

УДК 621.311.153.001.24

**ПРОГНОЗИРОВАНИЕ ГРАФИКОВ НАГРУЗКИ
PREDICTION OF LOAD SCHEDULES**

К.А. Марчук

Научный руководитель – Попкова Н.А., старший преподаватель
Белорусский национальный технический университет,

г. Минск, Республика Беларусь

popkova@bntu.by

Marchuk K.A.

Supervisor – Popkova N.A., Senior Lecturer

Belarusian national technical university, Minsk, Belarus

***Аннотация:** В работе рассматривается возможность применения моделей машинного обучения для решения задачи прогнозирования графиков электрической нагрузки. Выявлены закономерности в исходных и полученных данных, проведено сравнение их с действительными, произведен поиск лучшей модели на основании выбранной метрики.*

***Abstract:** The paper considers the possibility of using machine learning models to solve the problem of predicting electrical load curves. Regularities in the initial and received data are revealed, they are compared with the real ones, and the best model is searched based on the selected metric.*

***Ключевые слова:** графики нагрузки, корреляционная матрица, RMSLE.*

***Keywords** load graphs, correlation matrix, correlation, RMSLE.*

Введение

Предсказание будущих событий — перспективная, но и интересная задача во многих сферах: начиная с прогноза погоды и заканчивая предсказанием роста акций, курса валют. Машинное обучение уже сегодня позволяет значительно сократить время и трудозатраты на принятие управленческих решений, например, ввод дополнительного числа генерирующих устройств или же вывод. Прогнозирование графиков нагрузки дает возможность изначально предсказать загрузку генерирующих аппаратов и распределить загрузки между ними таким образом, чтобы себестоимость электроэнергии была минимальна [1].

Достижению высокой точности прогнозирования мешают много факторов:

- внезапное изменение температуры;
- отсутствие возможности учета случайных событий;
- зависимость от вида дня;
- точность предсказанной погоды;

Для прогнозирования графиков нагрузки используются следующие виды данных: история потребления и прогноз температуры. Для начала обучения необходимо проверить данные на аномалии, наиболее подходящим решением является корреляционная матрица в виде тепловой карты. Данная матрица показывает, как переменные в наборе данных связаны друг с другом и как они перемешаются относительно друг друга. Значение корреляции колеблется от -1

до +1. Нулевой коэффициент корреляции означает, что две переменные не зависят друг от друга. Положительная корреляция указывает на то, что переменные движутся в одном направлении, а отрицательная корреляция указывает на противоположное [2].

