

НЕЙРОСЕТЕВОЕ ПРОГНОЗИРОВАНИЕ ГЕНЕРАЦИИ ЭЛЕКТРОЭНЕРГИИ СОЛНЕЧНЫМИ ПАНЕЛЯМИ

С.М. СТЕПАНОВ, Н.А. ИСКРА

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

Поступила в редакцию 20 декабря 2017

Аннотация. Предметом исследования данной статьи является анализ влияния применения различных методов регрессии на качество краткосрочного предсказания генерации электрической энергии солнечными панелями. Для решения задачи предсказания выбраны многослойный перцептрон и деревья принятия решений. При постановке эксперимента используются реальные данные о генерации электрической энергии. Наилучший показатель коэффициента детерминации составил 0,94.

Ключевые слова: солнечная энергия, модель предсказания, регрессия, многослойный перцептрон, деревья принятия решений.

Abstract. The main purpose of this paper is analysis of various regression methods application on quality of short-term solar PV forecasting. Multilayer perceptron and decision trees were chosen in order to solve prediction problem. Real historical data on solar PV forecasting are used as experimental datasets. The best coefficient of determination was 0.94.

Keywords: solar power, prediction model, regression, multilayer perceptron, decision trees.

Doklady BGUIR. 2018, Vol. 113, No. 3, pp. 26-31
Neural network forecasting of energy generation by solar panels
S.M. Stepanov, N.A. Iskra

Введение

Прогнозирование генерации электрической энергии солнечными панелями требует знаний об активности солнца, условиях окружающей среды, атмосферных явлениях, процессе рассеивания солнечной энергии, и технических характеристиках солнечной электростанции. Выходное напряжение солнечной панели зависит от мощности светового потока, падающего на солнечную панель, и технических характеристик солнечной панели. Информация с предсказаниями генерации электроэнергии важна для эффективной работы солнечной электростанции, а также для эффективного управления электрической сетью и продажами электроэнергии энергетическим компаниям.

Проблема генерации электрической энергии тесно связана с проблемой прогнозирования погодных условий и атмосферных явлений. В связи с этим выделяют 3 типа прогнозирования генерации электрической энергии солнечными панелями: краткосрочное (до 3–4 ч); среднесрочное (до 7 дней); долгосрочное (от 1 месяца до 1 года).

Краткосрочное прогнозирование генерации электроэнергии обычно используется специалистами для мониторинга стабильности энергосети на солнечных электростанциях, которые имеют возможность изменять количество электрической энергии, отдаваемой в энергосеть. Наиболее удобным является прогноз на ближайшие 15–30 мин. Для такого прогноза очень важно иметь точный прогноз погоды за целевой период. Для этого используют несколько техник.

1. Статистические техники, базирующиеся на обработке временных рядов метеорологических данных, которые используются в качестве данных для обучения и изменения параметров модели. В данном случае используются модель авторегрессии скользящего среднего (ARMA, ARIMA), нейронные сети, метод опорных векторов и т. д. [1, 2].

2. С момента запуска спутников наблюдения за атмосферными явлениями, например, Meteosat, появилась возможность производить краткосрочный прогноз на основе спутниковых снимков и их автоматического анализа [3]. В таком случае краткосрочный прогноз происходит практически в режиме реального времени.

Среднесрочный прогноз также используется специалистами для мониторинга энергосети и планирования количества электроэнергии, которую солнечная электростанция может отдать в энергосеть. Такой прогноз базируется на различных прогностических метеорологических моделях для получения значений погодных характеристик. Наиболее популярными системами на основе таких моделей являются Global Forecast System (GFS) и European Center of Medium Range Weather Forecasting (ECMWF).

Долгосрочный прогноз используется для предсказания количества электрической энергии, которую солнечная электростанция может произвести за месяц или год. Такими прогнозами пользуются руководства солнечных электростанций и энергетических компаний для обсуждения возможных контрактов друг с другом.

Целью данной статьи является исследование влияния применения различных методов регрессии на качество краткосрочного прогнозирования. В рамках данной статьи в качестве модели предсказания будут использоваться регрессионные модели на базе многослойного персептрона и деревьев принятия решений [4].

