

УДК 004.853

ИССЛЕДОВАНИЕ И РАЗРАБОТКА ГЕНЕРАТИВНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ АНАЛИЗА ФИНАНСОВЫХ ВРЕМЕННЫХ РЯДОВ

Жданов М.Е. (Университет ИТМО)

Научный руководитель – преподаватель, Михайленко А.Ю. (Университет ИТМО)

Введение. Генеративно-сопоставительные сети (GAN) — это семейство архитектуры нейронных сетей, которое добилось хороших результатов в создании изображений и успешно применяется для создания временных рядов и других типов финансовых данных. Цель этого исследования — представить обзор того, как работают эти GAN, их возможности и ограничения в текущем состоянии исследований с финансовыми данными, а также разработать собственную генеративную нейронную сеть с учетом устранения недостатков других моделей.

Основная часть. Были выявлены недостатки типовых моделей для анализа финансовых временных рядов торговых бирж:

1) Collapse mode. Может случиться так, что Генератор сойдет до настройки, при которой он всегда будет выдавать одни и те же выходные данные. Несмотря на то, что Генератор может обмануть Дискриминатор, он не может представить сложное распределение данных в реальном мире. Способ решения: использование метрики Вассерштейна внутри функции ошибки, позволяет дискриминатору быстрее обучаться выявлять повторяющиеся выходы, на которых стабилизируется генератор.

2) Проблема исчезающего градиента — это сложность обучения искусственных нейронных сетей с помощью методов обучения на основе градиента и обратного распространения ошибки. В таких сценариях каждый из весов нейронной сети получает обновление, пропорциональное частной производной функции ошибок относительно текущего веса на каждой итерации обучения. В некоторых случаях градиент будет небольшим, если вес не изменится. Это может помешать обучению нейронной сети.

3) На первом этапе никакие корректировки параметров сети Дискриминатора не производятся. Генератор обучается на определенное количество так называемых тренировочных шагов. Генератор обучается с использованием обратного распространения ошибки. Прогресс обучения зависит от текущего состояния Дискриминатора, потому что при обучении GAN могут возникать различные проблемы. Один из них заключается в том, что обратная связь Дискриминатора ухудшается, так как Генератор работает лучше во время обучения. Обучение должно быть завершено вовремя, иначе качество Генератора и Дискриминатора может снизиться из-за случайных возвращаемых значений Дискриминатора.

Сфера применения модели — торговая биржа.

Датасет - список всех акций компаний S&P 500. Набор данных включает в себя список всех акций, содержащихся в нем, и связанные с ними ключевые финансовые показатели, такие как цена, рыночная капитализация, прибыль, соотношение цены и прибыли, цена к балансовой стоимости.

Метрики для оценки качества модели: Recall, Inception Score и Fréchet Inception Distance.

Выводы. Проведен анализ недостатков генеративно-сопоставительных сетей, определена сфера применения, найден набор данных и определены метрики для оценки качества модели.

Список использованных источников:

1. Caspar Hogeboom, "Generation of Synthetic Financial Time-Series with Generative Adversarial Network", June 25, 2020, url: <https://new.ingwb.com/binaries/content/assets/insights/themes/empowering-with-advanced-analytics/internship-theses-at-wbaa/generation-of-synthetic->

financial-time-series-with-generative-adversarial-networks.pdf

2. Jake Frankenfield, "Artificial Neural Network (ANN)", Aug 28, 2020, url: <https://www.investopedia.com/terms/a/artificial-neural-networks-ann.asp>

3. "Neural Network", url: <https://deepai.org/machine-learning-glossary-and-terms/neural-network>

Жданов М.Е. (автор)

Подпись

Михайленко А.Ю. (научный руководитель)

Подпись