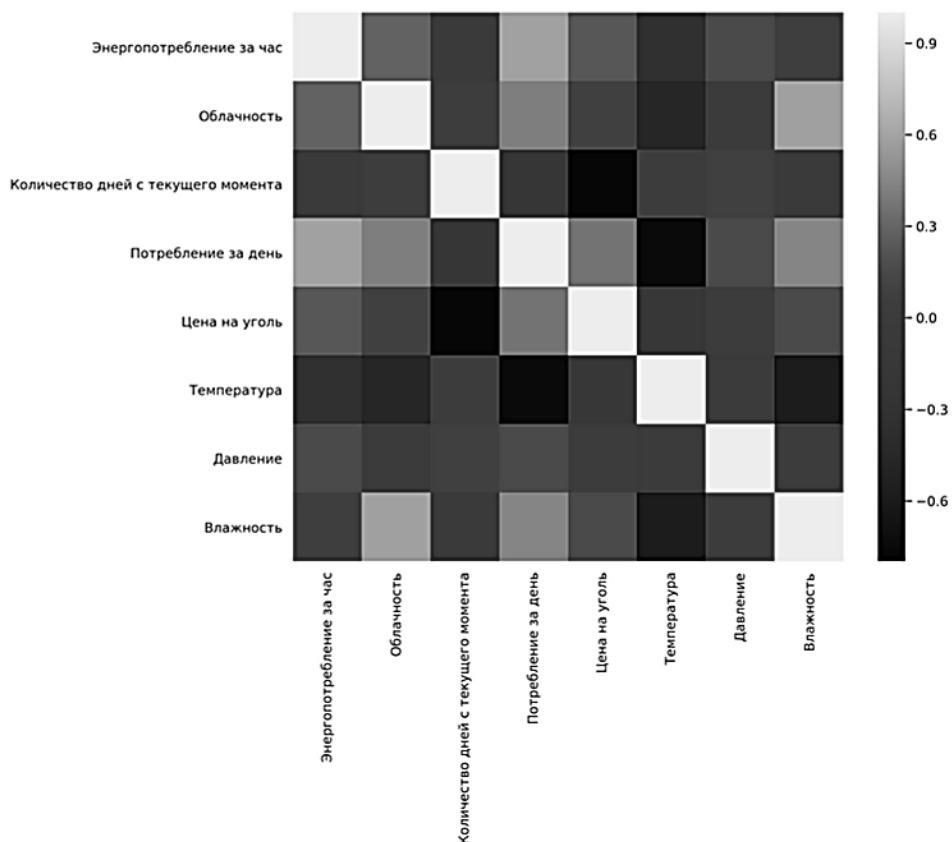


Рисунок 1 - Корреляционная матрица в виде тепловой карты

После предварительного просмотра данных, стоит задача с определением зависимостей и проверки этих зависимостей с действительностью. Как правило, потребление электроэнергии в зимний и осенний период больше, чем в весенний и летний, что и отражено на графиках энергопотребления показанных на рисунке 2.

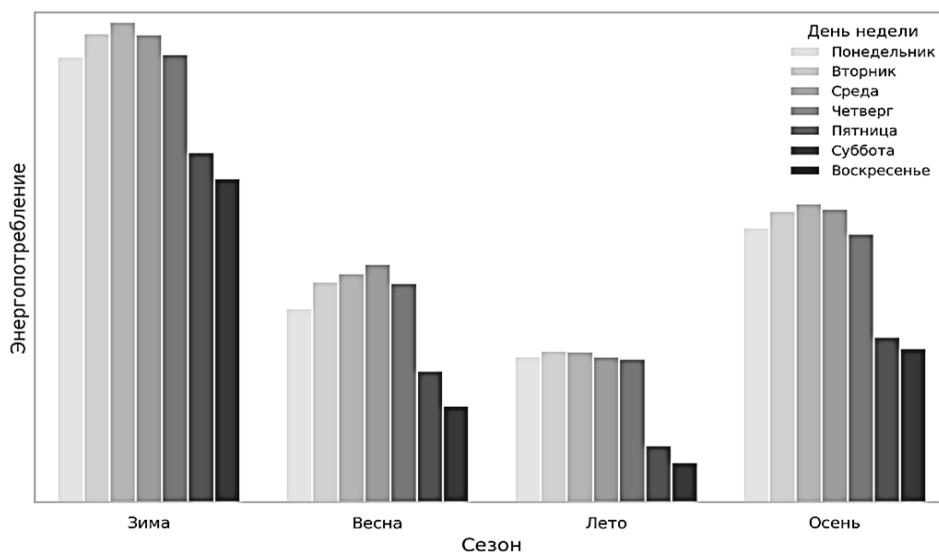


Рисунок 1 - Потребление электроэнергии от времени года

После обработки данных и установки всех зависимостей ставиться задача: выбора модели и метрики, по которой будет проводиться оценка модели. В качестве метрики используют r^2 (R - коэффициент детерминации, показывающий, как сильно условная дисперсия модели отличается от дисперсии реальных значений) и RMSLE (Root Mean Squared Logarithmic Error, или среднеквадратичная логарифмическая ошибка, которая является расстоянием между двумя точками на плоскости) [3].

Модель будет подобрана путем перебора моделей по лучшему качеству, а лучшие параметры для модели будут определены с помощью встроенных функций, результаты точности предсказания каждой модели приведены в таблице 1.

