

Анализ изменения точности прогнозов нейронной сети, в зависимости от используемых входных параметров

Мисник А.Е.
МОУ ВО «Белорусско-
Российский университет»
Могилев, Республика Беларусь
anton@misnik.by

Крутолевич С.К.
МОУ ВО «Белорусско-
Российский университет»
Могилев, Республика Беларусь
s_krutolevich@tut.by

Прокопенко С.А.
МОУ ВО «Белорусско-
Российский университет»
Могилев, Республика Беларусь
puss95@yandex.by

Боровых П.А.
RoninAI Lab
Нью-Йорк, США
peter.borovvkh@gmail.com

Васильев М.В.
RoninAI Lab
Могилев, Республика Беларусь
max.vasilveu@gmail.com

Abstract — The neural networks are tool for approximation universal series of data, but their precision highly depends on a adequate set of inputs. Cryptocurrencies experience high levels of volatility due to absence of agreed upon pricing methodologies behind its valuation. In this article we analyze approaches of obtaining additional inputs for neural networks and explore their influence on its precision.

Ключевые слова — нейронные сети, подбор входных параметров, анализ точности

I. ВВЕДЕНИЕ

Нейронные сети являются универсальным аппроксиматором данных, но точность их прогнозов в значительной степени зависит от адекватного набора входных данных. Из-за отсутствия согласованных методологий ценообразования криптовалюты характеризуются высоким уровнем волатильности. В данной статье анализируются подходы получения дополнительных входных данных для нейронных сетей и исследуется их влияние на точность прогнозов нейронной сети.

Для анализа выбрано изменение точности прогнозов нейронной сетью цены Биткойн (BTC), с использованием в качестве входных параметров рыночных данных, социальных и временных факторов.

Для сравнения в исследовании включены две архитектуры нейронных сетей: многослойный перцептрон (MLP) и сеть долгой краткосрочной памяти (LSTM). MLP является наиболее распространенной архитектурой, обеспечивающей приемлемую точность в большом количестве задач, а LSTM считается более точной в задачах прогнозирования временных рядов.

Результаты проверены и основаны на нейронных сетях, созданных RoninAI Lab. RoninAI использует различные нейронные сети для прогнозирования курса криптовалют, предоставляя практические данные и их анализ для данного исследования.

II. СБОР РЫНОЧНЫХ ДАННЫХ

В ряде источников упоминается возможность прогнозирования цены актива с использованием рыночных данных.

Чтобы проверить точность прогнозирования рыночных данных, мы начали со сбора исторических минутных рыночных данных за год с одной из старейших и крупнейших криптовалютных бирж - Kraken. Входные параметры рыночных данных приведены в таблице I.

ТАБЛИЦА I. ПАРАМЕТРЫ РЫНОЧНЫХ ДАННЫХ

Название параметра	Описание
ДАТА И ВРЕМЯ	Отметка времени данной минуты
ОТКРЫТИЕ	Цена актива в 0 секунду текущей минуты
ЗАКРЫТИЕ	Цена актива в 59 секунду текущей минуты
МАКСИМУМ	Максимальная цена актива в течение текущей минуты
МИНИМУМ	Минимальная цена актива в течение текущей минуты
ОБЪЕМ ОТ	Минимальный объем продаж в секунду
ОБЪЕМ ДО	Максимальный объем продаж в секунду

III. ОБУЧЕНИЕ НЕЙРОННЫХ СЕТЕЙ НА ОСНОВЕ РЫНОЧНЫХ ДАННЫХ

В качестве эталонной модели для анализа будем рассматривать MLP и LSTM нейронные сети, обученные на основе рыночных данных. [1,2,5]

После обучения нейронных сетей MLP и LSTM, использующих рыночные данные в качестве входных данных и цену BTC в следующую минуту в качестве выходных данных, модели продемонстрировали результаты, представленные на рис. 1.

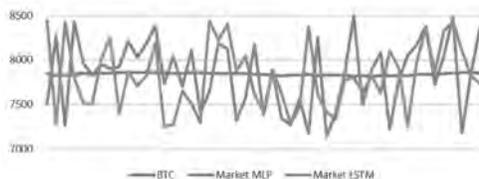


Рис. 1. Результаты обучения нейронных сетей на основе рыночных данных

Точность прогнозов обеих нейронных сетей оказалась достаточно низкой. LSTM показала чуть лучшую точность.

