МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение высшего образования «Балтийский федеральный университет имени Иммануила Канта»

ОТЧЁТ О ПРОХОЖДЕНИИ

УЧЕБНОЙ ПРАКТИКИ (НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ (ПОЛУЧЕНИЕ ПЕРВИЧНЫХ НАВЫКОВ НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЫ)

на базе Высшей школы компьютерных наук и прикладной математики образовательно-научного кластера «Институт высоких технологий» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Выполнила Бабина Полина Валерьевна

(ФИО обучающегося, курс, форма обучения)

Направление подготовки/специальность 02.04.03 «Математическое обеспечение и администрирование информационных систем»

(код, наименование)

Руководитель практики от университета Ткаченко Сергей Николаевич, к.т.н., руководитель цифровой кафедры ОНК «ИВТ», доцент БФУ им. И. Канта

Руководитель практики от профильной организации\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(ФИО, должность)

г. Калининград 2023 г

Оглавление

[Введение 3](#_Toc137078168)

[Глава 1. Исследование методов анализа временных рядов 4](#_Toc137078169)

[1.1. Гребневая регрессия (Ridge regression) 4](#_Toc137078170)

[1.2. Регрессия Лассо (Lasso regression) 6](#_Toc137078171)

[1.3. Дерево решений (Decision Tree) 8](#_Toc137078172)

[Глава 2. Анализ временного ряда при помощи машинного обучения 11](#_Toc137078173)

[2.1. Обработка данных 11](#_Toc137078174)

[2.2. Ridge Regression 15](#_Toc137078175)

[2.3. Lasso Regression 17](#_Toc137078176)

[2.4. Деревья решений 19](#_Toc137078177)

[Вывод 22](#_Toc137078178)

[Список литературы 24](#_Toc137078179)

**Введение**

Прогнозирование временных рядов является важной задачей в области анализа данных и машинного обучения. Оно находит широкое применение в различных сферах, таких как финансы, экономика, климатология, маркетинг и других, где требуется предсказание будущих значений на основе прошлых данных. Точность прогнозов временных рядов имеет большое значение для принятия стратегических решений, планирования и оптимизации ресурсов.

В данной практической работе мы сосредоточимся на анализе и сравнении трех методов прогнозирования временных рядов: Ridge Regression, Lasso Regression и Дерево решений. Эти методы представляют собой различные подходы к моделированию и прогнозированию временных рядов, и мы исследуем их производительность на основе ряда метрик, таких как коэффициент детерминации (R^2), средняя абсолютная процентная ошибка (MAPE) и среднеквадратичная ошибка (MSE).

Целью данного исследования является выявление наиболее эффективного метода прогнозирования временных рядов и определение его применимости в реальных сценариях. Результаты нашего исследования позволят понять, какая модель может достичь наиболее точных прогнозов временных рядов и какие факторы следует учитывать при выборе метода прогнозирования.

Для достижения этой цели необходимо выполнить следующие задачи: мы представим подробный обзор каждого метода, описывая их принципы работы, особенности и применение. Затем мы представим результаты экспериментов, сравнивая производительность каждого метода на основе выбранных метрик. В заключении мы подведем итоги и сделаем обобщенные выводы о применимости и эффективности рассмотренных методов прогнозирования временных рядов.

.

**Глава 1. Исследование методов анализа временных рядов**

**1.1. Гребневая регрессия (Ridge regression)**

Гребневая регрессия является одним из методов регуляризации, применяемых в статистике и машинном обучении, и используется для управления переобучением моделей регрессии.

Гребневая регрессия основана на добавлении регуляризационного члена к функции потерь в задаче регрессии. Регуляризация помогает сдерживать рост коэффициентов модели и уменьшать их значимость. В гребневой регрессии применяется L2-регуляризация, которая приравнивает квадраты коэффициентов к сумме квадратов остатков.

Основное преимущество гребневой регрессии заключается в способности снижать влияние мультиколлинеарности. Мультиколлинеарность возникает, когда признаки модели сильно коррелируют между собой, что может приводить к неустойчивым оценкам коэффициентов. Гребневая регрессия позволяет уменьшить влияние коррелированных признаков путем уменьшения их весов, что приводит к более стабильным и робастным оценкам.

Однако, гребневая регрессия также имеет свои ограничения. Во-первых, она не способна автоматически отбирать признаки, как это делает, например, Lasso-регрессия. Все признаки остаются включенными в модель с ненулевыми весами, хотя и сниженной значимостью. Если важно сузить модель до наиболее информативных признаков, то гребневая регрессия может быть менее подходящим выбором.