Многослойный персептрон

Персептрон – простейшая форма нейронной сети, предназначенная для классификации линейно-разделимых сигналов (т. е. образы можно разделить некоторой гиперплоскостью) [5]. Персептрон состоит из нейронов с настраиваемыми синаптическими весами и порогами. Также используется для прогнозирования результатов экспериментов на основе обучающей выборки. Структура однослойного персептрона представлена на рис. 1.



Рис. 1. Схема однослойного персептрона

Однослойный персептрон состоит из трех типов элементов: S , A , R . S -элементы представляют собой входные данные. A -элементы называются ассоциативными. Вес связи S - A может принимать значения -1 , 0 , $+1$. R -элемент называется сумматором, который подсчитывает сумму значений входных сигналов, принимая во внимание вес связи A - R . Однослойный персептрон с одним нейроном ограничен выполнением задачи разделения только двух классов. При увеличении количества нейронов, количество классов, которые может разделить классификатор увеличивается. Каждый нейрон имеет свою функцию активации, которая определяет присутствие или отсутствие сигнала.

Обучение классификатора – процесс настройки весов связей и порогов с целью уменьшения разности между желаемыми и получаемыми значениями на выходе. Классический метод обучения персептрона – метод коррекции ошибки. При таком методе вес связи не меняется до тех пор, пока текущая реакция классификатора остается правильной. При появлении неправильной реакции вес изменяется на шаг коррекции (единицу), а знак определяется противоположным от знака ошибки.

В некоторых случаях классы не являются линейно-разделимыми. Для таких задач обычно используется многослойный персептрон. Многослойный персептрон – это однослойный персептрон с дополнительными скрытыми слоями (нейронами). Схема многослойного персептрона показана на рис. 2.

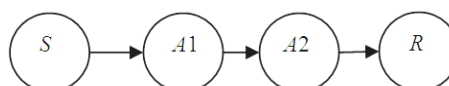


Рис. 2. Схема многослойного персептрона

В данном случае каждый из скрытых слоев выполняет свою часть по классификации входных параметров. При использовании персептрона в качестве регрессора ступенчатая функция активации нейрона заменяется на непрерывную функцию активации. В результате, персептрон выдает не бинарный результат (0 или 1), а реальное число.

Деревья принятия решений

Деревья принятия решений – средство поддержки принятия решений, использующееся в статистике и анализе данных для моделей прогнозирования [6]. Структура дерева представляет собой «листья» и «ветки». На ребрах («ветках») дерева записаны атрибуты, от которых зависит целевая функция. В «листьях» записаны значения целевой функции. Основная цель – это создать модель, которая предсказывает значения целевой переменной на основе нескольких переменных на входе.

Деревья принятия решений делятся на 2 типа:

- 1) дерево для классификации, когда предсказываемый результат является классом. Примером такого дерева может быть классификация солнечных электростанций;
- 2) дерево для регрессии, когда предсказываемый результат можно рассматривать как вещественное число, например, рентабельность солнечной электростанции.

Пример дерева принятия решения для классификации солнечных электростанций на основе количества солнечных дней, доступной площади и географической позиции представлен на рис. 3.

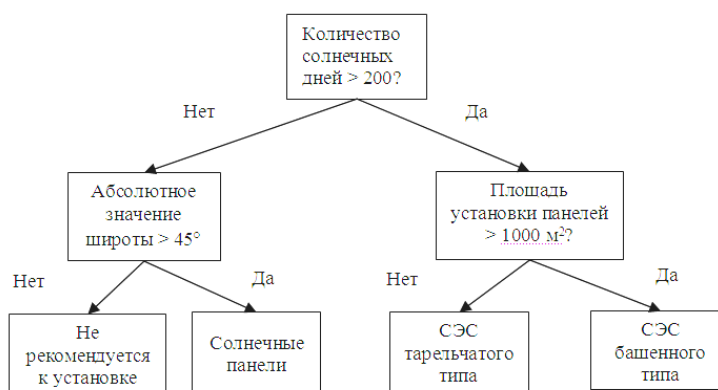


Рис. 3. Дерево принятия решений

Подготовка данных для эксперимента

Вектор информативных признаков для построения модели предсказания содержит следующие параметры: облачность, температура окружающей среды, коэффициент солнечной активности. Параметр предсказания – количество электрической энергии, произведенной солнечной панелью.

