

УДК 004.021

## ОПТИМИЗАЦИЯ МОДЕЛЕЙ ПРОГНОЗИРОВАНИЯ ПЛОЩАДЕЙ ЛЕСНЫХ ПОЖАРОВ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Безрукова Д.Е., студентка 5 курса факультета безопасности  
Научный руководитель – Катаева Е.С., старший преподаватель  
ФГБОУ ВО «ТУСУР»

*Ключевые слова:* машинное обучение, прогнозирование, классификация, регрессия, площади, лесные пожары.

*Работа посвящена повышению точности прогнозирования площадей лесных пожаров на основе метеорологических параметров с применение методов машинного обучения.*

**Введение.** В каждой стране имеются многочисленные лесные массивы, которые составляют значительную часть всех лесов мира. Значение леса для человека огромное, например, лес смягчает климат, является основным источником кислорода и древесины, поддерживает полноводность рек, препятствует распространению оврагов и т.д. Учитывая колоссальное значение лесов, в каждой стране принимаются все возможные меры для их сохранения и поддержания. На сегодняшний день одной из острой экологической проблемы являются лесные пожары, в связи с этим ряд ведомств уделяют большое внимание для борьбы с ними.

**Цель работы.** В связи с этим было принято решение выполнить прогнозирование с использованием методов машинного обучения (МО) для предотвращения будущих лесных пожаров из-за климатических особенностей в Португалии.

**Результаты исследования.** Анализируемый набор данных содержит 517 сведений о лесных пожарах в природном парке Монтесиньо [1] (северо-восточная часть Португалии) на основе метеорологических параметров по данным 2007 года.

Стоит отметить, что набор данных является актуальным и по сей день, так как климат Португалии неизменчивый, он определён особым

географическим расположением.

В ходе исследования по анализируемому набору данных с применение МО решались две задачи:

1. Задача классификации, то есть отнесение объектов к категории лесных пожаров по площади, охваченной огнем (га).

2. Задача регрессии, то есть по заданному набору признаков спрогнозировать значение целевой переменной – ареа.

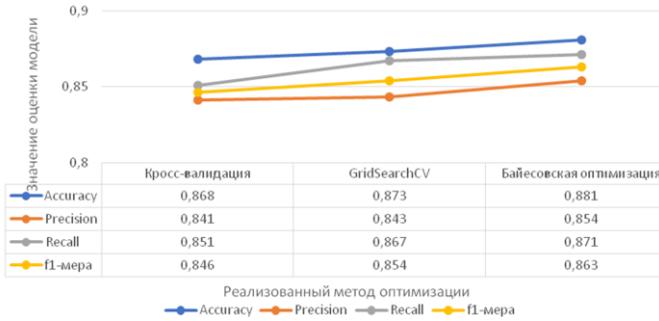
Реализация методов МО была выполнена на языке Python. В настоящей работе был реализован метод перекрестной проверки со значением  $N = 5$ . Значение  $N$  было установлено не случайно, так как оно обеспечивает низкое смещение входных данных и обеспечивает умеренную дисперсию [2].

Подбор оптимальных гиперпараметров в модели является одной из главных задач машинного обучения, так как от тщательного подбора колоссально зависит и оценка качества и надежность модели. В данной работе для оптимизации гиперпараметров моделей были использованы метод Поиск по сетке и Байесовская оптимизация.

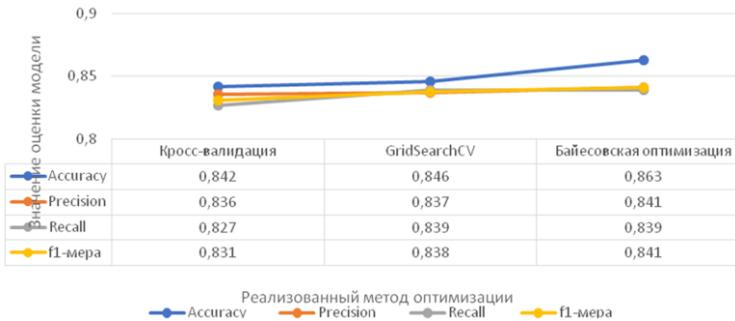
**Заключение.** В ходе исследования модель Случайный лес показала наилучшие результаты прогнозирования лесных пожаров по метеорологическим параметрам. Ансамблевые методы МО способствуют построению более эффективной и точной модели по сравнению с классическими методами МО.

На рисунках 1 – 2 представлены графики сравнения оценок качества моделей Случайный лес и Деревья лес для задачи классификации. На рисунке 3 представлен график оценок качества моделей Случайный лес и Деревья лес для задачи регрессии после применения методов оптимизации.

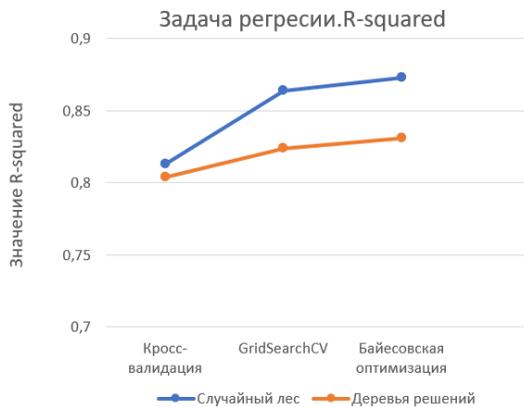
Таким образом, в рассматриваемых задачах наилучшим методом оптимизации гиперпараметров является Байесовская оптимизация. Преимуществом использования Байесовской оптимизации является то, что метод учитывает на каждой итерации предыдущие результаты оценок модели и с высокой вероятностью определяет наилучшую область пространства гиперпараметров для повышения производительности модели. Также по сравнению с Поиском по сетке Байесовская оптимизация требует меньших вычислительных затрат для оптимизации моделей.



**Рис. 1 – Задача классификация. Применение методов оптимизации для метода МО «Случайный лес»**



**Рис. 2 – Задача классификация. Применение методов оптимизации для метода МО «Деревья решений»**



**Рис. 3 – Задача регрессии-R<sup>2</sup>. Применение методов оптимизации**

**Библиографический список:**

1. Forest Fires Data Set. [Электронный ресурс]. – URL: <https://www.kaggle.com/datasets/elikplim/forest-fires-data-set> (дата обращения: 05.12.2022). – Текст: электронный.
2. Смещение против дисперсии в машинном обучении. [Электронный ресурс]. – URL: <https://learn.microsoft.com/ru-ru/dotnet/machine-learning/how-to-guides/train-machine-learning-model-cross-validation-ml-net> (дата обращения: 06.12.2022). – Текст: электронный.

**OPTIMIZATION OF FOREST FIRE AREA PREDICTION MODELS  
USING MACHINE LEARNING METHODS**

**Bezrukova D.E**

***Keywords:** machine learning, forecasting, classification, regression, areas, forest fires.*

*The work is devoted to improving the accuracy of forecasting forest fire areas based on meteorological parameters using machine learning methods.*