Кроме того, выбор значения гиперпараметра регуляризации в гребневой регрессии может быть сложной задачей. Гиперпараметр контролирует степень сжатия коэффициентов и может влиять на баланс между смещением и разбросом модели. Оптимальное значение гиперпараметра может быть найдено с помощью методов кросс-валидации или других подходов оптимизации.

В целом, гребневая регрессия является полезным инструментом для управления мультиколлинеарностью и создания более стабильных моделей регрессии. Она может быть особенно полезной, когда в данных присутствуют сильно коррелированные признаки. Однако, для задач отбора признаков или в случае требования автоматической регуляризации, другие методы, такие как Lasso-регрессия, могут быть более подходящими. [2]

Применение гребневой регрессии (Ridge regression) к временным рядам имеет свои особенности и преимущества. Временные ряды обладают зависимостью между значениями в разные моменты времени, и гребневая регрессия может помочь учесть эту зависимость.

Одним из подходов к применению гребневой регрессии к временным рядам является включение лаговых переменных в модель. Вместо использования исходных признаков временного ряда, можно использовать их лаговые значения на предыдущих временных шагах. Это позволяет учесть зависимость между значениями ряда в разные моменты времени и уловить временные зависимости.

При применении гребневой регрессии к временным рядам также может быть полезным нормирование признаков. Нормирование признаков помогает уравновесить их вклад в модель и избежать смещения в сторону признаков с большей дисперсией.

Однако, при использовании гребневой регрессии для временных рядов необходимо учитывать особенности данных. Например, если в ряде присутствуют тренды или сезонные компоненты, то гребневая регрессия может не справиться с их моделированием эффективно.

Таким образом, гребневая регрессия может быть полезным инструментом для моделирования временных рядов, особенно в случаях, когда присутствует зависимость между значениями ряда на разных временных шагах. Однако, для более точного учета временных зависимостей и особенностей временных рядов, могут потребоваться специализированные методы, которые учитывают динамику временных рядов более эффективно.

**1.2. Регрессия Лассо (Lasso regression)**

Lasso является одним из методов регуляризации, применяемых в статистике и машинном обучении, и широко используется для отбора признаков и построения моделей прогнозирования.

Основной идеей Lasso является добавление регуляризационного члена к функции потерь в задаче регрессии. Регуляризация позволяет сдерживать рост коэффициентов модели и уменьшать их значения. Lasso-регрессия использует L1-регуляризацию, которая приравнивает модули коэффициентов к сумме модулей остатков.

Одним из преимуществ Lasso-регрессии является возможность автоматического отбора признаков. Это достигается за счет свойства L1-регуляризации, которое приводит к установлению некоторых коэффициентов признаков в ноль. Таким образом, Lasso позволяет определить наиболее важные признаки, исключив незначимые.

Lasso может быть эффективным методом в случаях, когда количество признаков превышает количество наблюдений, или когда признаки сильно коррелируют друг с другом. В таких ситуациях Lasso может помочь справиться с проблемой мультиколлинеарности и избежать переобучения модели.

Однако, есть некоторые ограничения Lasso. Например, Lasso не способен улавливать зависимости между значениями признаков в разные моменты времени в случае анализа временных рядов. Для моделирования временных зависимостей и работы с временными рядами часто применяются специализированные методы, такие как авторегрессионные модели или рекуррентные нейронные сети.

В целом, Lasso-регрессия представляет собой мощный инструмент для отбора признаков и построения прогностических моделей, особенно в случаях с большим количеством признаков и проблемой мультиколлинеарности. Однако, в зависимости от контекста и типа данных, может потребоваться использование более специализированных методов. [1]

Lasso-регрессия в контексте временных рядов имеет свои особенности и ограничения. Во-первых, временные ряды обладают зависимостью между значениями в разные моменты времени, и эта зависимость может быть не учтена при использовании обычного Lasso.

Одним из способов учета временных зависимостей в Lasso-регрессии является включение в модель лаговых переменных. То есть, вместо использования исходных признаков временного ряда, используются их лаговые значения на предыдущих временных шагах. Это позволяет учесть зависимость между значениями ряда в разные моменты времени.

Однако, включение большого числа лаговых переменных может привести к проблеме размерности и увеличению сложности модели. Поэтому важно выбирать оптимальное количество и комбинацию лагов, чтобы модель была информативной и не переобучалась.

