

НЕЙРОННЫЕ СЕТИ В РЕКОМЕНДАТЕЛЬНЫХ СИСТЕМАХ

NEURAL NETWORKS
IN RECOMMENDATION SYSTEMS

A. Fimin
A. Somenkova

Summary. At the present stage of social development, recommendation systems are an integral element providing interaction between the user and the producer of information. Such systems are closely connected with public life and are used for various purposes, including e-commerce, social networks, and mobile phone applications. In the context of the global development of artificial intelligence based on deep learning of neural networks, there is an objective need to analyze the possibilities of using neural networks in recommendation systems.

The author concludes that if compared with traditional models of training recommendation systems, models based on neural networks and deep learning provide great opportunities for recommendation systems. This is because neural networks, due to the possibilities of in-depth analysis of various disparate data, as well as their ability to learn from their mistakes and improve in the learning process, allow them to obtain very accurate data. Hence, recommendation systems can be more adapted to the needs of a particular user and will allow you to select preferences even in those goods, works, services that the user has not previously ordered or was not interested in. However, the main problem in this area is the correct choice of a neural network and training technology, as well as the selection of a data set on which the training of the corresponding neural network will be carried out. The spontaneous choice of a neural network for a recommendation system, without taking into account the above parameters, can lead to errors in the analysis of data by a neural network and its long-distance learning, and as a result lead to incorrect recommendations for users.

Keywords: neural networks, recommendation systems, user preferences, information filtering, data analysis, deep learning.

Фимин Антон Андреевич

Технический руководитель разработки программного обеспечения, IT компания Артком Венчур, Германия
Anton.fimin@outlook.com

Соменкова Анжела Максимовна

Главный инженер по разработке ПАО Сбербанк, Сербия
miss.gelberg93@gmail.com

Аннотация. На современном этапе общественного развития рекомендательные системы являются неотъемлемым элементом, обеспечивающим взаимодействие между пользователем и производителем информации. Такие системы тесно связаны с общественной жизнью и используются для различных целей, среди которых электронная коммерция, социальные сети, приложения для мобильных телефонов. В условиях глобального развития искусственного интеллекта, основанного на глубоком обучении нейронных сетей, возникает объективная необходимость анализа возможностей использования нейронных сетей в рекомендательных системах.

Автор приходит к выводу о том, что, если сравнивать с традиционными моделями обучения рекомендательных систем, модели, в основу которых заложены нейронные сети и глубокое обучение, обеспечивают для рекомендательных систем гораздо большие возможности. Обусловлено это тем, что нейронные сети за счет возможностей глубокого анализа различных разрозненных данных, а также их способности обучаться на своих ошибках и совершенствоваться в процессе обучения, позволяют получать очень точные данные. Отсюда рекомендательные системы могут быть более адаптированы под потребности конкретного пользователя и позволят осуществлять подборку предпочтений даже в тех товарах, работах, услугах, которые ранее пользователь не заказывал или которыми не интересовался. Однако основной проблематикой в данной области является правильный выбор нейронной сети и технологии обучения, а также подбор набора данных, на которых будет осуществляться обучение соответствующей нейронной сети. Стихийный выбор нейронной сети для рекомендательной системы без учета вышеназванных параметров может привести к ошибкам при анализе данных нейронной сетью и ее дальнейшем обучении, и как следствие привести к неправильным рекомендациям для пользователей.

Ключевые слова: нейронные сети, рекомендательные системы, предпочтения пользователей, фильтрация информации, анализ данных, глубокое обучение.

В основе рекомендательных систем лежит фильтрация информации посредством массивного интеллектуального анализа данных, который может решить проблему информационной перегрузки и предоставить контент, который может удовлетворить потребностям пользователей с учетом их персонализации. Кроме того, рекомендательные системы действуют как платформа, помогающая пользователю не только находить необходимую информацию, но и позволяющая рекламодателю представить информацию именно тому пользователю, который в этой информации заинтересован.

В специализированной литературе рекомендательные системы определяют как комплекс сервисов и программ, который анализирует предпочтения пользователей и пытается предсказать, что может их заинтересовать [4, 10] и как модели, предсказывающие предпочтения пользователей по нескольким продуктам в самых разных областях, таких, как видео- и музыкальные сервисы, электронная коммерция и платформы социальных сетей [1, 2]. Таким образом, суть рекомендательных систем сводится к помощи пользователям различных приложений и систем находить товары, работы и услуги, которые их интересуют, исходя из взаимосвязи между ними.

