

ОБРАБОТКА ГРАФИЧЕСКОЙ И ЛЕКСИЧЕСКОЙ ИНФОРМАЦИИ ДЛЯ ФОРМИРОВАНИЯ БАЗ ЗНАНИЙ О ЛАНДШАФТЕ

Н.В. Гомзов

ФГБОУ ВО «Тамбовский государственный технический университет»

Аннотация. В статье рассматривается проблема обработки графической и лексической информации для формирования баз знаний о ландшафте. Рассматривается возможность использования нейро-нечеткой сети, которая обеспечивает эффективную обработку гетерогенных данных и позволяет получать более точные результаты анализа.

Ключевые слова: ландшафт, обработка данных, нейронная сеть, нечеткая логика, анализ.

PROCESSING OF GRAPHIC AND LEXICAL INFORMATION FOR THE FORMATION OF KNOWLEDGE BASES ON LANDSCAPES

N.V. Gomzov

Tambov State Technical University

Abstract. The article addresses the problem of processing graphic and lexical information for the formation of knowledge bases on landscapes. It considers the possibility of using a neuro-fuzzy network, which enables effective processing of heterogeneous data and obtaining more accurate analysis results.

Keywords: landscape, data processing, neural network, fuzzy logic, analysis.

В современном мире, где технологии играют важную роль в различных сферах, сбор и обработка данных стали ключевыми аспектами успешного функционирования многих систем. Это особенно актуально для таких областей, как экология, геология, сельское хозяйство и других, где информация о ландшафте имеет решающее значение.

Важно исследовать различные методы и системы, которые в настоящее время используются для обработки данных о ландшафте. Это сравнение не только создаст контекст для предложенного подхода на основе нейро-нечеткой сети, но и подчеркнет его уникальные преимущества и возможные области применения.

Одним из наиболее распространенных методов является машинное обучение, которое охватывает широкий спектр алгоритмов, используемых для задач классификации и прогнозирования. Традиционные методы машинного обучения, такие как деревья решений, машины опорных векторов и случайные

леса, часто применяются в анализе ландшафтов. Эти алгоритмы ценятся за свою интерпретируемость и эффективность при работе с структурированными данными. Однако они также сталкиваются с такими проблемами, как необходимость в значительной инженерии признаков и чувствительность к качеству данных. Хотя эти методы могут давать полезные результаты, они не всегда способны полностью учесть сложность взаимосвязей, присущих данным о ландшафте.

Другая важная категория — это глубокое обучение, особенно с использованием сверточных нейронных сетей (CNN). Эти сети приобрели популярность благодаря своей способности обрабатывать и анализировать данные изображений, что делает их особенно подходящими для задач, связанных со спутниковыми изображениями или аэрофотосъемкой. Модели глубокого обучения могут автоматически извлекать признаки из необработанных данных, снижая необходимость в ручной предварительной обработке. Однако у них есть свои сложности, такие как требование больших наборов данных и значительные вычислительные ресурсы. Сложность этих моделей иногда может приводить к переобучению, особенно когда объем данных для обучения ограничен.

Традиционные методы геоинформационных систем (ГИС) также играют важную роль в анализе ландшафтов. ГИС уже много лет является основным инструментом пространственного анализа, предлагая возможности для картирования, анализа и визуализации географических данных. Эти методы основаны на заранее определенных моделях и алгоритмах для оценки пространственных отношений и закономерностей. Хотя инструменты ГИС очень эффективны для ряда применений, им может не хватать адаптивности и гибкости, которые предлагают более современные методы машинного и глубокого обучения. Это ограничение может затруднять их использование для решения сложных и динамичных задач, связанных с изменениями ландшафта.

Кроме того, наблюдается растущая тенденция к использованию гибридных подходов, которые объединяют различные методы для повышения общей эффективности обработки данных о ландшафте. Например, интеграция методов машинного обучения с ГИС позволяет использовать сильные стороны обеих систем, обеспечивая более всесторонний анализ. Ансамблевые методы, объединяющие несколько алгоритмов обучения, также могут улучшить точность предсказаний, смягчая недостатки отдельных моделей.

Анализ существующих технологий, показывает уникальные преимущества предложенной системы на основе нейро-нечеткой сети. Этот подход не только улучшает способность обрабатывать гетерогенные данные, но и предлагает большую гибкость и точность в анализе характеристик ландшафта. В конечном итоге понимание сильных и слабых сторон этих методов позволит развивать будущие применения в анализе ландшафтов, прокладывая путь для более эффективных стратегий управления природными ресурсами и исследований окружающей среды.

Сбор и обработка данных о ландшафте могут способствовать созданию баз знаний, которые, в свою очередь, формируют основу для реализации систем поддержки принятия решений. Такие системы могут применяться в различных областях, таких как оптимизация использования земельных ресурсов, прогнозирование природных катастроф, планирование транспортных маршрутов и многое другое.