Еще одним ограничением Lasso-регрессии в контексте временных рядов является чувствительность к выбросам. Если во временном ряду присутствуют выбросы, которые сильно влияют на оценку коэффициентов модели, то Lasso может не справиться с их сдерживанием эффективно. В таких случаях может быть полезно применение других методов регуляризации или алгоритмов, устойчивых к выбросам.

Применение Lasso-регрессии к временным рядам может быть полезным для отбора признаков и упрощения модели, особенно в случаях с большим количеством признаков.

**1.3. Дерево решений (Decision Tree)**

Дерево решений — это графическая модель, которая используется в задачах классификации и регрессии для принятия решений на основе набора признаков.

"Корень" дерева представляет собой начальный узел, от которого отходят ветви к другим узлам. "Ветви" соединяют узлы дерева. Каждая ветвь представляет собой специфическое условие или правило, основанное на значении одного из признаков данных. "Листья" дерева представляют собой конечные узлы, где принимается окончательное решение.

Дерево решений строится поэтапно. Признаки, которые наилучшим образом разделяют данные и дают наибольший прирост информации, выбираются для построения ветвей. Этот процесс продолжается до тех пор, пока не будет достигнут критерий останова, например, достижение определенной глубины дерева или недостаточного количества объектов в узле.

Одним из основных преимуществ является их интерпретируемость. Деревья решений позволяют понять, какие признаки наиболее важны для принятия решений и какие условия приводят к определенным классам или значениям целевой переменной. Они также могут обрабатывать различные типы данных и автоматически учитывать взаимодействия между признаками.

Однако, деревья решений также имеют свои ограничения. Они могут быть склонны к переобучению, особенно при построении глубоких деревьев. Это может привести к недостаточной обобщающей способности модели. Для борьбы с переобучением используются различные методы, такие как обрезка дерева, использование ансамблевых методов (например, случайный лес) и регуляризация.

Деревья решений являются мощным инструментом для принятия решений в задачах классификации и регрессии. Они обладают интерпретируемостью и способностью работать с различными типами данных. Однако, для достижения более высокой обобщающей способности и устойчивости к переобучению, могут использоваться дополнительные методы и техники. [3]

Применение деревьев решений к временным рядам имеет свои особенности и требует учета специфики данных. Временные ряды характеризуются зависимостью между значениями в разные моменты времени, и деревья решений могут быть адаптированы для учета этой зависимости.

Одним из способов применения деревьев решений к временным рядам является использование лаговых переменных. Вместо использования исходных признаков временного ряда, можно использовать значения ряда на предыдущих временных шагах в качестве признаков. Это позволяет учесть зависимость между значениями ряда в разные моменты времени и уловить временные закономерности.

При построении дерева решений для временных рядов также можно использовать различные критерии разделения, учитывающие временную структуру данных. Например, можно использовать различные метрики для измерения сходства между временными рядами на разных временных шагах, такие как автокорреляция или схожесть формы кривых.

Однако, при применении деревьев решений к временным рядам необходимо учитывать некоторые ограничения. Во-первых, деревья решений могут быть чувствительны к шуму и выбросам в данных временного ряда, что может привести к переобучению. Поэтому рекомендуется проводить предварительную обработку данных и удаление выбросов.

Кроме того, деревья решений могут быть недостаточно гибкими для моделирования сложных временных зависимостей и трендов. В таких случаях могут быть более эффективными специализированные методы, такие как авторегрессионные модели (ARIMA), GARCH или рекуррентные нейронные сети (RNN), которые учитывают динамику временных рядов более точно.

В итоге, деревья решений могут быть полезным инструментом для моделирования временных рядов, особенно при использовании лаговых переменных и учете временной структуры данных. Однако, для более точного учета сложных временных зависимостей и особенностей временных рядов, могут потребоваться специализированные методы, которые более полно учитывают временную природу данных.

Для прогнозирования временного ряда будут использованы все вышеперечисленные модели машинного обучения.

**Глава 2. Анализ временного ряда при помощи машинного обучения**

**2.1. Обработка данных**

Первым шагом в процессе создания моделей машинного обучения была обработка данных. Загружаем файл "data.csv" в объект pandas DataFrame и выводим строки таблицы (Рис. 1). Для наглядности отображаем исходный временной ряд (Рис. 2).