Длительное время существовало всего три вида рекомендательных систем: рекомендательные системы на основе контента (в основу заложены предпочтения пользователя), рекомендательные системы на основе фильтрации (в основу заложены сходства между предпочтениями пользователя), гибридные рекомендательные системы (основаны на сочетании рекомендательных систем на основе контента и на основе фильтрации). Однако развитие технологий искусственного интеллекта и тех возможностей, которые позволяет реализовать искусственный интеллект, обучаемый на основе нейронных сетей, предопределяет необходимость анализа перспектив использования нейронных сетей в рекомендательных системах.

В специализированной литературе отмечается о том, что нейронные сети глубокого обучения являются наиболее подходящими для их использования в рекомендательных системах. Обусловлено это свойствами глубокого обучения. Глубокое обучение — это современное решение для многих задач машинного обучения, таких, как компьютерное зрение или проблемы с естественным языком, и оно превосходит альтернативные методы обучения [10]. Так, нейронная сеть, основанная на технологии глубокого обучения, может изучать не только латентные характеристики пользователей и их предпочтения, но и сложные нелинейные взаимодействия между пользователями и их предпочтениями на основе глубокого анализа [3, 5, 6, 9, 10]. Соответственно, результативность работы рекомендательных систем, основанных на нейронных сетях глубокого обучения, становится гораздо выше результативности традиционно выделяемых рекомендательных систем.

С указанным подходом нельзя не согласиться, поскольку нейронные сети за счет возможностей глубокого анализа различных разрозненных данных, а также их способности обучаться на своих ошибках и совершенствоваться в процессе обучения, позволяют получать очень точные данные. Если адаптировать это к рекомендательным системам, то здесь можно констатировать, что рекомендательные системы, основанные на нейронных сетях глубокого обучения, могут быть более адаптированы под потребности конкретного пользователя и позволяют осуществлять подборку предпочтений даже в тех товарах, работах, услугах, которые ранее пользователь не заказывал или которыми не интересовался.

Основной проблематикой в данной области является правильный выбор нейронной сети и технологии обучения, а также подбор набора данных, на которых будет осуществляться обучение соответствующей нейронной сети. При этом выбор нейронной сети и технологии обучения, а также подбор набора данных должен осуществляться исходя из назначения рекомендательной системы и области ее использования. Так, например, ре-

комендательные системы, используемые в музыкальных приложениях и анализирующие музыкальные предпочтения клиента, будут иметь один набор данных, необходимых для обучения нейронных сетей, а рекомендательные системы, используемые для целей электронной коммерции и выдачи контекстной рекламы, другой набор данных. Стихийный выбор нейронной сети для рекомендательной системы без учета вышеназванных параметров может привести к ошибкам при анализе данных нейронной сетью и ее дальнейшем обучении, и как следствие привести к неправильным рекомендациям для пользователей.

В специализированной литературе исследовались рекомендательные системы на основе сверточных нейронных сетей, где в качестве данных для анализа использовались хэштеги рекомендаций пользователей [8]. Авторы исследования рассматривали хэштег как возможность классификации по нескольким меткам, в то время как сверточная нейронная сеть использовалась в качестве инструмента извлечения информации для учета особенностей, присущих постам пользователей. Предложенная модель состояла из глобального канала и локального канала привлечения внимания. Глобальный канал состоял из сверточного слоя и объединяющего слоя, а локальный канал внимания состоял из слоя внимания и объединяющего слоя. Недостатками использования хэштегов в качестве данных, необходимых для обучения, здесь видится то, что хэштеги не позволяют сделать максимально полный анализ предпочтения пользователей, поскольку связаны только с частью информации в изображении и в тексте.

Исходя из вышесказанного, можно сделать промежуточный вывод о том, что последние тенденции включают в себя применение нейронных сетей на основе глубокого обучения к рекомендательным системам. На практике такие рекомендательные системы уже используются многими крупными международными компаниями, например, такими, как: AirBnB, Facebook, LinkedIn.

Вместе с тем, область рекомендательных систем сложна и возможности использования нейронных сетей в таких системах зависят от множества факторов, в том числе назначения использования рекомендательных систем.

