

ПРОГНОЗИРОВАНИЕ ЭЛЕКТРИЧЕСКОЙ НАГРУЗКИ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

Хэ Цинчжоу

Томский политехнический университет, ИШИП, 8ЕМ21, e-mail: cinchzhou1@tpu.ru

Аннотация

В контексте современного общества, значение электроэнергии для функционирования страны непрерывно возрастает. В рамках данной работы представлена разработка интегрированной модели машинного обучения. Данная модель синтезирует метеорологические, временные и прочие релевантные характеристики с целью эффективного моделирования и прогнозирования нагрузки на электросеть.

Ключевые слова: прогнозирование нагрузки на электросеть, машинное обучение, lstm, mape.

Введение

В современном обществе электроэнергия приобретает всё большую важность, что стимулирует появление новых энергетических предприятий и ведет к увеличению производства и потребления электроэнергии. Следовательно, актуализируется проблема прогнозирования нагрузки на электростанции, требующая срочного решения [1].

В рамках данной работы проведен эксплораторный анализ данных для определения методологии среднесрочного прогнозирования электрической нагрузки. В процессе анализа аномальные данные были обработаны и нормализованы, а временная шкала изменена с часового интервала на дневной. Кроме того, метеорологические факторы были определены с использованием тепловой карты и анализа значимости, что позволило эффективно уменьшить объем данных. В качестве тестового набора выбраны данные за последние два месяца, тогда как остальные данные служили обучающим набором.

Прогнозирование электрической нагрузки и предварительная обработка данных

Прогнозирование электрической нагрузки заключается в определении будущих тенденций изменения электрической нагрузки и потребления электроэнергии, исходя из исторических данных о нагрузке [2].

Для анализа в данной работе используются данные, полученные с платформы Kaggle. Информация об электрической нагрузке включает в себя сведения о потреблении энергии отдельными домохозяйствами. Метеорологические данные, полученные из Датского метеорологического института (DMI), охватывают 10 различных метеорологических параметров, как указано в таблице 1, при этом все они представлены в числовой форме. Сбор данных происходит каждый час, обеспечивая 24 измерения в день.

Таблица 1

Таблица типов метеорологических характеристик

Метеорологическое состояние	Описания	Единица
pressure_at_sea	Атмосферное давление на среднем уровне моря	hPa
precip_dur_past10min	Количество минут осадков за последний 10-минутный период	minutes
wind_speed	Измерение 10-метровой средней скорости ветра над местностью	m/s
temp_dew	Температура точки росы на высоте 2 м над уровнем моря	°C
Pressure	Атмосферное давление на уровне станции	hPa
visib_mean_last10min	Последняя 10-минутная средняя видимость	m

temp_dry	Температура воздуха на высоте 2 метра	°C
Humidity	Относительная влажность местности	%
cloud_cover	Облачный покров	%
visibility	повышенная видимость	m

Оценка важности признаков методом Random Forest

При построении модели дерева решений, важность признаков может быть определена с помощью параметра `feature_importances` [3]. В методе случайного леса для каждого дерева отбираются нецелевые данные, чтобы рассчитать их ошибку. Затем к признакам всех образцов добавляется шум, и ошибка пересчитывается. Значительное изменение в значении ошибки указывает на то, что признак существенно влияет на предсказание, что позволяет оценить его важность.

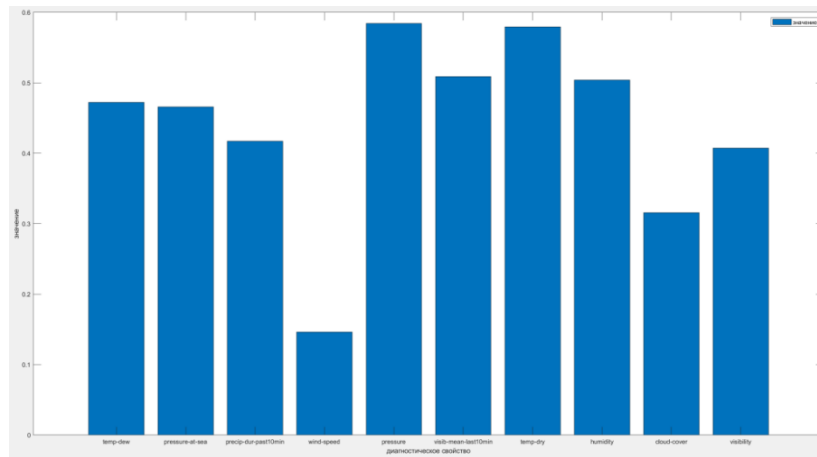


Рис. 1. Важность признаков Random Forest

Из рисунка следует, что признаки “pressure, temp_dew, humidity, visib-mean-last10min, и temp_dry” обладают высокой значимостью, в то время как значимость признака “wind-speed” ниже 0.2 и значительно уступает другим, поэтому “wind-speed” исключается из рассмотрения.