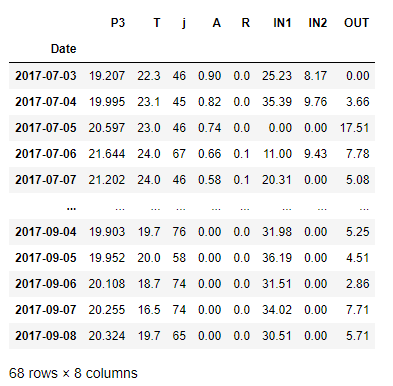


Рисунок 1. Исходные данные

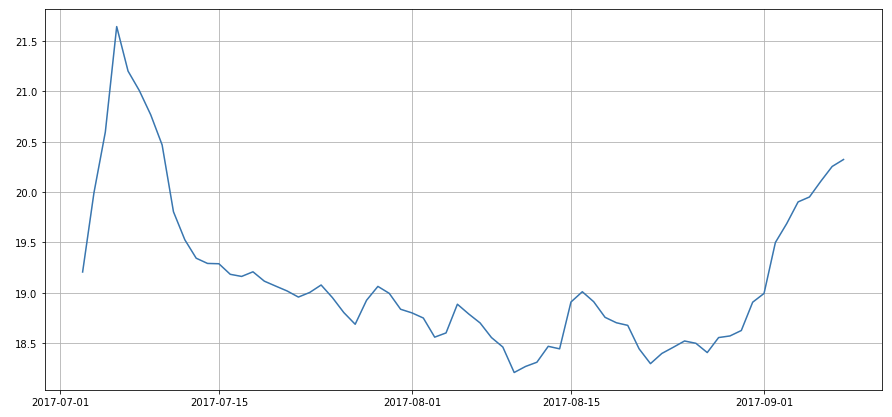


Рисунок 2. Исходный временной ряд

Корреляция переменных: Каждая ячейка матрицы показывает числовое значение корреляции между соответствующей парой переменных. Значения корреляции варьируются от -1 до 1, где -1 указывает на сильную отрицательную корреляцию, 1 - на сильную положительную корреляцию, и 0 - на отсутствие корреляционной связи.

Интерпретация корреляции: По графику можно определить, какие переменные имеют высокую положительную или отрицательную корреляцию между собой. Если две переменные имеют положительную корреляцию, это означает, что они имеют тенденцию изменяться в одном направлении. Если две переменные имеют отрицательную корреляцию, это означает, что они имеют тенденцию изменяться в противоположных направлениях.

По данным корреляционный матрицы и гистограммы можно увидеть, что наибольшая корреляция с целевой переменно P3 имеет A, R, T (Рис.3).

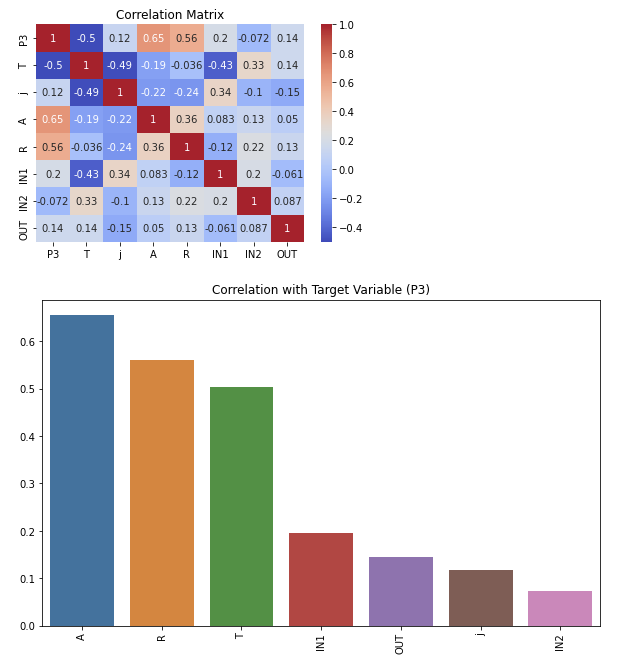


Рисунок 3. Корреляция

Анализ корреляционной матрицы и гистограммы позволяет выделить переменные A, R и T как наиболее значимые признаки, связанные с целевой переменной P3. Они могут оказывать значительное влияние на P3.

График коэффициентов модели регрессии представляет собой визуализацию значений коэффициентов, полученных в результате обучения модели регрессии.

Каждый столбец на графике представляет отдельный фактор (или переменную) модели регрессии, а высота столбца указывает на величину коэффициента, который отражает влияние соответствующего фактора на целевую переменную.