Как уже было отмечено ранее, существует множество различных методов разработки рекомендательных систем, таких, как правила ассоциации, фильтрация на основе контента или совместной фильтрации, матричная факторизация или, например, обучение древовидной модели для прогнозирования вероятности взаимодействия.

Преимуществами же использования нейронных сетей, основанных на глубоком обучении, является то, что

такие модели обеспечивают высокую точность прогноза поведения пользователя и, как следствие, позволяют делать более точные рекомендации [7]. Это достигается за счет того, что глубокое обучение может использовать дополнительные данные при увеличении пропускной способности нейронных сетей, модель может повысить производительность за счет увеличения объема данных.

Кроме того, нейронные сети гибки в своей конструкции. Например, за счет глубокого обучения можно обучить нейронную сеть решению нескольких задач пользователя и анализа его поведения (многозадачное обучение), поставив перед сетью различные вопросы, например, такие, как: «добавит ли пользователь товар в корзину?» или «начнет ли оформление заказа товара?». Каждая цель помогает модели извлекать информацию из данных, при этом цели могут поддерживать друг друга.

Другие подходы к проектированию рекомендательных систем на основе нейронных сетей включают добавление мультимодальных данных в рекомендательную модель. Например, можно решить эту задачу, обработав изображения продукта с помощью сверточной нейронной сети или описание продукта с помощью модели обработки естественного языка. Также необходимо учитывать и тот факт, что фреймворки глубокого обучения высоко оптимизированы для обработки от терабайт до петабайт данных для всех видов доменов, что позволяет повысить эффективную адаптацию нейронных сетей для рекомендательных систем.

Теперь акцентируем внимание на встраивании слоев нейронной сети. Слои встраивания представляют категории с плотными векторами и используются, как правило, для обработки естественного языка: слова со схожим значением имеют похожий вектор встраивания. Эта же методика может применяться к рекомендательным системам. Так, самая простая рекомендательная система будет основана на пользователях и товарах. Вопрос, который будет стоять перед системой: какие товары рекомендовать пользователю? В качестве исходных данных есть идентификаторы пользователей и товаров. Поскольку слова мы заменяем на пользователей и элементы, необходимо использовать две таблицы встраивания (рисунок 1).

Чтобы получить итоговую оценку, то есть вероятность того, что пользователь взаимодействует с элементом, необходимо вычислить точечное произведение между внедрением пользователя и внедрением элемента. В качестве последнего шага для преобразования выходных данных с вероятностью от 0 до 1 можно применить функцию активации сигмоида:

$$\text{Точечный продукт: } u * v = \sum r_i * t_i$$

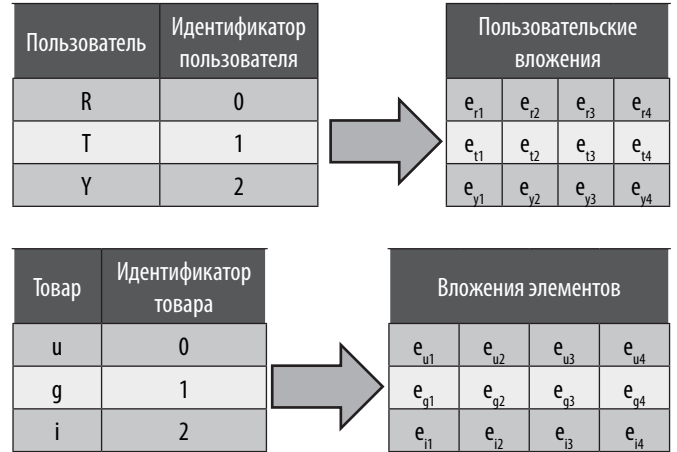


Рис. 1. Встраивание таблиц с размерностью 4

Вышеприведенный пример является элементарной моделью, поскольку в рекомендательных системах можно использовать и более глубокие модели с полностью подключенными слоями. Обусловлено это тем, что производительность нейронных сетей основана на глубоких архитектурах с несколькими нелинейными уровнями. За счет этого можно расширять предыдущую модель, передавая выходные данные встраиваемых слоев через несколько полностью связанных слоев через функцию активации нейросети (ReLU). Здесь основная задача будет заключаться в том, как именно объединить два вектора встраивания (либо объединить только векторы встраивания, либо поэлементно умножить векторы друг на друга, аналогично скалярному произведению).