IV. ДОПОЛНЕНИЕ ВХОДНЫХ ПАРАМЕТРОВ ВРЕМЕННЫМ ФАКТОРОМ

Первым дополнительным фактором, включенным в исследование, является экспертный фактор времени, отражающий уровень активности бирж. Переобучение нейронных сетей с дополнительным временным фактором привело к результатам, показанным на рис. 2.

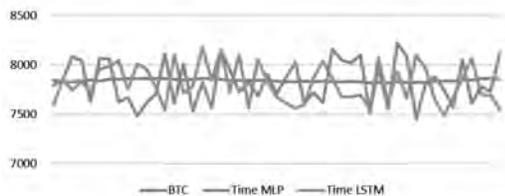


Рис. 2. Результаты обучения нейронных сетей на основе рыночных данных и временного фактора

Несмотря на достаточно низкий уровень точности, прогнозы, сделанные как MLP, так и LSTM, оказались значительно лучше по сравнению с прогнозами, основанными только на рыночных данных.

V. ДОПОЛНЕНИЕ ВХОДНЫХ ПАРАМЕТРОВ СОЦИАЛЬНЫМ ФАКТОРОМ

Считается, что большинство игроков на рынке криптовалюты являются розничными инвесторами, которые принимают свои инвестиционные и торговые решения, в частности, глядя на изменение графиков курса криптовалют. Пример такого графика представлен на рисунке 3.

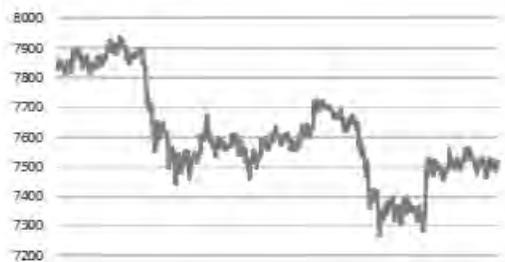


Рис. 3. График изменения курса BTC

В рамках исследования было решено проверить влияние факторов визуального восприятия на прогноз нейронной сети.

Рассмотрим алгоритм, который позволяет нам идентифицировать некоторые сложные показатели числового ряда, которые мы можем использовать для оценки субъективного восприятия трейдером визуального графика.

В общем, любой ряд данных может рассматриваться как сумма линейных и гармонических компонентов.

Цель дальнейшего исследования заключается в исследовании алгоритма выделения этих составляющих и их нормирования. [4,6,11]

Предложенный алгоритм предполагает проведение следующих этапов.

1. Определение минимального элемента e_{min} ряда E.
2. Проведение операции вычитания

$$E' = e_i - e_{min} \quad (1)$$

3. Аппроксимация ряда E' полиномом первого порядка вида

$$e_i = a(0) + a(1) * n, \quad (2)$$

где $a(0)$, $a(1)$ – коэффициенты аппроксимации; n – дискретные значения оси времени.

4. Вычисление коэффициента относительного изменения линейной составляющей за период T по формуле:

$$e_1 = a(1) * T * 100 / e_{min} \quad (3)$$

Величина e_1 является относительной и не зависит от абсолютного значения ряда. Если величина $e_1 > 0$ линейная составляющая увеличивается.

5. Для оценки гармонической составляющей ряда проведем прямое преобразование Фурье (FFT) для ряда E'.

6. Определим модули амплитуд колебаний в частотной области A(w). Проведём фильтрацию частот по значениям амплитуд.

Для анализа эффективности предложенного алгоритма проведём ряд экспериментов в среде MATLAB.

Шаг 1. Линейная составляющая постоянная. Гармоническая отсутствует. Результаты представлены на рисунке 4. В качестве легенды на второй график выводится значение коэффициента e_1 . В данном случае $e_1 = 0$.



Рис. 4. Анализ эффективности, шаг 1

Шаг 2. Линейная составляющая растет. Гармоническая отсутствует. Результаты представлены на рисунке 5. Коэффициент $e_1 = 20\%$. Модули спектра имеют значения в области низких частот (1-2 Гц.). Следует помнить, что основная полоса частот для модуля вещественной последовательности лежит в интервале $0 < k < N/2 - 1$.

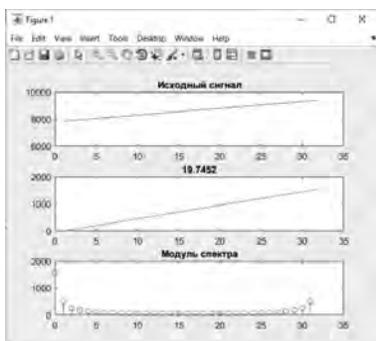


Рис. 5. Анализ эффективности, шаг 2

Поэтому частоты выше 15 Гц в нашем примере должны игнорироваться. При необходимости учитывать более высокие частоты увеличим число точек дискретизации времени N .