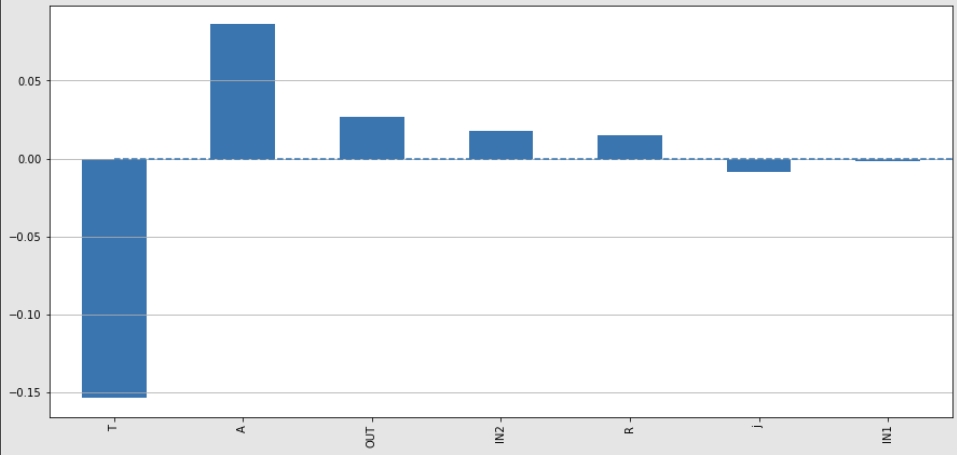


Рисунок 4. График коэффициентов модели регрессии

Наиболее положительное влияние: Факторы, представленные положительными коэффициентами, оказывают наибольшее положительное влияние на целевую переменную. Самый значимый из них - фактор, соответствующий самому высокому положительному коэффициенту. В данном случае, это фактор "A".

Наиболее отрицательное влияние: Факторы, представленные отрицательными коэффициентами, оказывают наибольшее отрицательное влияние на целевую переменную. Самый значимый из них - фактор, соответствующий самому низкому отрицательному коэффициенту. В данном случае, это фактор "T".

Нулевые коэффициенты: Горизонтальная линия на уровне нуля указывает на факторы, коэффициенты которых равны нулю или очень близки к нулю. Это означает, что эти факторы не имеют значимого влияния на целевую переменную. В данном случае, это фактор "INT1".

Далее выполним предварительную обработку данных и подготовку выборки для прогнозирования (Рис.5). Разделим признаки и целевую переменную, а также разделим на обучающую (Рис.6) и тестовую выборки (Рис.7), где использованы значения на последние 12 дней. Для нормализации признаковых данных масштабируем данные.

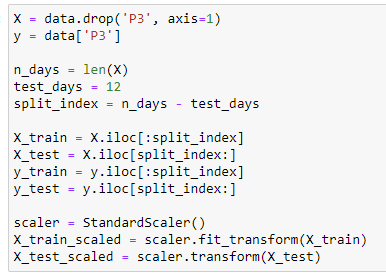


Рисунок 5. Подготовка данных

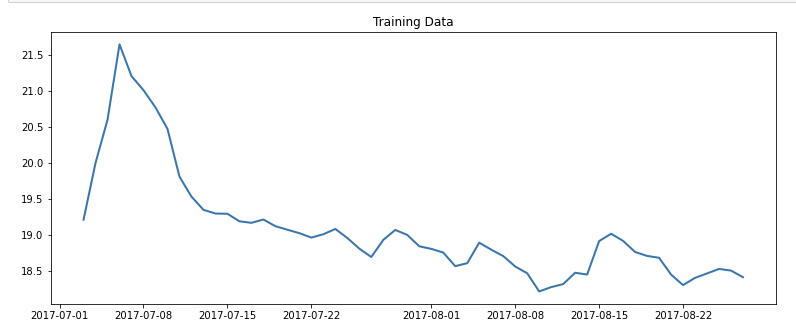


Рисунок 6. График обучающих данных

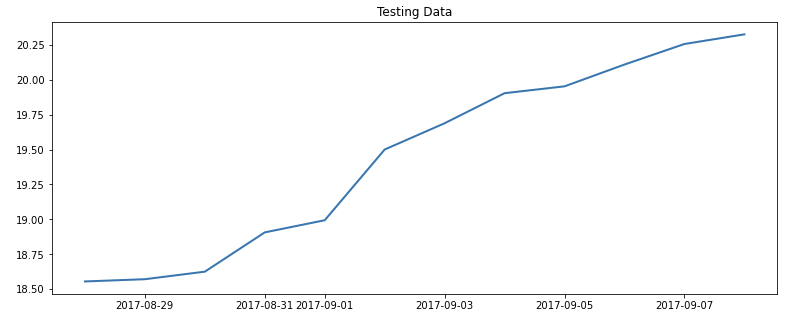


Рисунок 7. График тестовых данных

**2.2. Ridge Regression**

Для построения модели регрессии используется алгоритм Ridge Regression. Создаем модель регрессии Ridge с параметром alpha = 60 (Рис.8). Этот параметр выбран как наиболее оптимальный. Данный параметр поможет модели достичь лучшей производительности и точности при прогнозировании.

