

0.1. Костылева О.Д., Парамонов В.В. О моделировании экстремальных климатических явлений на основе методов машинного обучения

В работе рассматривается задача прогнозирования возникновения экстремальных климатических явлений. Данная задача представляет интерес, поскольку, даже в пределах одного государства резкие изменения климата является актуальной проблемой. В ряде стран поддерживается политика климата, которая регулярно корректируется, в зависимости от результатов новых исследований, климатической политики других стран и общественных настроений в целом [1]. Построение, предлагаемой прогнозной модели состоит из трех этапов:

- извлечение и организация данных;
- моделирование с применением различных методов машинного обучения;
- оценка качества полученных моделей.

Прогнозная модель, разрабатывалась и апробировалась для территории Новой Зеландии. В модели использованы два набора данных — для формирования признаков (включает измерения различных характеристик на выбранной территории), и классификации (используется для установления зависимостей). Источниками данных являются информационные ресурсы National Oceanic and Atmospheric Administration¹ и ClimateDataGuide².

Анализ данных, выявление зависимостей и прогнозирование экстремальных климатических явлений проводится с использованием методов машинного обучения с учителем. Для этого рассматривались метод случайного леса и метод опорных векторов. Был разработан модуль на языке Python, в котором ежедневные осадки суммируются по сезонам и разбиваются на классы «dry» («сухой»), «normal» («нормальный») и «wet» («влажный») в зависимости от значений. Ретроспективные данные о температуре поверхности воды, относительной влажности и количестве осадков, используемые для классификации, были получены с ресурса National Oceanic and Atmospheric Administration (NOAA) в бинарном формате NetCDF [2]. Для их извлечения также был разработан программный модуль.

При обучении модели использовался пакет SciKit Learn. Эксперименты показали, что точность прогноза на обучающей и контрольной выборках для метода случайного леса составляет 92–98% и 75%, а для метода опорных векторов — 95–98% и 58–66% соответственно. Для последнего метода была выполнена проверка работы модели на более простом разбиении на классы, а именно на 2 класса «dry» и «wet». Как и ожидалось, точность повысилась (70–75%), однако оценка прогноза с двумя

классами не столь эффективна, поскольку присутствует большая разница в значениях.

По результатам исследования созданы две модели [3], позволяющие с достаточно высокой точностью прогнозировать вероятность возникновения экстремальных климатических явлений. В итоге были предложены модели, основанные на двух методах машинного обучения. Наиболее качественным оказался метод случайного леса.

Работа выполнена в Институте математики и информационных технологий Иркутского государственного университета

Список литературы

- [1] Экстремальные погодные явления возглавили список глобальных рисков. Климатический центр Росгидромета: Главная геофизическая обсерватория имени А.И. Воейкова, 2013–2020. [Электронный ресурс]. URL: <http://cc.voeikovmgo.ru/ru/novosti/sobytiya/164-ekstremalnye-pogodnye-yavleniya-vozglavili-spisok-globalnykh-riskov> (дата обращения 01.02.2020).
- [2] What Is netCDF? 2020. [Электронный ресурс]. URL: <https://www.unidata.ucar.edu/software/netcdf/docs/faq.html> (дата обращения 01.04.2020).
- [3] NYADZI E., WERNERS E.S., BIESBROEK R. ET AL. Verification of Seasonal Climate Forecast toward Hydroclimatic Information Needs of Rice Farmers in Northern Ghana // Wea. Climate Soc. 2019. N. 11. P. 127–142.

¹<https://www.noaa.gov>

²<https://climatedataguide.ucar.edu>