Для обработки графической информации могут использоваться различные методы, такие как геоинформационные системы (ГИС), дистанционное зондирование и другие. ГИС позволяют анализировать пространственные данные и выявлять закономерности, которые могут быть полезны для принятия решений.

Лексическую информацию можно обрабатывать с помощью методов текстового анализа, которые позволяют извлекать характеристики ландшафта из текстовых данных. Например, ключевые термины, связанные с определенными типами почв или растительности, могут быть идентифицированы и использованы для построения базы знаний.

Несмотря на разнообразие методов обработки информации, данные часто оказываются неполными или неточными. В таких случаях могут быть использованы системы на основе нечеткой логики для работы с неопределенными данными. Они могут быть интегрированы с методами нейронных сетей, что позволит системе игнорировать пробелы в данных или неточности.

Предлагаемая нейро-нечеткая сеть сочетает в себе сильные стороны как нейронных сетей, так и систем на основе нечеткой логики, создавая надежную платформу для работы с гетерогенными типами данных. Архитектура сети включает в себя входной слой, несколько скрытых слоев и выходной слой, каждый из которых выполняет определенную функцию в процессе обработки информации.

Входной слой предназначен для приема различных типов данных, включая графическую информацию, такую как спутниковые изображения, аэрофотоснимки и топографические карты, а также лексическую информацию, полученную из текстовых описаний, предоставленных экспертами. Кроме того, интегрируются метеорологические данные, что добавляет контекст для анализа ландшафта. Такой многослойный ввод позволяет системе учитывать широкий спектр факторов, влияющих на характеристики ландшафта.

Скрытые слои сети выполняют основную обработку данных. Эти слои используют активационные функции, такие как Rectified Linear Unit (ReLU) или сигмоидальные функции, для моделирования сложных взаимосвязей между входными переменными. За счет включения нескольких скрытых слоев сеть способна захватывать сложные взаимодействия внутри данных. Нейро-нечеткий подход позволяет интегрировать нечеткую логику, что помогает системе управлять неопределенностью и двусмысленностью, присутствующей в данных. Системы нечеткого вывода могут преобразовывать четкие входные

данные в нечеткие множества, которые затем обрабатываются через сеть для получения более точных результатов.

Выходной слой генерирует финальные результаты, которые могут включать классификации ландшафтов по кластерам или правила для принятия решений. В зависимости от применения эти результаты могут быть адаптированы для предоставления конкретных выводов, таких как прогнозирование изменений ландшафта, определение оптимальных мест для строительства или информирование стратегий охраны природы. Использование функций, таких как `softmax`, в выходном слое позволяет проводить многоцелевую классификацию, что способствует комплексной категоризации характеристик ландшафта.

Для обеспечения эффективности нейро-нечеткой сети необходимо реализовать надежный процесс обучения. Этот процесс обычно включает использование алгоритмов обратного распространения ошибок для корректировки весов и смещений внутри сети на основе ошибок, рассчитанных на выходе. Различные методы оптимизации, такие как стохастический градиентный спуск или алгоритм Adam, могут повысить скорость сходимости и точность в процессе обучения.

С экономической точки зрения, интеграция передовых технологий обработки данных, таких как нейро-нечеткая сеть, может привести к повышению эффективности управления ресурсами. Например, в сельском хозяйстве система может оптимизировать распределение земли, воды и удобрений на основе точного анализа данных. Максимизируя урожайность и минимизируя потери, фермеры могут значительно увеличить свою прибыль. Способность системы предоставлять рекомендации по оптимальным методам ведения хозяйства также может снизить затраты, связанные с применением ресурсов и рабочей силой, что способствует более устойчивой экономике сельского хозяйства.

Кроме того, в сфере градостроительства нейро-нечеткая сеть может поддерживать более обоснованный процесс принятия решений. Анализируя такие факторы, как демографические данные, оценки воздействия на окружающую среду и потребности в инфраструктуре, планировщики могут проектировать города, способствующие устойчивому развитию и улучшению качества жизни жителей. Эффективное развитие городов может привести к увеличению стоимости недвижимости, созданию новых деловых возможностей и улучшению общественных услуг, что в конечном итоге способствует экономическому процветанию.

Система может способствовать активным мерам реагирования на экологические вызовы, такие как изменение климата или природные катастрофы. Предоставляя своевременные и точные прогнозы, сообщества могут внедрять эффективные стратегии смягчения последствий и адаптации. Эта возможность не только защищает жизни и имущество, но и укрепляет доверие общества к институтам и органам управления.

Базовая архитектура нейро-нечеткой сети для обработки графической и лексической информации не имеет принципиальных отличий показана на рис. 1.

