

## **АЛГОРИТМЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ДЛЯ ПРЕДСКАЗАНИЯ КОСМИЧЕСКОЙ ПОГОДЫ**

И.В. Улитин<sup>1,2</sup>, М.А. Барулина<sup>1,3</sup>

<sup>1</sup>Институт проблем точной механики и управления РАН, г. Саратов

<sup>2</sup>Саратовский национальный исследовательский университет им. Н.Г. Чернышевского

<sup>3</sup>Пермский государственный национальный исследовательский университет

[ulitin.ivan3004@mail.ru](mailto:ulitin.ivan3004@mail.ru)

Прогноз космической погоды является актуальной задачей, которая может быть решена с помощью анализа данных, получаемых со спутников. Геомагнитные бури вызываются взаимодействием солнечного ветра с магнитным полем Земли. Возникающие в результате возмущения геомагнитного поля могут нанести ущерб системам GPS, спутниковой связи, передаче электроэнергии и т. д. Эти возмущения измеряются индексом времени бури возмущения, или индексом Dst, который представляет собой отклонение вариации магнитного поля от спокойного уровня, усредненное по значениям, измеренным на контрольной цепочке расположенных в низких широтах магнитных станций. Индекс Dst вычисляется в Мировом центре данных (World Data Center) в Киото (Япония) на основе спутниковых данных о параметрах солнечного ветра и межпланетного магнитного поля. При этом прогнозирование значений Dst на ближайшие часы или дни является важной задачей, так как может способствовать уменьшению степени негативного воздействия геомагнитных бурь на объекты техносферы. **Цель настоящей статьи** - исследовать возможность применения различных алгоритмов искусственного интеллекта в задаче прогнозирования параметра Dst. Полученные результаты обучения далее проанализировать, сравнить и сформулировать на их основе выводы.

Исследуемые данные получены с двух спутников, NASA ACE и NOAA DSCOVR [1, 2]. Эти данные о космической погоде включают показания датчиков, относящиеся как к межпланетному магнитному полю, так и к плазме солнечного ветра. Межпланетное магнитное поле (ММП) играет огромную роль в том, как солнечный ветер взаимодействует с магнитосферой Земли.

Во время солнечного минимума магнитное поле Солнца является диполем. Дипольное поле Солнца имеет силу около 50 Гс, в то время как магнитное поле Земли примерно в 100 раз слабее. В период максимальной солнечной активности на солнечном диске появляются крупные пятна, которые являются областями выхода сильных магнитных полей. Магнитное поле Солнца является динамичным и сложным по своей структуре в силу того, что мощность силовых линий пятен в несколько сотен раз превышает мощность силовых линий фонового магнитного поля. Магнитное поле Солнца переносится солнечным ветром через Солнечную систему, пока не достигает гелиопаузы. Поскольку Солнце поворачивается вокруг своей оси, межпланетное магнитное поле имеет форму спирали, которая называется Спиралью Паркера. Важным параметром, включенным в набор данных, является величина  $B_t$  межпланетного магнитного поля, которая обозначает его общую мощность. Умеренные значения напряженности межпланетного магнитного поля начинаются с 15 нТ, однако для средних широт желательно 25 нТ и более. Межпланетное магнитное поле представляет собой векторную величину с трехосевой составляющей, две из которых ( $B_x$  и  $B_y$ ) ориентированы параллельно эклиптике. Компоненты  $B_x$  и  $B_y$  не важны

для авроральной активности, но величина  $V_z$  перпендикулярна эклиптике и создается волнами и другими возмущениями солнечного ветра.

Используемый в работе набор данных представляет собой записанные спутниками с частотой раз в минуту такие параметры космической погоды, как объемная скорость солнечного ветра (solar wind bulk speed), температура ионов солнечного ветра (solar wind ion temperature), плотность протонов солнечного ветра (solar wind ion temperature), геоцентрические координаты солнечной магнитосферы (geocentric solar magnetospheric coordinates), геоцентрические координаты солнечной эклиптики (geocentric solar ecliptic coordinates) и т.д. Набор данных содержит 8392320 записей. При этом записей, для которых определен индекс Dst - 139872. Это связано с тем, что Dst вычислялся раз в час, в то время как остальные параметры записывались раз в минуту. Надо отметить, что значение индекса Dst может быть как отрицательным, когда возмущение направлено к Земле, так и положительным. В безветренных условиях значения Dst примерно равны 0.

Рассматриваемые данные были проанализированы с использованием инструментария языка Python. Для обучения моделей были сконструированы новые показатели на основе представленных в наборе по аналогии с тем, как было описано в [3]. После обработки данных со спутников количество элементов выборки, используемой для обучения, стало соответствовать количеству имеющихся значений Dst - 139872. Далее набор данных был разделен на обучающий, тестовый и проверочный наборы, которые составили соответственно 82/12/6 процентов от исходного множества.