Шаг 3. Линейная составляющая уменьшается. Гармоническая отсутствует. Результаты представлены на рисунке 6. Значение коэффициента $e_1 < 0$.

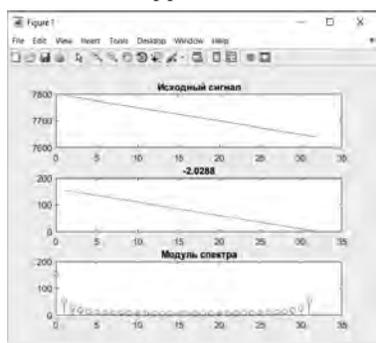


Рис. 6. Анализ эффективности, шаг 3

Шаг 4. Линейная составляющая отсутствует. Гармоническая присутствует при частоте 3 Гц - амплитуда 20. Результаты представлены на рисунке 7.

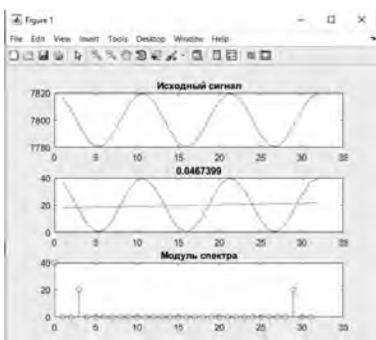


Рис. 7. Анализ эффективности, шаг 4

Шаг 5. Линейная составляющая возрастает. Гармоническая присутствует при частоте 3 Гц - амплитуда 400. Результаты представлены на рисунке 8.

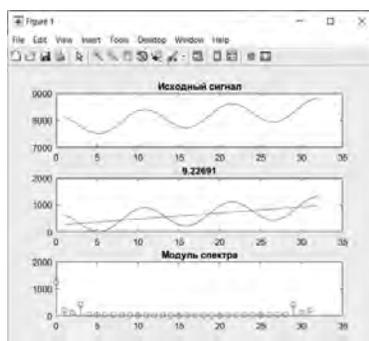


Рис. 8. Анализ эффективности, шаг 5

Применим предложенный алгоритм к данным, представленным на рисунке 3.

Результат представлен на рисунке 9. На рисунке видно, что линейная составляющая прогнозирует тренд на снижение со скоростью 4,1 % за период. Модули амплитуд гармонических составляющих оказывают влияние в частотах до 10 Гц.

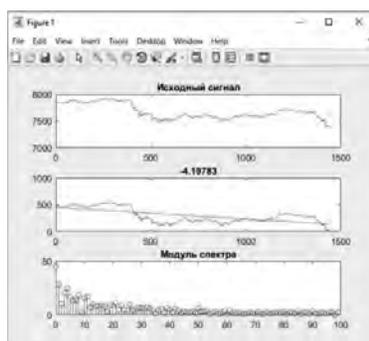


Рис. 9. Анализ эффективности, график курса BTC

Предложенный алгоритм выделяет линейную и гармонические составляющие числового ряда.

Предложенный коэффициент e_1 – может служить мерой оценки тренда изменений значений числового ряда.

Модули амплитуд колебаний в частотной области после прямого преобразование Фурье могут служить мерой оценки колебательной составляющей ряда данных.

В нашем случае мы расширили входы нейронных сетей 4 новыми параметрами: коэффициент e_1 и 3 диапазона гармонических составляющих: от 0 до 10 Гц, от 10 до 30 Гц, 30 Гц и выше.

После расчета вышеупомянутых параметров для каждого ряда нашего набора данных сети MLP и LSTM прошли переобучение. Результаты обучения показаны на рисунке 10.

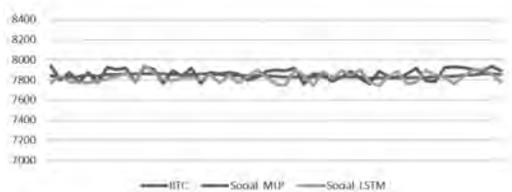


Рис. 10. Результаты обучения нейронных сетей на основе рыночных данных, временного и социального факторов

Точность прогнозов в сетях MLP и LSTM оказалась намного лучше по сравнению с предыдущими результатами.