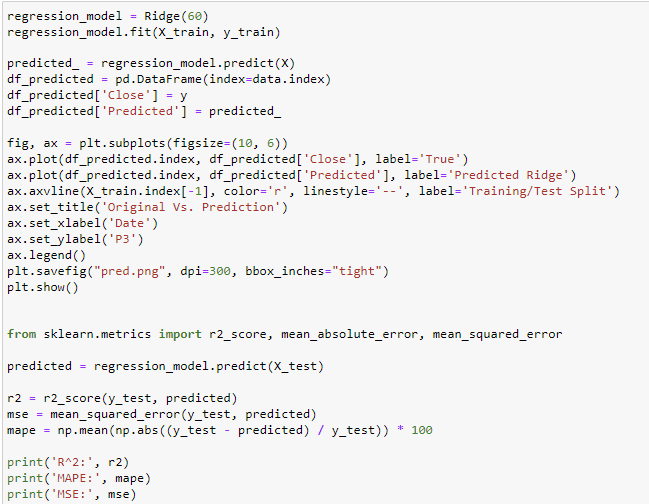


Рисунок 8. Ridge Regression

После обучения модели используются тестовые данные для предсказания значений целевой переменной. Модель использует обученные параметры для вычисления прогнозов на основе входных признаков (Рис.9).

Для оценки качества модели используются различные метрики, такие как:

1. коэффициент детерминации (R^2). Коэффициент детерминации является мерой, которая показывает, какую часть вариации зависимой переменной (в данном случае целевой переменной P3) объясняет модель. Он принимает значения от 0 до 1, где 0 означает, что модель не объясняет вариацию целевой переменной, а 1 означает, что модель полностью объясняет вариацию. Чем ближе значение (R^2) к 1, тем лучше модель;
2. средняя абсолютная процентная ошибка (MAPE). MAPE измеряет среднюю абсолютную процентную разницу между прогнозируемыми значениями и фактическими значениями временного ряда. Эта метрика учитывает процентное отклонение и может быть полезной, если важно оценить точность прогноза в процентном отношении.
3. среднеквадратическая ошибка (MSE). MSE представляет собой среднее значение квадрата разницы между предсказанными значениями и истинными значениями целевой переменной. MSE измеряет среднеквадратичную величину ошибки предсказания и часто используется в регрессионных моделях. Чем ниже значение MSE, тем лучше модель.

Эти метрики позволяют оценить точность и предсказательную способность модели на тестовых данных.

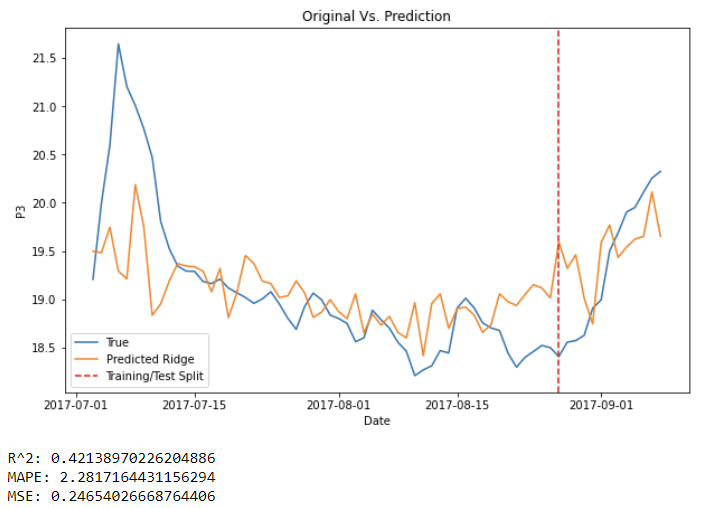


Рисунок 9. График и метрики Ridge Regression

Модель Ridge Regression показывает относительно низкое значение (R^2), что указывает на то, что модель не смогла полностью улавливать сложности и вариации в данных. Однако, она демонстрирует хорошую точность в прогнозировании с низкой средней абсолютной процентной ошибкой (MAPE) и среднеквадратической ошибкой (MSE).

