

Н.В. Багнюк, к.т.н.,
К.Я. Бортник, к.т.н.,
Р.В. Боба

Луцький національний технічний університет, Луцьк

МАТЕМАТИЧНА ТРЕНОВАНА МОДЕЛЬ ДЛЯ АНАЛІЗУ ДАНИХ З ВИКОРИСТАННЯМ MACHINE LEARNING В УПРАВЛІННІ САЙТУ ЕЛЕКТРОННОЇ КОМЕРЦІЇ

Аналіз даних є важливим завданням у різних сферах, зокрема у фінансах, охороні здоров'я та маркетингу. Алгоритми машинного навчання широко використовуються для аналізу даних завдяки їхній здатності навчатися на основі даних і робити прогнози. Продуктивність моделей машинного навчання залежить від кількох факторів, включаючи якість даних, вибір функцій і вибір алгоритмів [1].

Послідовність імплементації:

1. Збір даних: відстежування дій користувачів на сайті.
2. Перетворення даних: дані перетворюються в матрицю.
3. Розбиття даних: дані розбиваються на тренувальну та тестову вибірки.
4. Вибір моделі: вибираємо модель для рекомендацій на основі матричної факторизації, яка дозволяє розкласти матрицю на дві менші матриці.
5. Навчання моделі: застосовуємо алгоритм навчання для навчання моделі на тренувальній вибірці.
6. Тестування та оцінка: оцінюємо точність моделі на тестовій вибірці, використовуючи метрики.

Для рекомендаційних систем в електронній комерції використовуються різні методи machine learning, такі як матрична факторизація, рекурентні нейронні мережі та інші [2].

Для побудови рекомендаційних систем зазвичай використовуються відкриті набори даних, що містять інформацію про покупки, оцінки та відгуки користувачів. Перш за все, необхідно провести підготовку даних, яка включає в себе очистку, обрізку та стандартизацію даних. Наступним кроком є розбиття даних на тренувальну та тестову вибірки.

Далі вибирається модель для рекомендацій. Зазвичай використовуються алгоритми матричної факторизації, такі як

Singular Value Decomposition або Alternating Least Squares. Також можна використовувати deep learning моделі, такі як neural collaborative filtering, яка використовує Convolutional Neural Network для аналізу товарів та користувачів.

Після вибору моделі необхідно провести навчання моделі на тренувальній вибірці. Для цього застосовуються алгоритми навчання, такі як стохастичний градієнтний спуск, який забезпечує оптимізацію параметрів моделі.

Наступним кроком є тестування та оцінка точності моделі на тестовій вибірці. Це можна зробити за допомогою різних метрик, таких як середньоквадратична помилка (RMSE) або середній ранговий показник (MRR).

Після навчання моделі та її перевірки на точність можна розпочати використання її для створення рекомендацій.

Рекомендації на основі користувача передбачають рекомендації для конкретного користувача на основі його попередніх покупок та оцінок. Модель аналізує подібних користувачів та рекомендує товари, які були куплені або оцінені цими користувачами.

Рекомендації на основі товару передбачають рекомендації товарів, які можуть зацікавити користувача на основі подібних товарів, які він купував або оцінював. Модель аналізує подібні товари та рекомендує інші товари, які були куплені або оцінені користувачами, що купували подібні товари.

Однією з основних переваг моделей машинного навчання для рекомендаційних систем є те, що вони можуть бути покращені з часом. Коли користувач робить нові покупки або оцінки, модель може бути знову тренувана з новими даними.

ВИКОРИСТАНІ ДЖЕРЕЛА

1. Ли К., Чен Т., Лю Ю. та Ма Ю. (2019). *Collaborative filtering with social exposure: a modular approach based on graph convolutional networks. Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 425-434.

2.. Alpaydin, E. (2020). *Introduction to machine learning (3rd ed.)*. MIT press.