

Использование методов машинного обучения для геопространственного моделирования типов болот на территории природного парка «Нумто»

Научный руководитель – Юртаев Андрей Александрович

Московченко Михаил Дмитриевич

Аспирант

Тюменский государственный университет, Институт экологической и сельскохозяйственной биологии (Х-БИО), Тюмень, Россия

E-mail: simonreise@rambler.ru

В Западной Сибири болота занимают около 104 млн га, оказывая значительное влияние на глобальный баланс углерода и на климат Земли. Так как настоящее время на севере Западной Сибири практически повсеместно рост температуры воздуха превышает среднемировые показатели [3], это приводит к деградации мерзлых болот. В зоне прерывистого распространения многолетней мерзлоты в Западной Сибири бугристые болота чаще всего являются мерзлыми, тогда как остальные типы болот чаще являются немерзлыми. Деструктивное влияние климатических изменений на мерзлые болота делает актуальной разработку моделей, позволяющих проводить крупномасштабное картографирование типов болот на больших территориях для оценки площади болот, которые могут быть подвержены деградации.

Целью данной работы является разработка оптимальной модели машинного обучения для картографирования типов болот на территории природного парка «Нумто».

Для создания модели типов болот в качестве предикторов были использованы следующие источники: спутниковые снимки Landsat 8 за 4 времени года и ЦММ ASTER GDEM в качестве предикторов и карта типов болот природного парка «Нумто» в качестве целевой переменной. Данные были предобработаны и разбиты на тренировочный набор, на котором осуществлялось обучение моделей, и валидационный и тестовый наборы, на которых осуществлялся контроль качества моделирования. Для достижения цели были обучены и сравнены между собой 15 различных моделей машинного обучения, в том числе 4 модели классического машинного обучения и 11 моделей глубокого обучения (нейронных сетей).

Среди всех моделей классического машинного обучения наилучших результатов (точность 79,95% на тренировочном наборе и 75,1 и 75,73% на валидационном и тестовом наборах), сравнимых с результатами нейросетей, достигла модель градиентного бустинга [1], остальные модели показали худшую точность. Среди нейросетей наилучший результат показала DeepLabV3 [2] с экстрактором признаков ResNet50 (точность 80,14% на тренировочном наборе и 75,28 и 75,76% на валидационном и тестовом наборах). Также было выявлено, что большее количество параметров у нейросети не всегда приводит к улучшению результатов моделирования, а сверточные нейронные сети имеют преимущество над нейросетями с механизмом внимания.

При дальнейшем сравнении этих двух моделей было выявлено, что с моделированием бугристых болот лучше справилась нейронная сеть, тогда как при моделировании других типов болот – олиготрофных немерзлых, мезотрофных и евтрофных, обе модели показали примерно одинаковую точность. Построение карт по результатам моделирования показало, что применение методов классического машинного обучения, не учитывающих пространственную структуру данных и проводящих попиксельную классификацию, приводит к тому, что границы ареалов размываются, часто представляют собой градиенты, появляются много небольших (несколько пикселей) участков какого-либо класса, часто внутри ареалов одного класса есть небольшие включения других классов. Применение

для моделирования нейронных сетей позволяет сохранить исходную структуру и степень генерализации данных.

Так как точность на тренировочных, валидационных и тестовых данных отличалась не сильно, можно сделать вывод, что модели практически не переобучились, следовательно, могут быть использованы для моделирования типов болот и на смежных территориях. Модели были успешно применены для картографирования типов болот на нескольких смежных с исследуемой территорией участках.

Источники и литература

- 1) Chen T., Guestrin C. Xgboost: A scalable tree boosting system // Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. – 2016. С. 785-794.
- 2) Chen L. C. et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs //IEEE transactions on pattern analysis and machine intelligence. – 2017. – Т. 40. – №. 4. – С. 834-848.
- 3) Kirpotin S.N. Western Siberia in a changing climate. //International Journal of Environmental Studies. 2014. – Т.71. – №. 5. – С. 591-594.

Иллюстрации

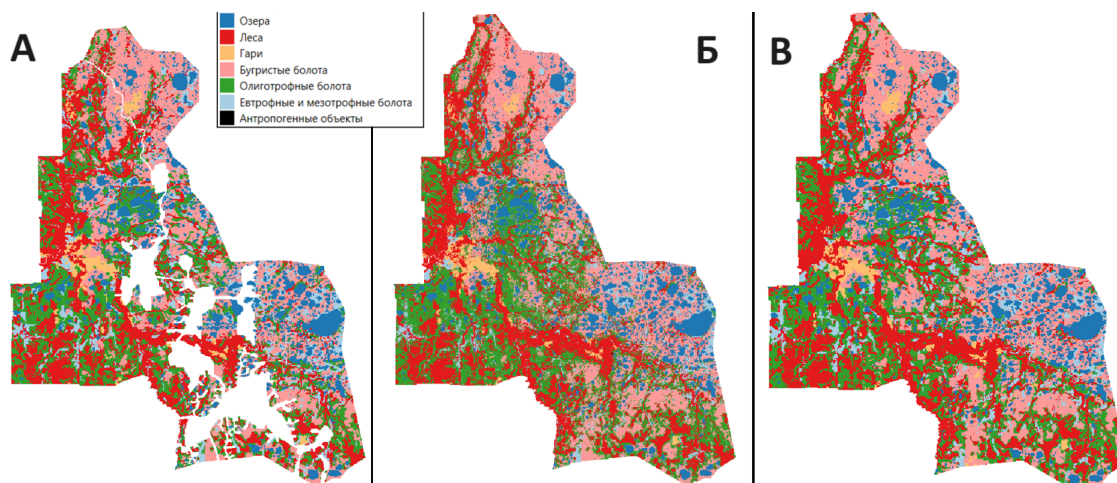


Рис. : Результат картографирования: А - исходных данных, Б - результатов моделирования градиентного бустинга, В - результатов моделирования нейросети DeepLabV3