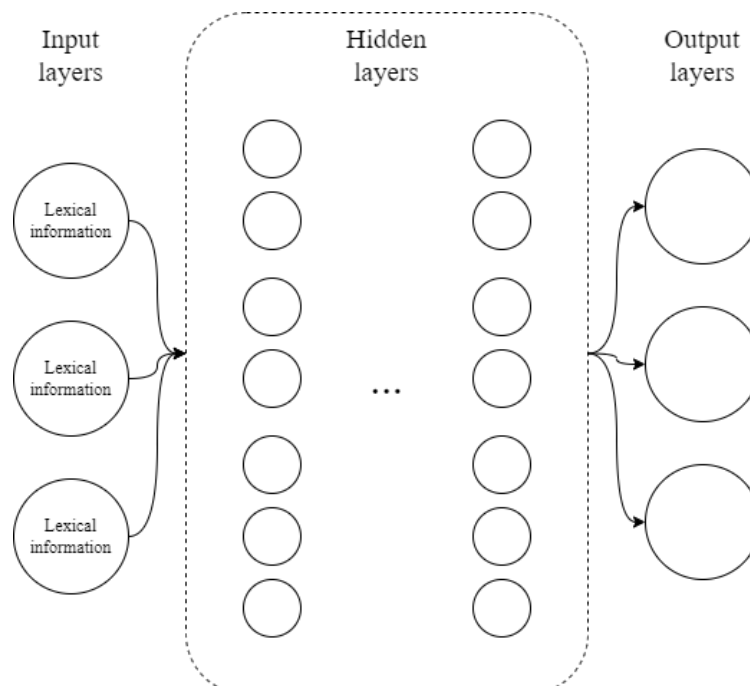


Рисунок 1 – Структура нейронечеткой сети для обработки графической и лексической информации

Входной слой получает графическую и лексическую информацию о ландшафте, а также показания метеорологических станций. Графическая информация может быть представлена в виде спутниковых снимков, аэрофотоснимков, топографических карт и других источников. Лексическая информация включает в себя описания ландшафта, полученные от экспертов или из текстовых источников.

Последующие скрытые слои нейронной сети обрабатывают разнородную информацию, полученную на входном слое. Скрытые слои нейронной сети позволяют эффективно обрабатывать сложные взаимосвязи между графической и лексической информацией.

В результате обработки информации формируется описание ландшафта, распределённое на ранее заданные классы, которое может быть использовано для решения различных задач, таких как прогнозирование изменений ландшафта, определение оптимальных мест для строительства, а также формирования баз знаний для систем оптимизации принятия решений в задачах на данном ландшафте.

Рассмотрим частный случай предлагаемой нейронной сети, результатом которой является набор правил кластеризации ландшафта вида:

P_1 : ЕСЛИ $x_1^{(n)}$ есть {малый} И $x_2^{(n)}$ есть {малый} И $x_3^{(n)}$ есть {малый}
ТО $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$ принадлежит к классу C_1

P_2 : ЕСЛИ $x_1^{(n)}$ есть {малый} И $x_2^{(n)}$ есть {малый} И $x_3^{(n)}$ есть {средний}
 ТО $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$ принадлежит к классу C_2

...

P_{13} : ЕСЛИ $x_1^{(n)}$ есть {большой}
 ТО $x^{(n)} = (x_1^{(n)}, x_2^{(n)}, x_3^{(n)})$ принадлежит к классу C_{13}

Таким образом, предложенная система обработки информации является важным инструментом для изучения и понимания ландшафтных особенностей. Она позволяет эффективно обрабатывать данные, формировать правила кластеризации и получать более точные результаты анализа. Это, в свою очередь, способствует более глубокому пониманию процессов, происходящих в природе, и разработке эффективных стратегий управления природными ресурсами.

Дальнейшие исследования могут быть направлены на расширение возможностей системы обработки информации, улучшение алгоритмов кластеризации, а также применение разработанной системы для решения конкретных задач в области экологии, сельского хозяйства и других областях, связанных с изучением ландшафта.

Список литературы

1. Борисов В.В. Нечеткие модели и сети / В.В. Борисов, В.В. Круглов, А.С. Федулов. – 2-е изд., – М.: Горячая линия – Телеком, 2012. – 284 с.
2. Круглов В. В., Дли М. И. Интеллектуальные информационные системы. – Издательство Физико-математической литературы, 2002. – 227 с.
3. Chaib-draa B., Decharnais J. A relational model of cognitive maps // Int. J. Human-Computer Studies. – 1998 г. – Т. 49. – С. 181-200.

References

1. 1. Borisov V.V., Kruglov V.V., Fedulov A.S. Fuzzy models and networks / V.V. Borisov, V.V. Kruglov, A.S. Fedulov. – 2nd ed. – Moscow: Hot Line – Telecom, 2012. – 284 p.
2. Kruglov V.V., Dly M.I. Intelligent information systems. – Publishing House of Physical and Mathematical Literature, 2002. – 227 p.
3. Chaib-draa B., Decharnais J. A relational model of cognitive maps // Int. J. Human-Computer Studies. – 1998. – Vol. 49. – P. 181-200.