Ещё одним перспективным направлением является добавление метаданных в нейронную сеть. Так, обычно в качестве входных данных нейронной сети используют только идентификатор пользователя и продукта, как уже было отмечено ранее, однако при анализе данных часто становится доступна дополнительная информация о пользователе или его предпочтениях. Дополнительной информацией о пользователе может быть пол, возраст, город (адрес), время с момента последнего посещения или номер банковской карты, используемой для оплаты, марка товара, ценовая категория товаров и так далее. Вышеуказанная дополнительная информация может использоваться для входных данных нейронной сети и позволит осуществлять более эффективный анализ предпочтений и построение на таких предпочтениях более точных рекомендаций.

Подводя итог, отметим, что, если сравнивать с традиционными моделями обучения рекомендательных систем, модели, в основу которых заложены нейронные сети и глубокое обучение, обеспечивают для рекомендательных систем гораздо большие возможности. Обусловлено это тем, что нейронные сети за счет возможностей глубокого анализа различных разрозненных данных, а также их способности обучаться на своих

ошибках и совершенствоваться в процессе обучения, позволяют получать очень точные данные. Отсюда рекомендательные системы могут быть более адаптированы под потребности конкретного пользователя и позволят осуществлять подборку предпочтений даже в тех товарах, работах, услугах, которые ранее пользователь не заказывал или которыми не интересовался. Однако основной проблематикой в данной области является правильный выбор нейронной сети и технологии обу-

чения, а также подбор набора данных, на которых будет осуществляться обучение соответствующей нейронной сети. Стихийный выбор нейронной сети для рекомендательной системы без учета вышеназванных параметров может привести к ошибкам при анализе данных нейронной сетью и ее дальнейшем обучении, и как следствие привести к неправильным рекомендациям для пользователей.

ЛИТЕРАТУРА

1. Ботов Д.С., Меньшикова Н.В., Портнов И.В. Разработка рекомендательной системы для музыкального сервиса на основе методов контентной и коллаборативной фильтрации // Информационные технологии и системы: труды VI Международной научной конференции (1–5 марта 2017 г.). Челябинск: Челяб. гос. ун-т, 2017. — С. 37–46.
2. Воробьева М.С., Боганюк Ю.В., Якубов Р.М. Разработка рекомендательной системы навыков для студента на основе текстов его работ // Математическое и информационное моделирование: материалы Всероссийской конференции молодых ученых (17–21 мая 2021 г.). — Тюмень: Тюмен. гос. ун-т, 2021. — С. 123–127.
3. Зайцева С.А., Смирнов В.А. Ценностные регулятивы «рекомендательной системы» в контексте распространения нейро-сетевых моделей // Вестник ИвГУ. Серия: Гуманитарные науки. — 2022. — №1. — С.152–160.
4. Исламова С.А., Липатникова Н.В. Обзор рекомендательных систем // Modern Science. — 2019. — № 4–3. — С. 174–176.
5. Колебцев В.И., Белов Ю.С. Метрики, необходимые для составления ранжированного списка новостей для рекомендательной системы // Научное обозрение. Технические науки. — 2021. — № 5. — С. 10–14.
6. Ниничук М.М., Намиот Д.Е. Обзор методов построения рекомендательных систем на основе сессий // International Journal of Open Information Technologies. — 2023. — №5. — С.23–32.
7. Якупов Д.Р., Намиот Д.Е. Рекомендательные системы на основе сессий — модели и задачи // International Journal of Open Information Technologies. — 2022. — №7. — С.128–155.
8. Gong Y., Zhang Q. Hashtag recommendation using attention based convolutional neural networks. Proceedings of the 25th International Joint Conference on the Artificial Intelligence. New York, USA, 2016:2782–2788.
9. Li Y., Liu T., Jiang J. et al. Hashtag recommendation with topical attention-based LSTM. Proceedings of the 26th International Conference on Computational Linguistics. — Osaka, Japan, 2016:3019–3029.
10. Song, Ningxin. (2020). Analysis of Recommendation Systems Based on Neural Networks. Journal of Physics: Conference Series. 1634. 012051. 10.1088/1742-6596/1634/1/012051.

© Фимин Антон Андреевич (Anton.fimin@outlook.com); Соменкова Анжела Максимовна (miss.gelberg93@gmail.com)

Журнал «Современная наука: актуальные проблемы теории и практики»