Облачность – это количество облаков в данный момент времени. Может иметь значения от нуля до единицы, где ноль – ясная погода, а единица – небо полностью затянуто тучами и солнечный свет не доходит до поверхности земли. Например, при значении облачности 0,9, 90 % неба покрыто облаками, что значительно уменьшает мощность светового потока, который падает на солнечную панель.

Температура окружающей среды – это мгновенное значение температуры окружающей среды, измеряемое в градусах Цельсия (°C). Максимальная температура, присутствующая в обучающей и тестовой выборке – +34,06 °C, минимальная температура – –9,25 °C.

Коэффициент солнечной активности – это значение, которое показывает активность солнца в данный момент времени. Может иметь значения от нуля до единицы, где ноль – солнце находится за горизонтом (от заката до рассвета). Максимальное значение этого коэффициента достигает после двух с половиной часов после рассвета и начинает уменьшаться за два с половиной часа до заката. График изменения коэффициента солнечной активности представлен на рис. 4. В рассматриваемом случае рассвет произошел в 7 ч утра, а закат – в 18:30.

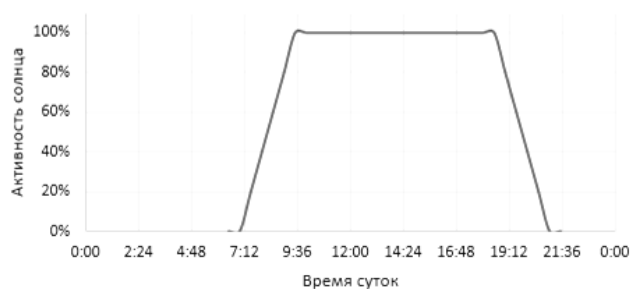


Рис. 4. График изменения коэффициента солнечной активности

Количество электрической энергии, произведенное солнечной панелью, измеряется в ваттах (Вт). В обучающей и тестовой выборке количество энергии измеряется в МВт. Данные для проведения экспериментов получены с официального сайта бельгийского подразделения европейской энергетической компании ELIA для Брюсселя за 2013–2016 годы. Максимальное значение – 38,99 МВт, минимальное – 0,01 МВт за весь период. Размер обучающей выборки – 71030 элементов, а размер тестовой выборки – 71 элемент.

Наиболее широко используемой метрикой правильности обучения регрессоров является коэффициент детерминации (*R*-квадрат) – это доля дисперсии зависимых переменных, объясняемая построенной моделью. Другими словами, этот коэффициент показывает насколько хорошо построенная модель соответствует тем данным, на основе которых она обучалась. Этот коэффициент принимает значения от нуля до единицы, причем ноль – модель не соответствует данным, единица – модель полностью соответствует данным [7]. Вектор информативных признаков и вектор параметра предсказания представлены в табл. 1.

Таблица 1. Векторы информативных признаков и параметра предсказания

Название признака	Значение
Облачность	0,56
Температура окружающей среды	22,2
Коэффициент солнечной активности	1,0
Количество электрической энергии, произведенной солнечной панелью	35,72

В дополнение к этому, данные были редуцированы: строки с нулевым значением количества сгенерированной энергии были удалены из обучающей и тестовой выборки. Такие значения приходились на ночные часы (от захода и до рассвета, т. е. при коэффициенте активности солнца, равном нулю).

Проведение экспериментов

Для постановки и проведения экспериментов использовалась библиотека машинного обучения *scikit-learn*. В данном случае использовался отдельный дистрибутив под названием *Anaconda 2* [8], который содержит в себе все необходимые зависимости и пакеты для работы не только с нейронными сетями, но и другими областями машинного обучения.