Таблица 1 – Показатели качества для каждой модели

Метрика	LightGBM	Catboost	XGBoost	Prophet	SARIMAX
r^2	0.94137	0.93984	0.92909	0.81435	0.73778
MSLE	0.02468	0.02477	0.01219	0.00829	0.00658

Как видно из таблицы 1, что лучшая модель LightGBM, результаты прогноза представлены на рисунке 3.

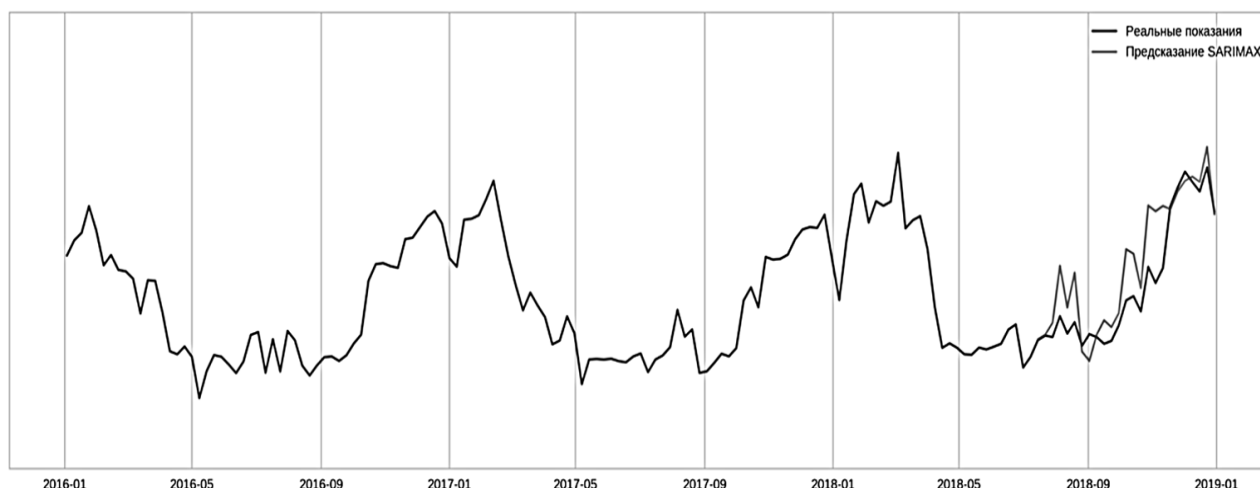


Рисунок 3 — График прогноза потребления

В крупных городах, как правило, есть почти линейная зависимость потребляемой мощности от температуры наружного воздуха. Если потребление не отвечает этим сложившимся графикам, значит вмешался какой-то фактор, который невозможно предсказать.

Вывод

На основании вышеизложенного можно сделать вывод, что недоотпуск электроэнергии потребителям в случае возникновения аварии приводил к серьезной погрешности исследуемой модели. Наиболее перспективным вариантом будет проведение предсказания нагрузки по районам, предприятиям, кварталам города добавив к уже имеющимся признакам, новые, которые будут отвечать за вероятность появления аварии, а для нахождения суммарной нагрузки, нужно будет обучить модель на суммарных предсказаниях других моделей.

Литература

1. Управление нагрузками // [электронный ресурс] – Режим доступа: <https://forca.ru/knigi/arhivy/ekspluatatsiya-elektricheskikh-sistem-3.html>. Дата доступа: 14.04.2022.
2. Тепловая карта корреляции в Seaborn // [электронный ресурс] – Режим доступа: <https://www.delftstack.com/ru/howto/seaborn/correlation-heatmap-seaborn-python>. Дата доступа: 14.04.2022.
3. Машинное обучение в энергетике, или не только лишь все могут смотреть в завтрашний день // [электронный ресурс] – Режим доступа: <https://habr.com/ru/company/lanit/blog/487944>. Дата доступа: 14.04.2022.