



НАЦИОНАЛЬНЫЙ ЦЕНТР
КОГНИТИВНЫХ РАЗРАБОТОК
УНИВЕРСИТЕТА ИТМО

НЦКР



Веб-сервис предиктивной аналитики социо-экономических данных

руководство пользователя

Санкт-Петербург, 2020 г.



ОГЛАВЛЕНИЕ

1. Общие сведения.....	4
2. Функциональное назначение.....	4
3. Описание логической структуры.....	4
4. Используемые технические средства.....	5
5. Вызов и загрузка.....	5
6. Входные данные.....	5
7. Выходные данные.....	6



1. ОБЩИЕ СВЕДЕНИЯ

Веб-сервис предиктивной аналитики социо-экономических данных предназначен для исследования моделей краткосрочного прогнозирования временных рядов с применением данных из открытых источников и непосредственного построения прогнозов.

Сервис разработан на языке Python 3.6. Программное обеспечение, необходимое для функционирования программного модуля: интернет браузер (Google Chrome, Internet Explorer, Safari) с доступом в интернет.

2. ФУНКЦИОНАЛЬНОЕ НАЗНАЧЕНИЕ

Веб-сервис предиктивной аналитики социо-экономических данных реализует следующую функциональность:

- построение прогнозов временных рядов на выбранный промежуток времени;
- анализ предиктивной способности и возможной причинно-следственной связи между различными финансово-экономическими величинами;
- построение прогнозов временных рядов на выбранный промежуток времени с учётом сопутствующих факторов;
- анализ точности прогнозирования временного ряда;
- визуализация результатов прогнозирования.

3. ОПИСАНИЕ ЛОГИЧЕСКОЙ СТРУКТУРЫ

При заходе на главную страницу веб-сервиса пользователю доступны:

- поле выбора целевого временного ряда из доступных на данный момент в базе данных в виде выпадающего списка,
- поле выбора метода прогнозирования в виде переключающейся радиокнопки;
- поле выбора метода отбора рядов-предикторов в виде переключающейся радиокнопки;
- поле ввода значения в поле количества предсказываемых значений, которое принимает целочисленные значения от 1 до 6 (включительно);
- поле ввода значения в поле количества предикторов, которое принимает целочисленные значения от 0 до 49 (включительно).
- кнопка «Submit», по нажатию которой отправляются выбранные пользователем параметры на сервер, и запускается процесс расчёта.

Для выбора метода предсказания пользователю доступны следующие модели:

- **VAR** — модель векторной авторегрессии;
- **MSSA** — модель сингулярного спектрального анализа;
- **LSTM** — модель рекуррентной нейронной сети с ячейками краткосрочной долгосрочной памяти;
- **LocalApprox** — модель локальной аппроксимации временного ряда.

Для отбора оптимальных рядов предикторов доступны:

- **Granger** — метод Грейнджера;
- **Cross cor** — метод кросс-корреляции;
- **CCM** — метод конвергентного совместного представления;
- **VAR** — метод анализа качества прогноза методом векторной авторегрессии.



Расчёт по нажатию кнопки «Submit» начинается только после выбора и заполнения пользователем всех полей (выбор метода предсказания, выбор горизонта прогнозирования и выбор числа предикторов). Далее отправляется POST запрос на сервер, содержащий все выбранные параметры, и начинается расчёт, исполняющийся на удалённом сервере. Расчёт включает в себя вычисление списка предикторов в виде словарей, содержащих название датасета, временной лаг и значение метрики соответствующего теста, также включает 10 экспериментов для вычисления средних значений метрик MAE, MAPE, SMAPE, финальное предсказание согласно выбранному горизонту прогнозирования и подсчёт времени работы методов и моделей. После окончания расчёта главная страница обновляется, и пользователю становятся доступны выходные данные модели.

4. ИСПОЛЬЗУЕМЫЕ ТЕХНИЧЕСКИЕ СРЕДСТВА

Веб-сервис запускается в отдельной вкладке интернет-браузера (Google Chrome, Internet Explorer, Safari) с доступом в интернет.

5. ВЫЗОВ И ЗАГРУЗКА

Веб-сервис запускается в отдельной вкладке интернет-браузера по ссылке на сервер, на котором развёрнуто приложение Django.

Вид основной страницы веб-сервиса представлен на рисунке 5.1.