В процессе эксперимента многослойный перцептрон и дерево принятия решений обучались на одной и той же выборке данных, но с различными настройками сети (например, функция активации или максимальная глубина дерева). После обучения вычислялся коэффициент детерминации для заданных параметров с помощью метода *score*, реализация которого уже присутствует в библиотеке для каждого типа регрессора.

При создании объекта перцептрона (*MLPRegressor*) необходимо задать параметры, которые напрямую влияют на точность работы нейронной сети при прогнозировании новых данных. Могут использоваться следующие параметры.

1. *Hidden_layer_sizes* – количество скрытых слоев и количество нейронов в каждом слое. В данном эксперименте количество скрытых слоев – 1, количество нейронов – 3.

2. *Activation* – функция активации для *A*-элементов скрытых слоев нейронной сети. Может принимать значения:

- *identity* – отсутствие функции активации, т. е. $f(x) = x$;
- *logistic* – логистическая сигмоидная функция, т. е. $f(x) = 1 / (1 + \exp(-x))$;
- *tanh* – гиперболический тангенс, т. е. $f(x) = \tanh(x)$;
- *relu* – *rectifier linear unit* – функция рампы, т. е. $f(x) = \max(0, x)$.

3. Solver – метод оптимизации весов связей нейронов. Может принимать значения:
 – lbfgs – алгоритм семейства квази-ньютоновских методов;
 – sgd – стохастический градиентный спуск;
 – adam – алгоритм, основанный на стохастическом градиентном спуске, предложенный исследователями Kingma, Diederik и Jimmy Ba [9].

Результаты эксперимента представлены в табл. 2.

Таблица 2. Точность прогноза многослойного персептрона с различными параметрами

№ п/п	Функция активации	Метод оптимизации	Коэффициент детерминации
1	identity	lbfgs	0,49
2	identity	sgd	0,49
3	identity	adam	0,48
4	logistic	lbfgs	0,56
5	logistic	sgd	0,57
6	logistic	adam	0,58
7	tanh	lbfgs	0,56
8	tanh	sgd	0,57
9	tanh	adam	0,57
10	relu	lbfgs	0,56
11	relu	sgd	0,55
12	relu	adam	0,58

При создании объекта DecisionTreeRegressor необходимо задать 2 параметра, которые влияют на точность прогнозирования новых данных. В эксперименте используются следующие параметры.

1. Criterion – критерий остановки при построении дерева. Может принимать значения:

- mse – mean square error – среднеквадратичное отклонение;
- friedman_mse – среднеквадратичное отклонение, улучшенное исследователем J.H. Friedman [10];
- mae – среднее абсолютное отклонение.

2. Max_depth – максимальная глубина дерева. Если это значение не задано, дерево будет расти неограниченно, т. е. нет никаких ограничений на глубину дерева.

Результаты эксперимента представлены в табл. 3.

Таблица 3. Точность прогноза дерева принятия решений с различными параметрами

№ п/п	Критерий	Максимальная глубина	Коэффициент детерминации
1	mse	10	0,57
2	mse	20	0,84
3	mse	30	0,94
4	mse	40	0,94
5	mse	None	0,94
6	friedman_mse	10	0,57
7	friedman_mse	20	0,84
8	friedman_mse	30	0,94
9	friedman_mse	40	0,94
10	friedman_mse	None	0,94
11	mae	10	0,55
12	mae	20	0,81
13	mae	30	0,93
14	mae	40	0,94
15	mae	None	0,94

Заключение

Наивысший коэффициент детерминации показывает дерево принятия решений для любого из критериев остановки построения дерева – 0,94. Такой результат характеризует очень хорошую построенную модель с высоким уровнем точности прогнозов. Наилучший коэффициент детерминации многослойного персептрона – 0,58 для сигмоидной функции активации и улучшенного алгоритма градиентного спуска для расчета весов связей нейронов. Такой коэффициент характеризует модель средней точности с высокой степенью вероятности ошибок. Таким образом, деревья принятия решений лучше подходят для краткосрочного предсказания генерации электрической энергии солнечными панелями.

