;

- Глубокие нейронные сети (DNN): обработка сложных нелинейных зависимостей в данных.

3. Учет факторов при прогнозировании

3.1. Технические индикаторы:

- Скользящие средние, индикаторы стохастичности, относительной силы и др.

3.2. Экономические показатели:

- ВВП, инфляция, ставки, цена нефти, объемы экспорта и импорта, безработица и другие макроэкономические факторы, влияющие на валютные курсы.

3.3. Геополитические события:

- Политическая стабильность, международные отношения и другие события, которые могут повлиять на рынок.

3.4. Объемы торгов:

- Информация о том, сколько активов было куплено или продано.

3.5. Учет ставок:

- Разница в процентных ставках между разными странами.

4. Обучение и оценка:

- Разделение данных на обучающий и тестовый наборы для оценки производительности;

- Обучение сети: использование обучающего набора для настройки весов нейронной сети;

- Оценка производительности: использование метрик, таких как среднеквадратичная ошибка (MSE), коэффициент детерминации (R²) и другие, для оценки точности модели.

5. Адаптация к изменениям на рынке:

- Обновление модели: периодическое обновление модели с использованием новых данных для адаптации к изменениям на финансовых рынках.

Аналитики могут подчеркивать важность разных факторов, которые следует учитывать при анализе и прогнозировании курсов финансовых активов.

Заключение

В данной статье были рассмотрены различные методы и модели построения нейронных сетей для прогнозирования курсов финансовых активов. Было проведено исследование, в ходе которого изучено влияние различных факторов на точность при прогнозировании курса рубля и определены наиболее эффективные методы.

Рассмотрены разные параметры нейронных сетей, включая количество скрытых нейронов, их взаимосвязь и влияние на точность прогноза. Результаты исследования показали, что, помимо количества скрытых нейронов, важными факторами являются правильный выбор входных данных, методы нормализации данных, а также параметры обучения и оценки модели. Более того, использование нескольких структур нейронных сетей и их комбинаций может привести к еще более точным прогнозам курса.

Финансовые рынки чрезвычайно сложны и подвержены воздействию множества факторов. Опытные аналитики используют сочетание фундаментального и технического анализа, а также учитывают макроэкономические и геополитические события для более точного прогнозирования, при этом важность различных факторов будет меняться со временем.

Список литературы

1. Artificial intelligence (AI) market size worldwide in 2021 with a forecast until 2030. URL: <https://www.statista.com/statistics/1365145/artificial-intelligence-market-size/> (дата обращения: 10.12.2023).

2. Концепция технологического развития на период до 2030 года. Утв. Правительством РФ от 20 мая 2023 г. № 1315-п. // СПС «Гарант».

3. Мусаев Т.А., Федоров О.В., Шагеев С.Р., Прохорова М.В. Интеллектуальные системы учета как инструмент снижения потерь электрической энергии // Строительство: новые технологии – новое оборудование. 2021. № 2. С. 52–55.

4. Будущее e-commerce: как Amazon применяет генеративный ИИ для оптимизации товарных описаний. URL: <https://vc.ru/u/2204210-neyrosfera/838753-budushchee-e-commerce-kak-amazon-primenyaet-generativnyu-ii-dlya-optimizacii-tovaryh-opisaniy?from=rss> (дата обращения: 12.12.2023).

5. Манжула В.Г., Федяшов Д.С. Нейронные сети Кохонена и нечеткие нейронные сети в интеллектуальном анализе данных // Фундаментальные исследования. 2011. № 4. С. 108–114.

6. Червяков Н.И., Тихонов Э.Е. Применение нейронных сетей для задач прогнозирования и проблемы идентификации моделей прогнозирования на нейронных сетях // Нейрокомпьютеры: разработка, применение. 2003. № 10-11. С. 25–31.

7. Качалов О.И., Миронов А.Н., Володина А.М. Сравнение точности прогнозирования стоимости акций компании с использованием свёрточной нейронной сети (CNN), простой рекуррентной нейронной сети (SIMPLE RNN) и рекуррентной нейронной сети на базе долгой краткосрочной памяти (LSTM) // ИТ-Стандарт. 2018. № 4 (17). С. 49–61.

8. Белобородова Н.А. Модели прогнозирования роста производства с использованием нейронных сетей // Известия высших учебных заведений. Нефть и газ. 2002. № 4 (34). С. 110–117.

9. Мухаметшин Р.Р. Инновационный подход к оценке и управлению инвестиционными рисками на фондовом рынке РФ с использованием нейронных сетей // Вестник Казанского технологического университета. 2009. № 5. С. 164–171.

10. Костина Л.Н., Гареева Г.А. Нейронные сети в задачах прогнозирования временных рядов // Инновационная наука. 2015. Т. 2. № 6 (6). С. 70–73.

11. Пальчиков И.А. Применение нейронных сетей при управлении бизнесом // Инновации, технологии и бизнес. 2021. № 2 (10). С. 63–67.

