

Анализ растительности на спутниковых снимках с использованием машинного обучения

Е.В. Журавская

Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия.

Обоснование. Анализ растительности представляет собой важную задачу в экологии, сельском хозяйстве, климатических исследованиях и других сферах. Мультиспектральные снимки, полученные со спутников, позволяют нам извлечь информацию о земной поверхности из разных спектральных каналов. Для анализа таких данных можно использовать машинное обучение.

Цель — реализовать прогнозирование вегетационного индекса NDVI на основе многолетних спутниковых данных и моделей машинного обучения, а также оценить точность моделей.

Методы. Для дальнейшего анализа будут использованы снимки с веб-платформы Earth Explorer [1], где можно найти данные из коллекции Landsat Collection 2 Level-2 — снимки, полученные с помощью спутника Landsat 8. Вручную были отобраны снимки с 2015 по 2024 год в летний период, так как именно в этот временной промежуток растительность наиболее хорошо видна. Также были выбраны снимки с облачностью менее 10 %. Для их отбора использовался фильтр «Cloud Cover» на платформе Earth Explorer. Далее с помощью кода на языке программирования Python был вычислен вегетационный индекс NDVI. Данный индекс широко используется в анализе растительности [2, 3]. Он был рассчитан для каждого пикселя на снимке, которые далее были сохранены в формате .tif. Для минимизации шумов и уменьшения размерности данные были агрегированы в области 3×3 пикселя. Для каждой области был построен временной ряд NDVI. Прогнозирование NDVI на 2024 год производилось с использованием трех моделей: линейной регрессии, случайного леса (Random Forest) и LSTM.

Результаты. Была проведена оценка точности моделей с использованием метрик MSE, MAE и R^2 (табл. 1). Время обучения моделей измерялось при выполнении кода на ноутбуке с процессором Intel Core i3-1215U (12-е поколение, 8 логических потоков) и 8 ГБ оперативной памяти. Вычисления проводились без использования GPU. Все эксперименты были реализованы в среде Python с использованием библиотек scikit-learn (линейная регрессия и случайный лес) и TensorFlow/Keras (модель LSTM). Код запускался в редакторе Visual Studio Code. Наилучшие результаты показала модель линейной регрессии: $R^2 = 0,91184$ при времени обучения 0,21 с. Модель LSTM продемонстрировала перспективные результаты ($R^2 = 0,86091$), но потребовала значительно больше времени (704 с). Случайный лес показал наихудшую точность ($R^2 = 0,85603$) при среднем времени обучения 510 с.

Таблица 1. Оценка качества моделей

Модель	MSE	MAE	R^2	Время обучения, с
Линейная регрессия	0,00130	0,02553	0,91184	0,21
LSTM	0,00205	0,03249	0,86091	704,15
Случайный лес	0,00212	0,03065	0,85603	510,56

Выводы. Полученные результаты демонстрируют, что машинное обучение может быть эффективно в решении задачи прогнозирования NDVI по спутниковым снимкам. Линейная регрессия обеспечила лучшую точность. Модель LSTM требует дальнейшей настройки и, возможно, увеличения обучающей выборки. Подход может быть использован для экологического мониторинга, оценки деградации территорий и управления зелеными насаждениями.

Ключевые слова: вегетационный индекс; машинное обучение; дистанционное зондирование; прогнозирование; растительность.

Список литературы

1. usgs.gov [Электронный ресурс] EarthExplorer — U.S. Geological Survey. Режим доступа: <https://earthexplorer.usgs.gov> Дата обращения: 30.04.2025.
2. Блэкберн А.А., Золотой А.Л., Остапко В.М. Сравнительная оценка квазиприродных территорий Донецкого края с использованием индекса NDVI // Материалы IX Международного симпозиума: Степи Северной Евразии. Оренбург: Оренбургский государственный университет. 2021. № 9. С. 130–136. doi: 10.24412/cl-36359-2021-130-136 EDN: JGRRGP
3. Boori M.S., Choudhary K., Kupriyanov A.V. Crop growth monitoring through sentinel and landsat data based NDVI time-series // Computer Optics. 2020. Vol. 44, N 3. P. 409–419. doi: 10.18287/2412-6179-CO-635 EDN: WFXXE

Сведения об авторе:

Елизавета Владимировна Журавская — студентка, группа 6404-010302D, институт информатики и кибернетики; Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия. E-mail: eliza.zhuravskaya@gmail.com

Сведения о научном руководителе:

Рустам Александрович Парингер — кандидат технических наук, доцент кафедры технической кибернетики; Самарский национальный исследовательский университет имени академика С.П. Королева, Самара, Россия. E-mail: RusParinger@ssau.ru