Многослойный персептрон восприимчив к «выбросам», т. е. данным, которые сильно отличаются от общей выборки значений. Это может быть одной из возможных причин низкого коэффициента детерминации многослойного персептрона. Второй возможной причиной низкого

коэффициента может быть относительно небольшой объем данных для обучения. Для улучшения коэффициента детерминации многослойного персептрона предлагается убрать «шумные» данные из обучающей выборки и увеличить ее объем.

Список литературы

1. Huang R., Huang T., Gadh R. Solar Generation Prediction using the ARMA Model in a Laboratory-level Microgrid // IEEE Third International Conference on Smart Grid Communications. 2012. P. 528–533.
2. Bacher P. Short-term Solar Power Forecasting // Technical University of Denmark. 2008. P. 109.
3. Leung I.J.H., Jordan J.E. Image processing for weather satellite cloud segmentation // Electrical and Computer Engineering. 1996. P. 953–956.
4. Fine-grained photovoltaic output prediction using Bayesian Ensemble / Prithwish Chakraborty [et al.] // AAAI. 2012. P. 7.
5. Хайкин С. Нейронные сети. Полный курс. М.: Вильямс, 2006. 1104 с.
6. Tom M.M. Machine Learning. Boston, Massachusetts, 1997. 52 p.
7. Patrick K.W., Sonja S.T. A practical introduction to econometric methods: classical and modern. University of the West Indies Press, 2002. 307 p.
8. Anaconda [Электронный ресурс]. – URL: <https://www.anaconda.com> (дата обращения: 01.12.2017).
9. Diederik P.K., Jimmy B.A. A method of stochastic optimization // 3rd ICLR. San Diego, 2015. P. 15.
10. Friedman J.H. Greedy Function Approximation: A Gradient boosting machine // IMS Reitz Lecture. 1999. P. 44.

References

1. Huang R., Huang T., Gadh R. Solar Generation Prediction using the ARMA Model in a Laboratory-level Microgrid // IEEE Third International Conference on Smart Grid Communications. 2012. P. 528–533.
2. Bacher P. Short-term Solar Power Forecasting // Technical University of Denmark. 2008. P. 109.
3. Leung I.J.H., Jordan J.E. Image processing for weather satellite cloud segmentation // Electrical and Computer Engineering. 1996. P. 953–956.
4. Fine-grained photovoltaic output prediction using Bayesian Ensemble / Prithwish Chakraborty [et al.] // AAAI. 2012. P. 7.
5. Hajkin S. Nejrornyie seti. Polnyj kurs. M.: Vil'jams, 2006. 1104 s. (in Russ.)
6. Tom M.M. Machine Learning. Boston, Massachusetts, 1997. 52 p.
7. Patrick K.W., Sonja S.T. A practical introduction to econometric methods: classical and modern. University of the West Indies Press, 2002. 307 p.
8. Anaconda [Electronic resource]. – URL: <https://www.anaconda.com> (access date: 01.12.2017).
9. Diederik P.K., Jimmy B.A. A method of stochastic optimization // 3rd ICLR. San Diego, 2015. P. 15.
10. Friedman J.H. Greedy Function Approximation: A Gradient boosting machine // IMS Reitz Lecture. 1999. P. 44.

Сведения об авторах

Степанов С.М., магистрант кафедры электронных вычислительных машин Белорусского государственного университета информатики и радиоэлектроники.

Искра Н.А., старший преподаватель кафедры электронных вычислительных машин Белорусского государственного университета информатики и радиоэлектроники.

Адрес для корреспонденции

247210, Республика Беларусь,
г. Жлобин, мкр. 3, д. 50, кв. 7
тел. +375-2334-3-06-33;
e-mail: siargei.stepanov@gmail.com
Степанов Сергей Михайлович

Information about the authors

Stepanov S.M., master student of electronic computing machines department of Belarusian state university of informatics and radioelectronics.

Iskra N.A., senior lecturer of electronic computing machines department of Belarusian state university of informatics and radioelectronics.

Address for correspondence

247210, Republic of Belarus,
Zhlobin, microdistrict 3, 50–7;
tel. +375-2334-3-06-33;
e-mail: siargei.stepanov@gmail.com
Stepanov Sergey Michailovich