12. Атнабаев И.Ю. Прогнозирование цен акций Сбербанка на основе искусственных нейронных сетей // Перспективы развития информационных технологий. 2016. № 29. С. 7–11.

13. Лазарева И.Е. Моделирование уровня финансовой стабильности предприятий на основе нейронных сетей // Торговля и рынок. 2021. № 1 (57). С. 282–294.

14. Кочнев А.А. Применение искусственных нейронных сетей в прогнозировании // Научные горизонты. 2023. № 1 (65). С. 48–57.

15. Кумратова А.М., Чумаренко К.Э. Прогнозирование сложных процессов нейронными сетями //

Современная экономика: проблемы и решения. 2023. № 3 (159). С. 27–36. DOI: 10.17308/meps/2078-9017/2023/3/27-36.

16. Абрамов А.В. Прогнозирование валютного курса EUR/USD с использованием нейронных сетей // Ученые записки. Электронный научный журнал Курского государственного университета. 2012. № 4-2 (24). С. 71–76.

17. Ломакин Н.И., Максимова О.Н., Экова В.А. и др. Нейронные сети для прогнозирования курса доллара с использованием астрологических циклических индексов Гюшон и Ганю // Международный журнал прикладных и фундаментальных исследований. 2016. № 6-1. С. 133–136.

18. Гринева Н.В. Построение нейронной сети для прогнозирования цены опциона // Проблемы экономики и юридической практики. 2022. Т. 18. № 5. С. 190–199.

19. Bloomberg создал аналог модели искусственного интеллекта GPT в области финансов. URL: <https://www.forbes.ru/tekhnologii/487606-bloomberg-sozdal-analog-modeli-iskusstvennogo-intellekta-gpt-v-oblasti-finansov> (дата обращения: 12.12.2023).

20. Кочеваткина Э.Ф., Миляева Н.В., Устинова Н.Н. Применение нейронных сетей для прогнозиро-

вания экономических явлений // Modern Economy Success. 2022. № 2. С. 100–106.

21. Как Яндекс применил генеративные нейросети для поиска ответов. URL: <https://habr.com/ru/companies/yandex/articles/561924/> (дата обращения: 12.12.2023).

22. ESG and artificial intelligence (AI) – statistics & facts. URL: <https://www.statista.com/topics/11077/esg-and-ai/#topicOverview> (дата обращения: 10.12.2023).

23. Бардаков А.А., Корнилов Д.А. Реинжиниринг бизнес-процессов в структуре методов повышения эффективности деятельности организации // Управление экономическими системами: электронный научный журнал. 2019. № 8 (126). С. 3.

24. Global total corporate artificial intelligence (AI) investment from 2015 to 2022. URL: <https://www.statista.com/statistics/941137/ai-investment-and-funding-worldwide/> (дата обращения: 10.12.2023).

25. Корнилова Е.В., Захаров В.Я., Корнилов Д.А. Оценка устойчивого развития и формирование рейтинга устойчивости регионов страны // Развитие и безопасность. 2023. № 1 (16). С. 30–49. DOI: 10.46960/2713-2633_2023_1_36.

CLASSIFICATION AND APPLICATION OF NEURAL NETWORKS FOR FORECASTING FINANCIAL INDICATORS

A.I. Polaznov^{1,2}, D.A. Kornilov²

¹Atomenergoproekt Open Joint Stock Company

²Nizhny Novgorod State Technical University n.a. R.E. Alekseev

This article is a study of various methods for constructing neural networks for forecasting financial indicators, in particular for forecasting exchange rates. The relevance of this study is due to the need to develop effective models for forecasting exchange rates, stocks, cryptocurrencies and other financial indicators. The purpose of the study is to review neural networks, formulate their classification, select the optimal methods and models for using neural networks for forecasting financial indicators, including forecasting exchange rates. As part of the work, various neural network architectures were considered, including simple recurrent neural networks (RNN), convolutional neural networks (CNN), deep neural networks (DNN) and others. More than 40 methods were classified and their descriptions were presented. For each method, their advantages and disadvantages were analyzed from the point of view of course forecasting. The results of the study showed that the choice of the optimal method for predicting heading depends on the characteristics of a particular task. For example, RNNs performed better on problems with time series, while CNNs were effective on problems with spatial structure. The issues of increasing the accuracy of prediction and preventing overtraining of the neural network were considered. A block diagram of the SA algorithm is presented, which allows you to visualize the sequence of steps and operations. This work can be a useful guide for researchers and practitioners involved in course forecasting. The results of the work can contribute to more accurate and reliable development of forecasting models, which is of great importance for the financial sector and other areas where predicting exchange rates and improving the efficiency of financial planning is important.

Keywords: artificial intelligence, neural networks, forecasting, classification, process model, forecasting, exchange rate, ruble exchange rate, methods.