

ВОЗМОЖНОСТИ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ РАЗВИТИЯ ГОРОДСКИХ ТЕРРИТОРИЙ

Андрей Александрович Сивков

Сибирский государственный университет геосистем и технологий, 630108, Россия, г. Новосибирск, ул. Плахотного, 10, магистрант кафедры прикладной информатики и информационных систем, тел. (996)379-75-52, e-mail: Andrey-sa-396@mail.ru

Алексей Александрович Колесников

Сибирский государственный университет геосистем и технологий, 630108, Россия, г. Новосибирск, ул. Плахотного, 10, кандидат технических наук, доцент кафедры картографии и геоинформатики, тел. (913)725-09-28, e-mail: alexeykw@mail.ru

Проектирование городов всегда было пространственным процессом. Так как город представляет собой совокупность пространств и связей как надземных, так и подземных, то особенно важно вывести территориальное планирование на уровень пространственного моделирования. В работе были исследованы возможности методов машинного обучения для прогнозирования развития городских территорий, составлена математическая модель прогнозирования, произведена оценка её точности.

Ключевые слова: машинное обучение, дерево решений, сегментация, нейронные сети, классификация, геоинформатика, спутниковые снимки, развитие городов.

POSSIBILITIES OF MACHINE LEARNING METHODS FOR FORECASTING THE DEVELOPMENT OF URBAN TERRITORIES

Andrey A. Sivkov

Siberian State University of Geosystems and Technologies, 10, Plakhotnogo St., 630108, Novosibirsk, Russia, Graduate, Department of Applied Informatics and Information Systems, phone: (383)361-06-35, e-mail: Andrey-sa-396@mail.ru

Alexey A. Kolesnikov

Siberian State University of Geosystems and Technologies, 10, Plakhotnogo St., 630108, Novosibirsk, Russia, Ph. D., Associate Professor, Department of Cartography and Geoinformatics, phone: (383)361-06-35, e-mail: alexeykw@mail.ru

Urban design has always been a spatial process. Since the city is a combination of spaces and connections, both above and under ground, it is especially important to bring territorial planning to the level of spatial modeling. In the paper, the possibilities of machine learning methods for predicting the development of urban areas were investigated, a forecasting model was compiled, and its accuracy was evaluated.

Key words: machine learning, decision tree, segmentation, neural networks, classification, geoinformatics, satellite imagery, urban development.

Введение

В настоящее время машинное обучение является одним из наиболее популярных и развивающихся направлений среди научных исследований. В нём можно выделить такое направление, как обработка пространственных данных,

к которому можно отнести и прогнозирование развития городских территорий. Система прогнозирования городской застройки позволяет отображать сравнительную перспективность отдельных зон и районов исследуемой местности с точки зрения возможности дальнейшего планирования и расширения площадей жилищных территорий [1].

Методы и материалы

Для решения задачи по прогнозированию зон городской застройки были выбраны методы и инструменты машинного обучения. Такой выбор был сделан, поскольку математическое обоснование и интерпретация получаемых математических моделей, применяемых методов прогнозирования, а также точность получаемых прогнозов стали вполне реальными при использовании для расчетов вычислительной техники существующего уровня [2]. Используемая при исследовании методика прогнозирования основывается на анализе пространственных и временных закономерностей в виде заранее формируемой обучающей выборки. Она представляет собой регулярную сетку с набором семантических характеристик максимально подробно описывающих каждую из ячеек.

Создание модели прогнозирования проходит в несколько этапов. Они представлены в виде технологической схемы, которая приведена на рис. 1.