Модель PSO-LSTM

Алгоритм PSO-оптимизации для LSTM нейронных сетей включает следующие шаги:

1. Загрузка выборочных данных и определение параметров роя частиц, включая вес инерции, количество итераций и коэффициенты обучения.
2. Инициализация нейронной сети LSTM с использованием выборочных данных.
3. Расчет значения адаптации и выбор среднеквадратичной ошибки LSTM нейронной сети в качестве функции приспособленности.

$$E = \frac{1}{M} \sum_{l=1}^M \sum_{o=1}^q (y_o - d)^2 \quad (1)$$

В данном контексте: m - обозначает количество обучающих выборок; q - количество узлов на выходном слое сети; $y_o = [y_{o1}, y_{o2}, \dots, y_{oq}]^T$ представляет фактический выход сети; $d = [d_1, d_2, \dots, d_q]^T$ – это целевой выход сети. Частица стремится найти минимум фитнес-функции, что соответствует оптимальной работе нейронной сети LSTM.

4. Индивидуальное оптимальное положение каждой частицы ($pbest$) и общее оптимальное положение роя ($gbest$) определяются на основе вычисления фитнес-функции, согласно уравнению (1), что ведет к регулированию положений и скоростей частиц до достижения минимизации фитнес-функции и определения оптимального положения.

На графике сравниваются прогнозируемые значения двух моделей: красные линии отображают реальные значения электрической нагрузки, черные линии - прогнозы модели PSO-LSTM, а синие линии - прогнозы модели LSTM.

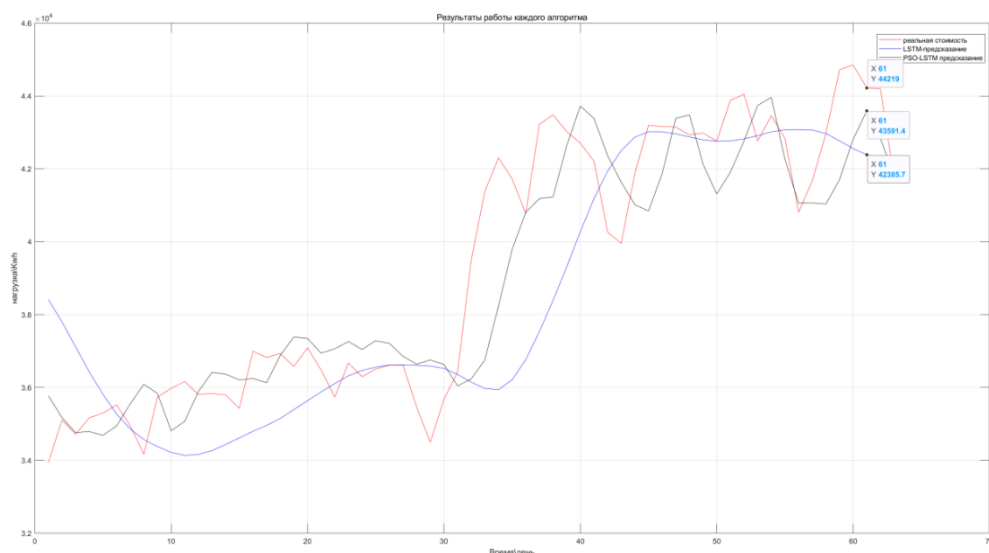


Рис. 2. Результаты прогнозирования электрической нагрузки

Таблица 2

Анализ результатов модели PSO-LSTM и LSTM

Моделирование	MAPE	R2	MAE	RMSE
PSO-LSTM	2.54%	0.867	1.009	1.300
LSTM	3.83%	0.647	1.513	2.115

На рис. 2 показано, что прогнозируемые и фактические значения, полученные с использованием интегрированной модели, совпадают по направлению, что демонстрирует преимущество этого подхода в ситуациях, где одиночная модель машинного обучения оказывается недостаточно эффективной для предсказания экстремальных значений.

Заключение

В данной статье исследуется применение моделей машинного обучения для анализа данных о электрической нагрузке. Анализ показал, что использование одиночной модели машинного обучения ограничено в точности прогнозирования и не в состоянии адекватно отразить нелинейные характеристики данных о нагрузке. В ответ на это была разработана интегрированная модель PSO-LSTM. Результаты показывают, что PSO-LSTM превосходит традиционные подходы, становясь предпочтительным вариантом для среднесрочного прогнозирования электрической нагрузки.

Список использованных источников

1. Ibrahim Bibi, Rabelo Luis, GutierrezFranco Edgar, ClavijoBuritica Nicolas. Machine Learning for Short-Term Load Forecasting in Smart Grids[J]. Energies. – 2022. – Vol. 15(21). – P. 8079–8079.
2. Mingjian Cui, Jianhui Wang, Meng Yue «Machine Learning-Based Anomaly Detection for Load Forecasting Under Cyberattacks», IEEE TRANSACTIONS ON SMART GRID. – 2019.
3. Muhammad Sajjad, Zulfiqar Ahmad Khan, Amin Ullah, Tanveer Hussain, Waseem Ullah, Mi Young Lee, Sung Wook Baik «A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting», IEEE ACCESS. – 2020.