

ОСНОВЫ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ВАЛЮТНЫХ КУРСОВ

Кондаков Дмитрий Сергеевич

студент, Тюменский Индустриальный Университет, РФ, г. Тюмень

Введение.

Прогнозирование валютных курсов является важной задачей для трейдеров, инвесторов и экономистов. Точные прогнозы помогают принимать обоснованные решения, минимизировать риски и максимизировать прибыль. Традиционные методы прогнозирования, такие как эконометрические модели, имеют свои ограничения, поэтому применение методов машинного обучения (ML) становится все более популярным. Машинное обучение предлагает широкий спектр инструментов и алгоритмов, которые могут значительно повысить точность прогнозов валютных курсов.

Основные методы машинного обучения.

Линейная регрессия.

Линейная регрессия является одним из самых простых и широко используемых методов машинного обучения для прогнозирования. Она предполагает линейную зависимость между входными признаками и целевой переменной. В контексте валютных курсов, линейная регрессия может использоваться для прогнозирования будущих значений на основе исторических данных.

Преимущества:

1. Простота и интерпретируемость модели.
2. Быстрое обучение на небольших наборах данных.

Недостатки:

1. Ограниченность в моделировании сложных нелинейных зависимостей.
2. Чувствительность к выбросам и аномалиям в данных.

Решающие деревья и случайные леса.

Решающие деревья представляют собой иерархическую структуру, где каждый узел выполняет проверку на основе одного из признаков данных, а листья содержат прогнозируемые значения. Случайные леса — это ансамблевый метод, который комбинирует несколько решающих деревьев для улучшения точности и устойчивости модели.

Преимущества:

1. Моделирование сложных зависимостей и взаимодействий между признаками.
2. Устойчивость к выбросам и способность работать с пропущенными данными.

Недостатки:

1. Возможность переобучения на небольших наборах данных.
2. Менее интерпретируемые результаты по сравнению с линейной регрессией.

Методы ансамблей.

Методы ансамблей, такие как Bagging и Boosting, используют комбинацию нескольких моделей для повышения точности прогнозов. Bagging (Bootstrap Aggregating) строит несколько моделей на случайных подвыборках данных и объединяет их результаты. Boosting последовательно обучает модели, корректируя ошибки предыдущих.

Преимущества:

1. Высокая точность за счет объединения сильных сторон отдельных моделей.
2. Способность снижать риск переобучения.

Недостатки:

1. Высокая вычислительная сложность и требовательность к ресурсам.
2. Сложность в интерпретации итоговой модели.

Нейронные сети.

Нейронные сети, вдохновленные биологическими нейронными сетями, состоят из слоев нейронов, соединенных между собой. Многослойные перцептроны (MLP) являются базовым типом нейронных сетей, способным решать задачи классификации и регрессии. Рекуррентные нейронные сети (RNN) и их разновидности, такие как LSTM, специально разработаны для обработки последовательных данных и временных рядов.

Преимущества:

1. Способность моделировать сложные нелинейные зависимости.
2. Эффективная работа с временными рядами и последовательными данными (особенно LSTM).

Недостатки:

1. Требовательность к количеству данных и вычислительным ресурсам.
2. Сложность настройки и необходимость в тщательной предобработке данных.

Процесс подготовки данных.

Сбор данных

Для прогнозирования валютных курсов требуется сбор качественных и актуальных данных. Источники данных могут включать финансовые платформы, биржи и экономические отчеты. Важно собирать данные с высокой частотой и разрешением, чтобы обеспечить точность прогнозов.

Предобработка данных

Предобработка данных включает несколько этапов:

1. Очистка данных: Удаление выбросов, заполнение пропусков и устранение аномалий.
2. Нормализация: Приведение данных к единой шкале для улучшения сходимости моделей машинного обучения.
3. Создание новых признаков: Извлечение дополнительных признаков из исходных данных, таких как скользящие средние, индикаторы волатильности и др.

Заключение.

В данной публикации были рассмотрены основные методы машинного обучения для прогнозирования валютных курсов. Линейная регрессия, решающие деревья, методы ансамблей и нейронные сети предлагают различные подходы к решению этой задачи, каждый

из которых имеет свои преимущества и ограничения.

Список литературы:

1. Нейросети и трейдинг. Практическая реализация // Хабр URL: <https://habr.com/ru/articles/494964/> (дата обращения: 09.06.2024).
2. Прогнозирование движения котировок с применением нейронных сетей // vc.ru URL: <https://vc.ru/money/323315-prognozirovanie-dvizheniya-kotirovok-s-primeneniem-neironnyh-setei> (дата обращения: 09.06.2024).
3. Машинное обучение: прогнозируем цены акций на фондовом рынке // Медиа нетологии URL: <https://netology.ru/blog/mashinnoe-obuchenie-prognoz-sen> (дата обращения: 09.06.2024).