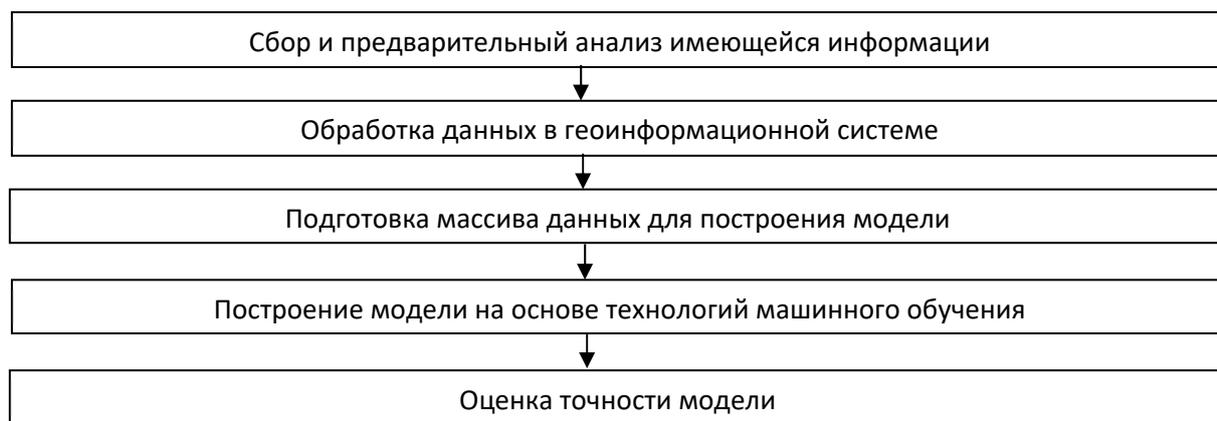


Рис. 1. Технологическая схема создания модели прогнозирования

На первом этапе из доступных источников были взяты спутниковые снимки на территорию Новосибирска, архивные карты и векторные слои из интернет-ресурса Open Street Map [3].

На втором этапе нужно было сформировать цифровую модель рельефа на исследуемую территорию, построить регулярную сетку точек и заполнить их данными, описывающими положение каждой точки относительно других объектов (удалённость от рек, дорог, населённых пунктов, лесов и сложных форм рельефа, высоту, уклон, экспозицию). Также с помощью архивных карт нужно было проанализировать приблизительное развитие городской территории до настоящего времени и создать полигональные объекты, описывающие границы городской территории в разные года.

На третьем этапе нужно было сформулировать способ классификации сформированной обучающей выборки и, на основе этого, выбрать метрику оценки полученного результата прогноза. С помощью пространственных запросов были созданы комбинации точек, в семантику которых, в зависимости от их положения относительно полигонов застройки, добавлен дополнительный атрибут, отражающий класс урбанизированности (0 – уже застроенная территория, 1 – застроенная на текущем этапе анализа, 2 – оставшаяся незастроенной). Далее размеченные записи нужно было выгрузить как файл формата csv для загрузки в R-Studio [4]. В качестве метрики качества было выбрано количество совпадений прогноза, относительно всей тестовой выборки [5].

На четвёртом этапе в программном обеспечении R Studio нужно было произвести построение дерева решений с указанием того, что решается задача классификации. Параметры построения дерева были оставлены по умолчанию, графическое представление полученной математической модели показано на рис. 2 [6].

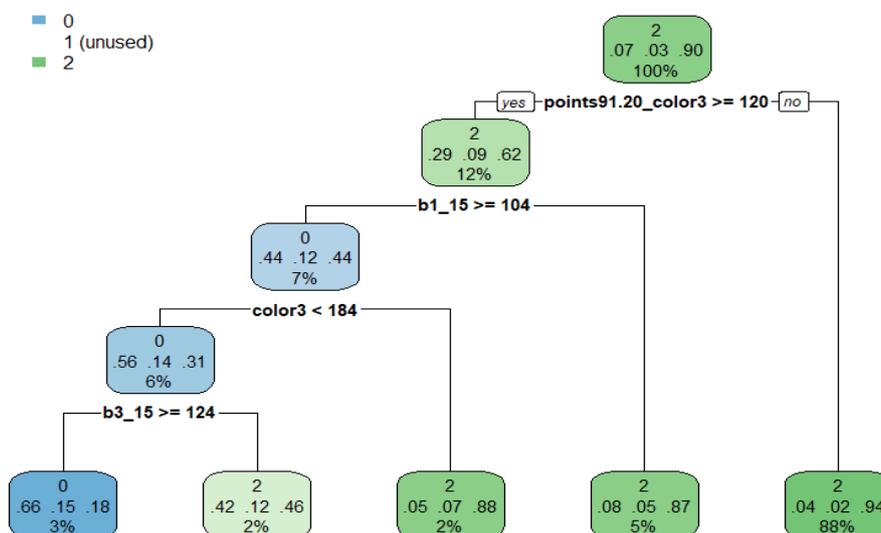


Рис. 2. Дерево решений

Результаты

На заключительном этапе нужно было произвести оценку точности модели. Для этого был использован код (`mean(labeled_points_renamed$class == t_pred)`). В результате чего значение ошибки получилось 0.9156262, значит точность созданной модели составляет более 90 % [7].

