

ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В АНАЛИЗЕ КЛИМАТИЧЕСКИХ ДАННЫХ

П. С. Гоч¹⁾, М. В. Королюк²⁾

¹⁾Белорусский национальный технический университет, Беларусь, Минск,
polinagoc12@gmail.com

²⁾Белорусский национальный технический университет, Беларусь, Минск,
korolukmx2@yandex.by

Глобальное потепление – одна из самых актуальных проблем современности, угрожающая экосистемам и будущему планеты. В работе рассмотрены методы анализа климатических данных с использованием машинного обучения. На основе исторических данных о температуре построена линейная регрессионная модель для выявления трендов и прогнозирования. Проект демонстрирует, как технологии ML помогают в понимании и решении глобальных экологических проблем.

Ключевые слова: глобальное потепление; машинное обучение; линейная регрессия; прогнозирование температуры.

USING MACHINE LEARNING METHODS IN CLIMATE DATA ANALYSIS

P. S. Goch¹⁾, M. V. Korolyuk²⁾

¹⁾Belarusian National Technical University, Belarus, Minsk, Minsk,
polinagoc12@gmail.com

²⁾Belarusian National Technical University, Belarus, Minsk,
korolukmx2@yandex.by

Global warming is one of the most pressing problems of our time, threatening ecosystems and the future of the planet. The paper considers methods of analysing climate data using machine learning. Based on historical temperature data, a linear regression model is built to identify trends and make predictions. The project demonstrates how ML technologies help in understanding and solving global environmental problems.

Keywords: global warming; machine learning; linear regression; temperature prediction.

Введение

Глобальное потепление является одной из наиболее значимых экологических проблем XXI века, оказывающей влияние на экосистемы, экономику и качество жизни людей. Рост средней температуры Земли, вызванный увеличением концентрации парниковых газов, приводит к таянию

ледников, повышению уровня моря и учащению экстремальных погодных явлений. Актуальность темы исследования обусловлена необходимостью прогнозирования климатических изменений и разработки стратегий для смягчения их последствий. В научной литературе активно обсуждаются методы анализа климатических данных, включая применение машинного обучения для выявления трендов и построения прогнозов [1]

Методология исследования

Цель исследования – анализ исторических климатических данных и прогнозирование изменений температуры с использованием методов машинного обучения. Гипотеза исследования заключается в том, что машинное обучение позволяет эффективно выявлять тренды изменения температуры и строить точные прогнозы. В качестве методов исследования использовались:

- 1) Линейная регрессия для анализа зависимости температуры от времени.
- 2) Метрики качества модели (среднеквадратичная ошибка).
- 3) Визуализация данных с использованием библиотек matplotlib и seaborn.

Линейная регрессия – это статистический метод, используемый для моделирования зависимости между одной зависимой переменной (y) и одной или несколькими независимыми переменными (x) [2]. В простейшем случае (линейная регрессия с одной переменной) она описывает прямую линию, которая наилучшим образом аппроксимирует данные. Уравнение линейной регрессии имеет вид:

$$y = mx + b$$

Формула для коэффициента наклона (m):

$$m = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

Формула для свободного члена (b)

$$b = \frac{\sum y}{n} - m \frac{\sum x}{n}$$

Рассмотрим климатические изменения за последние 100 лет на примере города Минска. Взяв данные[3] и применив формулы для нахождения коэффициента наклона и свободного члена, получаем уравнение вида:

$$y = 0.0234x - 40.1234$$

Результаты, их обсуждение, заключение

В результате исследования была построена линейная регрессионная модель, которая показала устойчивый рост средней температуры Земли за последние десятилетия. Коэффициент наклона регрессионной прямой составил $m = 0.0234$, что свидетельствует о значительном увеличении температуры. Визуализация данных подтвердила наличие положительного тренда (рисунок).

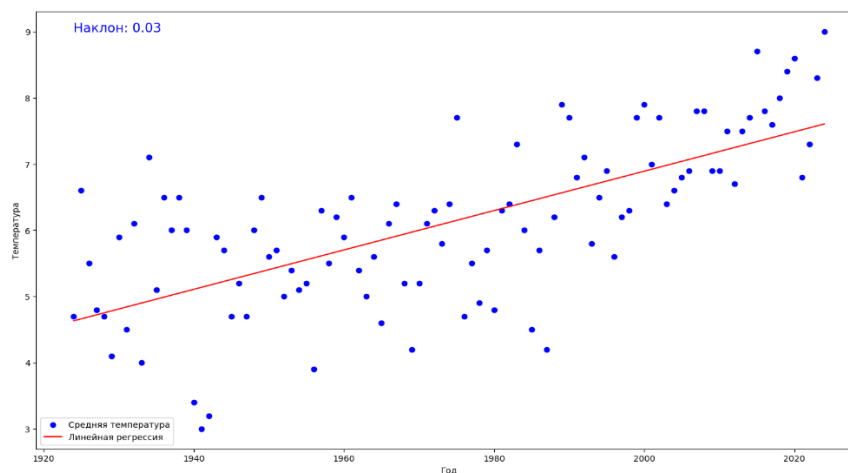


График изменения температуры по годам

Полученные результаты могут быть использованы для разработки мер по снижению выбросов парниковых газов и адаптации к изменению климата. В будущих исследованиях планируется расширить анализ, включив дополнительные факторы, такие как уровень CO_2 и ледниковый покров, а также использовать более сложные модели машинного обучения для повышения точности прогнозов.

Библиографические ссылки

1. Современные подходы к защите данных в распределённых системах [Электронный ресурс]. Режим доступа: <https://habr.com/ru/articles/804135>. Дата доступа: 10.03.2025.
2. Нилд Т. Математика для Data Science. Управляем данными с помощью линейной алгебры, теории вероятностей и статистики / М.: Издательский дом «Питер», 2021. 164 с
3. Архив погодных условий [Электронный ресурс]. Режим доступа: <http://www.pogodaiklimat.ru/history/26850.htm>. Дата доступа: 11.03.2025.