**2.3. Lasso Regression**

Для построения модели регрессии используется алгоритм Lasso Regression. Создаем модель регрессии Lasso с параметром alpha = 0.2 (Рис. 10). Этот параметр выбран как наиболее оптимальный. Данный параметр поможет модели достичь лучшей производительности и точности при прогнозировании.



Рисунок 10. Lasso Regression

После обучения модели используются тестовые данные для предсказания значений целевой переменной. Модель использует обученные параметры для вычисления прогнозов на основе входных признаков (Рис.11).

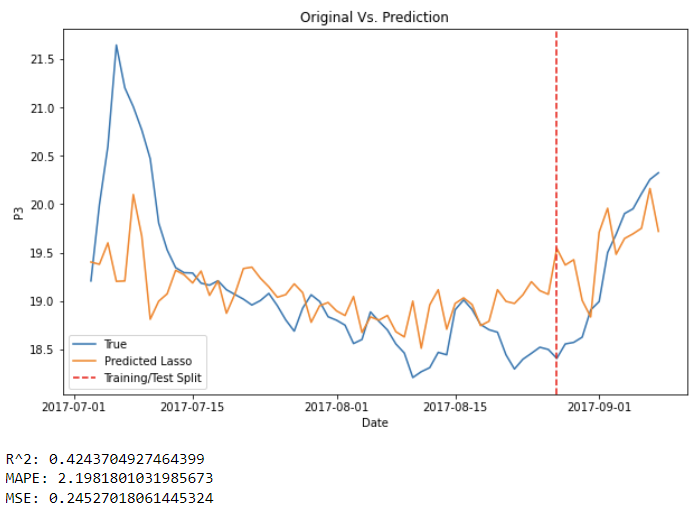


Рисунок 11. График и метрики Lasso Regression

Модель Lasso Regression показывает относительно низкое значение (R^2), что указывает на то, что модель не смогла полностью улавливать сложности и вариации в данных. Однако, она демонстрирует хорошую точность в прогнозировании с низкой средней абсолютной процентной ошибкой (MAPE) и среднеквадратической ошибкой (MSE).

**2.4. Деревья решений**

Выполняем настройку и оценку модели регрессии на основе дерева решений с использованием метода GridSearchCV для подбора оптимальных гиперпараметров (Рис.12).



Рисунок 12. Дерево решений

Задается диапазон значений для гиперпараметров модели в переменной param\_grid. В данном случае определены следующие гиперпараметры:

* max\_depth - максимальная глубина дерева решений.
* min\_samples\_split - минимальное количество образцов, необходимых для разделения узла.
* min\_samples\_leaf - минимальное количество образцов, необходимых для создания листа.
* max\_features - количество признаков, рассматриваемых при каждом разделении.

Создается объект GridSearchCV, который принимает модель, диапазон параметров и указание на количество складок для кросс-валидации (в данном случае cv=5).

Получаем лучшие параметры модели и создаем новый объект модели регрессии на основе дерева решений с их использованием. Наилучшие параметры, полученные после настройки модели регрессии на основе дерева решений с использованием GridSearchCV, состоят из следующих значений:

* max\_depth: 5
* max\_features: 'log2'
* min\_samples\_leaf: 1
* min\_samples\_split: 5

Эти параметры выбраны как наиболее оптимальные на основе кросс-валидации. Данные параметры помогут модели достичь лучшей производительности и точности при прогнозировании.

Обучаем модель и прогнозируем на тестовых данных с использованием обученной модели.

Вычисляем метрики для оценки качества модели, такие как коэффициент детерминации (R^2), среднеквадратическая ошибка (MSE) и средняя абсолютная процентная ошибка (MAPE).

Строим график, на котором отображаются фактические значения и предсказания модели (Рис.13).