Обсуждение

В работе были показаны возможности методов машинного обучения для прогнозирования развития городских территорий. Использование этих методов возможно в любых сферах деятельности, где есть необходимость предсказания

какого-либо явления. Это можно использовать как для коммерческих, так и для бытовых видов деятельности, для управленческих решений на муниципальном и региональном уровнях [8].

Заключение

С использованием методов машинного обучения была создана математическая модель, которая отображает прогнозирование развития городской территории на примере города Новосибирск. Данная модель позволяет формализовать исследовательскую задачу, содержание и процедуру исследования, а также может использоваться для пространственного анализа систем различной сложности и прогнозирования изменений в пространстве и времени [9].

На данный момент такая система неидеальна и может приводить к значительным погрешностям, поскольку исходные данные для каждой местности уникальны. Поэтому следующими шагами по улучшению модели могут быть расширение обучающей выборки на основе городов с другими географическими и климатическими характеристиками и расширение списка параметров, описывающих каждый элемент регулярной сетки [10].

Таким образом, содержание моделей прогнозирования не ограничивается предсказанием будущего, с их помощью можно показать и существующее в настоящее время состояние объекта или явления, например, для поиска еще неизвестных или малоизученных взаимосвязей компонентов, ретроспективного анализа, оценки влияния внешних факторов.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Гуняев А.Е. Обзор методов машинного обучения для прогнозирования наступления политических событий [Текст] / А.Е. Гуняев. – М. МГУ, 2018. – 23 с.
2. Гришинцев, А.Ю. Развитие и разработка методов и средств обеспечения систем автоматизированного проектирования распределённых геоинформационных систем [Текст]: / СПб: НИУИТМиО, 2016. – выпуск 279 с.
3. Использование машинного обучения для построения картографических изображений [Текст] / А. А. Колесников, П. М. Кикин, Д. В. Грищенко, Е. В. Комиссарова. – Новосибирск : СГУГиТ, 2016. – 15 с.
4. Колесников А. А., Кикин П. М., Комиссарова Е. В. Использование открытых семантических баз данных для получения пространственной информации // Интерэкспо ГЕО-Сибирь-2017. XIII Междунар. науч. конгр. : Междунар. науч. конф. «Геодезия, геоинформатика, картография, маркшейдерия» : сб. материалов в 2 т. (Новосибирск, 17–21 апреля 2017 г.). – Новосибирск : СГУГиТ, 2017. Т. 2. – С. 53–57.
5. Мошкин, В.С. Особенности интеграции механизмов логического вывода в онтологическую модель представления знаний с помощью swrl-правил [Текст]: / В.С. Мошкин. – Ульяновск: УГТУ, 2014. – выпуск 173 – 181 с.
6. Станкевич Т.Ю. Разработка метода оперативного прогнозирования динамики развития лесного пожара посредством искусственного интеллекта и глубокого машинного обучения [Текст] / Т.Ю. Станкевич. – Калининград: КГТУ, 2018. –120 с.
7. Тимина, И.Р. Корректировка гипотез прогноза для извлечения знаний о временных рядах [Текст] / И.Р. Тимина. – Ульяновск: УГТУ, 2014. – выпуск – 68 – 76 с.
8. Черкасов Д.Ю. Машинное обучение / [Текст] Д.Ю. Черкасов, В.В. Иванов. – М. МГУ, 2018. – 85с.

9. Wei Wei, «Research on the Application of Geographic Information System in Tourism Management», *Procedia Environmental Sciences*, vol. 12pp. 1104 – 1109, 2012.
10. B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik, “Simultaneous detection and segmentation,” in *European Conference on Computer Vision*. // Springer, 2014, pp. 297–312.

© *A. A. Сивков, А. А. Колесников 2020*