

# МЕТОДЫ КЛАССИФИКАЦИИ ДЛЯ ПРОГНОЗИРОВАНИЯ ДОСТУПНОСТИ ЗАРЯДНЫХ СТАНЦИЙ

*Рассматривается применение методов классификации для прогнозирования состояния зарядной станции на следующий день. Полученные результаты могут быть использованы для оптимизации общественной инфраструктуры для зарядки.*

## ВВЕДЕНИЕ

В данной статье рассмотрены методы классификации Random Forest (RF) и Gradient Boosting (GB) для решения задачи прогнозирования. Эти модели предоставляют возможность использовать бинарные данные обучения (занято или свободно), классифицировать результат и определять меру уверенности предсказаний.

### I. ПОДГОТОВКА ДАННЫХ

Для данной статьи был использован открытый набор данных, содержащий информацию об использовании 30-ти зарядных станций, расположенных в городе Боулдер (США), за период с 2018-го по 2022-й годы [1]. Для решения задачи прогнозирования методами классификации данные были дискретизированы в регулярные временные ряды с 20-минутными интервалами, и был установлен признак занятости (1 – занято, 0 – свободно). Для повышения точности прогнозирования были сгенерированы лаговые и индикаторные признаки на основе исходных данных. Индикаторные: месяц, день недели, 4-х часовые интервалы суток, государственные праздники и нерабочие дни, школьные каникулы, а также дни, предшествующие выходным и праздничным дням. Лаговые: 10 прошлых состояний зарядки, состояние зарядки в прошлые сутки с окном в два временных интервала, и такое же состояние неделю назад.

### II. ОПИСАНИЕ АЛГОРИТМОВ

RF и GB были выбраны благодаря их высокой эффективности использования в задачах прогнозирования. RF – это ансамблевый метод машинного обучения с учителем, который объединяет несколько деревьев решений, обученных на разных подмножествах данных. В процессе обучения метода RF каждый базовый алгоритм создаётся независимо от других, в то время как GB применяет концепцию последовательного построения линейной комбинации алгоритмов. Каждый следующий базовый алго-

ритм в GB обучается так, чтобы уменьшить общую ошибку всех своих предшественников последовательно. Для работы с классификаторами RF и GB использовалась библиотека Python `sklearn.ensemble`.

### III. МЕТРИКИ И РЕЗУЛЬТАТЫ РАБОТЫ КЛАССИФИКАТОРОВ

Для качественной оценки результатов работы моделей использовались метрики Accuracy, Recall, Precision, AUC и F1. Результаты работы классификаторов приведены в таблице 1.

Таблица 1 – Результаты работы классификаторов

Классифик.	Метрики				
	Accur.	Recall	Precit.	AUC	F1
Random Forest	0.979	0.938	0.968	0.964	0.952
RF + Log. Regression.	0.983	0.953	0.968	0.972	0.961
Gradient Boosting	0.979	0.938	0.968	0.964	0.952

В процессе анализа результатов работы классификаторов были выявлены индикаторные и лаговые признаки, которые наиболее сильно влияют на результат прогноза для конкретного набора данных о зарядной станции.

### IV. ВЫВОДЫ

Методы классификации могут быть очень полезны для прогнозирования доступности зарядных станций. Однако, чтобы получить точный и надёжный прогноз, необходимо использовать достаточно большой объём данных. Классификаторы могут работать плохо, когда данных мало, например, станция только начала работать или когда есть явный тренд.

1. EV Charging Station Energy Consumption Open Data [Электронный ресурс]. – Режим доступа: [https://open\\_data.bouldercolorado.gov/datasets/](https://open_data.bouldercolorado.gov/datasets/). – Дата доступа: 01.04.2023

*Голубева Ирина Александровна*, магистрант кафедры информационных технологий автоматизированных систем БГУИР, [holubeva.irina@gmail.com](mailto:holubeva.irina@gmail.com).

*Научный руководитель: Севернёв Александр Михайлович*, доцент кафедры информационных технологий автоматизированных систем БГУИР, кандидат технических наук, доцент, [severnev@bsuir.by](mailto:severnev@bsuir.by)