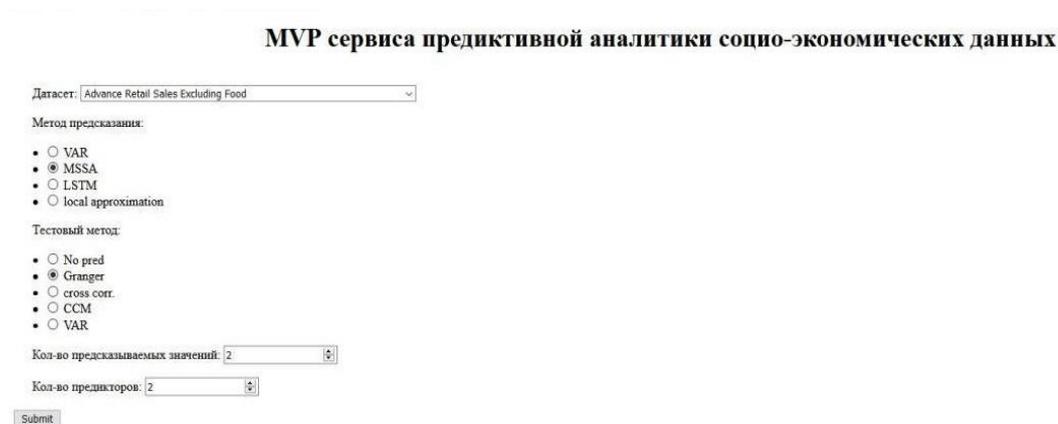


Рисунок 5.1 – Стартовое состояние основной страницы веб-сервиса

6. ВХОДНЫЕ ДАННЫЕ

Входными данными веб-сервиса являются временные ряды, выбираемые (загружаемые) пользователем сервиса в процессе взаимодействия с его интерфейсом.



Рисунок 6.1 – Пример выбора временного ряда для прогнозирования

7. ВЫХОДНЫЕ ДАННЫЕ

Выходными данными веб-сервиса являются следующие поля, содержащие агрегированную информацию о реализованном расчёте:

- Dataset — название датасета для прогнозирования;
- Method — метод предсказательного моделирования;
- Time — время выполнения в секундах;
- Predictors chosen by — выбранный метод отбора предикторов.

Если выбрано использование предикторов и количество предикторов больше 0, то выводится список предикторов в виде:

- Lag — длина интервала, оптимизирующего выбранный статистический метод выбора предиктора;
- Score — значение соответствующей статистики.

Если количество предикторов равно 0, то выдается сообщение: No predictor.

Также пользователю становятся доступны метрики качества построенной модели прогнозирования со значениями трёх различных метрик качества:

- MAE=<значение> (средняя абсолютная ошибка);
- MAPE=<значение> (средняя относительная ошибка);
- SMAPE=<значение> (симметричная средняя относительная ошибка);
- и значения сформированного прогноза на выбранный промежуток времени: <дата> <значение>.

Пример выходных данных модели представлен на рисунке 7.1.



Результаты прогнозирования

Datascraper Advance Retail Sales Excluding Food. Метод VAR

Time 2.1486 s. Predictors chosen by Granger

lag=6, score=0.0000, Production Electricity for China

lag=3, score=0.0000, Consumer Price Index for South Africa

Time 0.1406 s. Method examination result:

mean values in 10 cases: MAE=37940; MAPE=8.27%; SMAPE=8.42%

Time 0.0150 s. Prediction result:

2019-03 430137

2019-04 392514

2019-05 468769

2019-06 450104

Рисунок 7.1 – Пример выходных данных веб-сервиса

Также с правой стороны страницы генерируется график, на котором отображаются истинные значения временного ряда («Real»), предсказанные значения на тестовой выборке («Predicted»), а также предсказанные значения на тренировочной выборке («Train»), если выбраны методы MSSA или LSTM.

Пример сформированного графика представлен на рисунке 7.2.

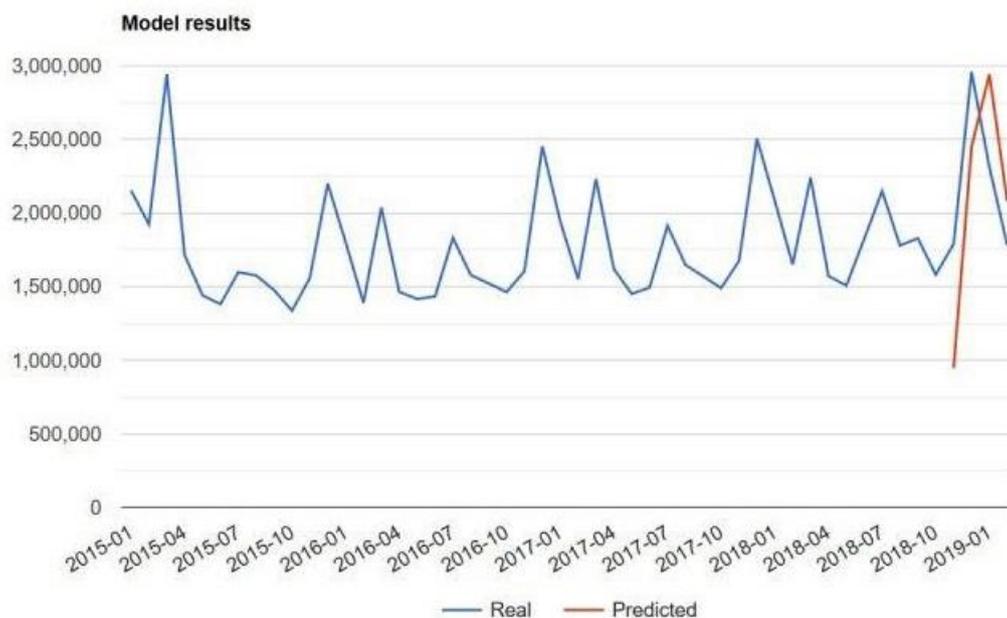


Рисунок 7.2 – Пример визуализации результатов работы модели MSSA для прогнозирования временного ряда веб-сервисом предиктивной аналитики социо-экономических данных