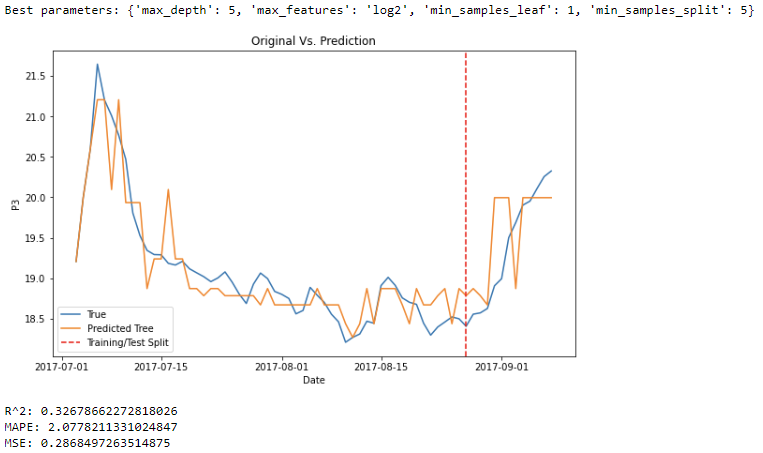


Рисунок 13. График и метрики дерева решений

Модель дерева решений, оцененная по данным метрикам, показывает небольшое значение (R^2), что указывает на недостаточную способность модели объяснить вариативность в данных. Однако, она демонстрирует хорошую точность с низкой средней абсолютной процентной ошибкой (MAPE) и среднеквадратической ошибкой (MSE). Это означает, что модель достаточно хорошо прогнозирует значения временного ряда с относительно небольшими отклонениями.

**Вывод**

Исходя из анализа трех моделей прогнозирования - Ridge Regression, Lasso Regression и Дерево решений, можно сделать следующие выводы, для данной временной задачи.

Ridge Regression и Lasso Regression показали сопоставимую производительность в предсказании временных рядов. Обе модели демонстрировали значение коэффициента детерминации (R^2) около 0.42, что указывает на то, что примерно 42% вариативности в данных можно объяснить с помощью этих моделей. Кроме того, средняя абсолютная процентная ошибка (MAPE) для обеих моделей составила около 2.2-2.3%, что говорит о низком уровне ошибки прогнозирования. Также стоит отметить, что среднеквадратичная ошибка (MSE) для обеих моделей была невысокой, около 0.24-0.25. Эти результаты свидетельствуют о том, что Ridge Regression и Lasso Regression являются надежными методами прогнозирования временных рядов и могут быть использованы для достижения высокой точности прогнозов.

С другой стороны, модель Дерева решений демонстрировала немного более низкую производительность по сравнению с линейными моделями. Значение коэффициента детерминации (R^2) для Дерева решений составило около 0.33, что указывает на то, что около 33% вариативности в данных может быть объяснено данной моделью. Однако, средняя абсолютная процентная ошибка (MAPE) для Дерева решений составила около 2.08%, что говорит о небольшой ошибке прогнозирования. Стоит отметить, что среднеквадратичная ошибка (MSE) для Дерева решений была немного выше, около 0.29. В целом, модель Дерева решений обладает потенциалом для прогнозирования временных рядов, но ее производительность немного ниже по сравнению с линейными моделями.

Таким образом, основываясь на проведенном анализе, можно заключить, что Ridge Regression и Lasso Regression представляют собой эффективные методы прогнозирования временных рядов, которые показывают высокую точность прогнозов. Дерево решений также может быть использовано для прогнозирования временных рядов, хотя его производительность немного ниже. Выбор конкретной модели следует осуществлять с учетом требований и контекста конкретной задачи прогнозирования.

|  |  |  |  |
| --- | --- | --- | --- |
|  | R^2 | MAPE | MSE |
| Ridge | 0.42138970226204886 | 2.2817164431156294 | 0.24654026668764406 |
| Lasso | 0.4243704927464399 | 2.1981801031985673 | 0.24527018061445324 |
| Дерево решений | 0.32678662272818026 | 2.0778211331024847 | 0.2868497263514875 |

Таблица 1. Общие результаты метрик



Рисунок 14. Общий график

**Список литературы**

1. Лассо-регрессия // Большая Российская Энциклопедия URL: https://bigenc.ru/c/lasso-regressiia-79d732 (дата обращения: 05.06.2023).
2. Гребневая регрессия // Большая Российская Энциклопедия URL: https://bigenc.ru/c/grebnevaia-regressiia-9aa84b (дата обращения: 05.06.2023).
3. Для чего начинающим аналитикам нужны деревья решений // Блог Яндекс Практикума URL: https://practicum.yandex.ru/blog/chto-takoe-derevo-reshenii-kak-ego-postroit/ (дата обращения: 05.06.2023).
4. Forecasting Timeseries Using Machine Learning & Deep Learning// medium URL: https://medium.com/mlearning-ai/forecasting-timeseries-using-machine-learning-deep-learning-446eccc6eb6d (дата обращения: 05.06.2023).
5. Expand your Time Series Arsenal with These Models // towardsdatascience URL: https://towardsdatascience.com/expand-your-time-series-arsenal-with-these-models-10c807d37558 (дата обращения: 05.06.2023).