Слайд 1.

Добрый день, уважаемые члены комиссии. Хочу предоставить вашему внимаю проект на тему «Предсказание временных рядов».

Машинное обучение уже стала новой эффективной стратегией, которую используют многие инвестиционные фонды для увеличения своих доходов.

Слайд 2.

Идея проекта заключается в том, чтобы взять несколько коррелирующий акций и предсказывать их стоимости исходя из их общей динамики.

В качестве коррелирующих акций можно подобрать акции компаний из одной отрасли, например банков.

При этом необходимо серьезно подойти к вопросу подбора этих акций. Стоит учесть страны компаний, их историю на рынке и прочее.

Слайд 3.

Для сбора были выбраны банки США.

Все данные для обучения были собраны с сайта investing.com.

Также дополнительно была собрана информация макроданные: ВВП, ставке труда, инфляция и прочее.

Далее все данные были приведены к нормальному виду, находящиеся в одних временных рамках.

Слайд 4.

На изображении продемонстрированы графики свечей банков за исследуемый период. Как видно из графиков тенденция банков приблизительно совпадают. Однако Сити банк наиболее отличен от других.

Слайд 5.

О чем и говорит матрица корреляций цен закрытия каждого из банков.

Слайд 6.

Исходя из построенных матриц корреляций таргетов и признаков для каждого банка, можно сделать выводы о необходимости и полезности каждого из признаков.

Слайд 7.

Также для проверки временных рядов на стационарность был проведен тест Дики-Фуллера.

Слайд 8.

Таким образом мы выяснили, что каждый из временных рядов целевых переменных не стационарен.

Слайд 9.

В качестве целевой переменной была выбрана - цена закрытия. Для будущего обучения моделей был выбран фреймворк ЭТНА. А именно планируется обучить модель линейной регрессии, АутоАримы, Кет буст в двух вариациях и Профит.

Слайд 10.

В дальнейшей разработке предполагается разработка интерфейса посредством телеграмм бота.

Функционал включает в себя подписку на получения предсказания стоимости акций с заданным интервалом.

Также подразумевается возможность получения стоимости заданной акции в заданный период.

Слайд 11.

Также в идеях провести бэктестирование модели на тестовых биржах, тем самым мы можем протестировать проделанную работу и подобрать наиболее выигрышную стратегию уже для торгов на реальной бирже.

P.S.

Рассматривая матрицы корреляций, можно сделать следующие выводы:

1. достаточно слабая корреляция между стоимостями акций Citi и Wells Fargo, однако другие связи достаточно сильные (например, для пары Wells Fargo & Bank of America к.к. составляет 0.84)
2. сильно коррелируют между собой признаки Real GDP, Nominal GDP и RD. такая связь возникает из-за специфики расчета данных показателей, RD и Nominal GDP имеют меньшую связь, поэтому можно попробовать создать новый признак - отклонение (процентное или абсолютное) номинального от реального ВВП
3. удалить Nominal GDP имеющий почти нулевую связь почти с каждой целевой переменной.
4. Сильную связь с Целевыми переменными имеет индекс S&P 500 (это объясняется в т.ч. тем, что данные компании входят в данный индекс)