VI. ДОПОЛНЕНИЕ ВХОДНЫХ ПАРАМЕТРОВ ФАКТОРОМ СОЦИАЛЬНЫХ СЕТЕЙ

В последнее время социальные сети стали мощным явлением, влияющим на мнение людей по различным вопросам.

Чтобы сделать прогнозы более точными, мы попытались использовать данные, полученные из социальных сетей.

Мы расширили наши входные данные параметрами, извлеченными из социальных сетей по темам, связанным с BTC. Результаты обучения нейронных сетей показаны на рисунке 11.

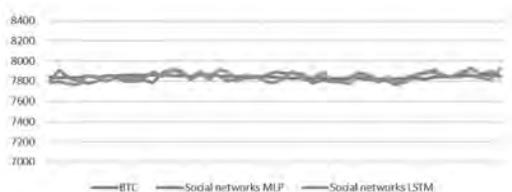


Рис. 11. Результаты обучения по рыночным данным, фактору времени, социальным факторам и факторам социальных сетей

Точность прогнозов сетей MLP и LSTM стала еще ближе к реальной цене BTC.

VII. ЗАКЛЮЧЕНИЕ

Дополнение рыночных данных релевантными параметрами может значительно повысить точность прогнозов.

Метод, описанный в разделе V, можно использовать на различных числовых рядах данных для получения дополнительных входных параметров.

В наших экспериментах сети LSTM показали несколько лучшую точность, чем MLP.

СПИСОК ЛИТЕРАТУРЫ

- [1] Хайкин С. Нейронные сети: полный курс, 2-е изд., испр. — М. ЖОО «И.Д. Вильямс», 2006. — 1104 с.
- [2] Галушкин А.И. Теория нейронных сетей. Кн. 1. М.: ИПРЖР, 2000. — 416 с.
- [3] Tang, Jiliang; Chang, Shiyu; Aggarwal, Charu; Liu, Huan (2015). "Negative Link Prediction in Social Media" (PDF). In Proceedings of ACM International Conference on Web Search and Data Mining.
- [4] «Исследование и разработка методов и моделей интеллектуального управления рисками в сложных организационно-технических систе-мах», НИР, 2011—2013 г.г., Минобрнауки России, Аналитическая ведом-ственная целевая программа «Развитие научного потенциала высшей шко-лы», в рамках тематического плана НИУ «МЭИ», Договор № 1043110. Отчет о НИР; Борисов В.В., Сеньков А.В., Мисник А.Е. и др. — М., 2012. № гос.рег. 01201067780. — 108 с.
- [5] Головкин В.А. Нейронные сети: обучение, организация, применение. — Кн.4. — М.: ИПРЖР. — 2001.
- [6] Борисов В.В., Мисник А.Е. Иерархическая нейросетевая модель систем теплоснабжения // Энергетика, Информатика, Инновации—2011 — ЭИИ-2011: сб. трудов Междунар. науч.-техн. конф. В 2 т. Т. 2 Секции 3, 4, 5. Смоленск: РИО филиала ГОУВПО МЭИ(ТУ) в г. Смоленске, 2011. — 291 с. С. 205—207.
- [7] Wierstra, Daan; Schmidhuber, J.; Gomez, F. J. (2005). "Evolino: Hybrid Neuroevolution/Optimal Linear Search for Sequence Learning". Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI), Edinburgh: 853–858.
- [8] Борисов В.В., Мисник А.Е. Комбинированный нейросетевой способ моделирования для оперативного управления сложными системами // Информационные технологии. — 2012. — № 7. — С. 69–72.
- [9] Мисник А.Е., Борисов В.В. Композиционное нейросетевое моделирование сложной технической системы // Тез. докл. XII Всероссийской научной конференции «Нейро-компьютеры и их применение», «НКП-2014», Москва, 18.04.2014. — С. 64–65.
- [10] Tax, N.; Verenich, I.; La Rosa, M.; Dumas, M. (2017). "Predictive Business Process Monitoring with LSTM neural networks". Proceedings of the International Conference on Advanced Information Systems Engineering (CAiSE): 477–492. arXiv:1612.02130 Freely accessible. doi:10.1007/978-3-319-59536-8_30.
- [11] Борисов В.В., Мисник А.Е. Использование типизированных нейросетевых моделей сложной технической системы // Информационные технологии, Энергетика и Экономика—2013 — Сб. трудов 10-й Междунар. Науч.-техн. конф. В 3 т. Т. 2 Секции 4, 5. Смоленск: РИО филиала МЭИ в г. Смоленске, 2013. — 282 с. С. 103—105.