Для предсказания значений индекса Dst были обучены следующие алгоритмы - XGBoost [4], Random Forest [5], Decision Tree [6] и Gradient Boosting [7]. Отметим, что в некоторых научных работах алгоритм случайного леса (Random Forest) считается наиболее предпочтительным алгоритмом машинного обучения для предсказания космической погоды [8]. Качество результатов обучения алгоритмов определяется посредством значений абсолютной метрики RMSE (корень из среднего квадрата ошибок), относительной метрики SMAPE, и коэффициент детерминации  $R^2$ .

На обработанных данных было проведено обучение ранее упомянутых алгоритмов. На рисунке 1 показаны результаты обучения с помощью XGBoost. В Таблице 1 показаны значения метрик относительно применения каждого алгоритма к тестовой части набора.

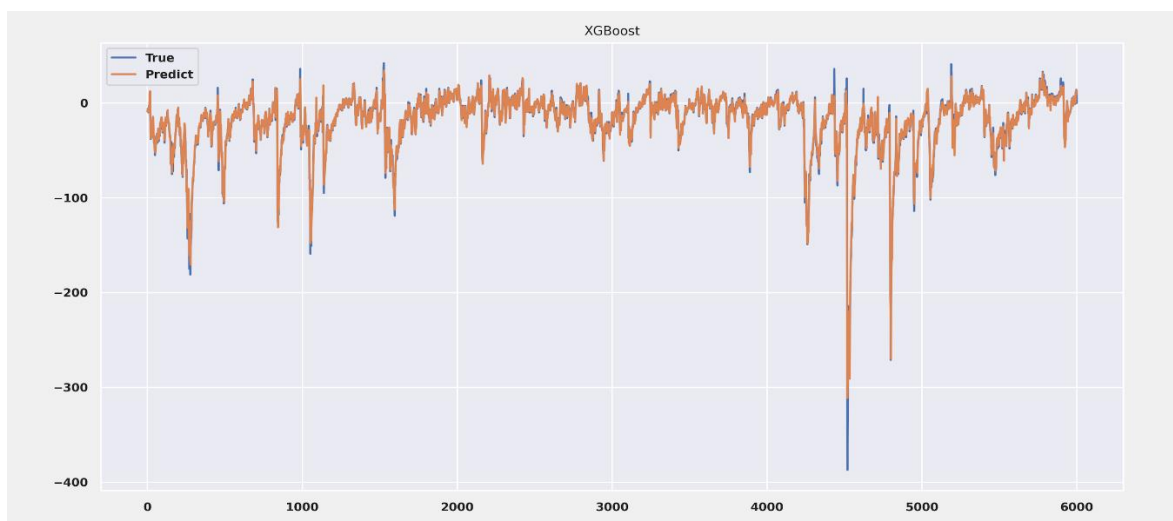


Рисунок 1 – Результаты обучения модели XGBoost.

Синие линии - истинное значение, оранжевые - предсказанные.

Таблица 1 – Значения метрики после обучения моделей

Название модели	Значение RMSE	Значение SMAPE	Значение R2
XGBoost	3.88	0.4248	0.96
Random Forest	4.03	0.4298	0.96
Decision Tree	5.43	0.5348	0.93
Gradient Boosting	3.89	0.4238	0.97

Таким образом, исходя из полученных значений метрики RMSE относительно каждого алгоритма, а также построенных графиков прогнозирования значений Dst, можно сделать вывод, что из представленных алгоритмов искусственного интеллекта с задачей прогнозирования параметра солнечного ветра Dst с учетом конкретной выборки наилучшим образом справляется алгоритм Gradient Boosting, имея наименьшее значение ошибок на тестовом наборе данных и наибольшее значение метрики R2 - примерно 0.97.

### Список литературы:

1. NASA Space Science Data Coordinated Archive: [сайт]. – URL: <https://nssdc.gsfc.nasa.gov/nmc/spacecraft/display.action?id=1997-045A> (дата обращения: 18.06.2023).
2. DSCOVR Space Weather Data Portal: [сайт]. – URL: <https://www.ngdc.noaa.gov/dscovr/portal/> (дата обращения: 20.06.2023).
3. Lang M., Owens M.J. A variational approach to data assimilation in the solar wind // Space Weather. 2019. Т. 17, №. 1. С. 59-83.
4. Xgboost: extreme gradient boosting / Т. Chen [et al.] // R package version 0.4-2. 2015. Т. 1, №. 4. С. 1-4.
5. Breiman L. Random forests // Machine learning. 2001. Т. 45. С. 5-32.
6. Breiman L. Classification and regression trees. Routledge, 2017.
7. Friedman J. H. Greedy function approximation: a gradient boosting machine // Annals of statistics, 2001. С. 1189-1232.
8. Islam M. M., Shareef H., Fayez Al Hassan E. S. Deep Learning Technique for Forecasting Solar Radiation and Wind Speed for Dynamic Microgrid Analysis // Przegląd Elektrotechniczny. 2023. Т. 99, №. 4.