

*Аронов Владимир Сергеевич
Студент 2 курса магистратуры
института передовых информационных технологий
Тульского государственного педагогического университета им. Л. Н. Толстого
г. Тула, Российская Федерация*

АНАЛИЗ ОЦЕНКИ РЕКОМЕНДАЦИЙ В ЗАВИСИМОСТИ ОТ ТИПА ДАННЫХ, ИСПОЛЬЗОВАННЫХ ДЛЯ ОБУЧЕНИЯ МОДЕЛЕЙ ИИ

Аннотация:

В данной статье проводится анализ оценки респондентами и математическими формулами рекомендаций в зависимости от типа данных, использованных для обучения моделей искусственного интеллекта – реальных (полученных в ходе опроса респондентов) и синтетических, сгенерированных программно. Для оценки рекомендаций был проведен опрос среди респондентов, которые выставили оценку для рекомендаций, полученных от модели ИИ. Так же были использованы формулы точности для моделей ИИ, позволяющие оценить качество предоставленной рекомендации. Данные об оценках были нормализованы для предоставления более наглядной информации в сравнении.

Ключевые слова:

Рекомендательные системы, машинное обучение, рекомендательные технологии

*Aronov Vladimir Sergeevich
Second year master's student
Institute of Advanced Information Technologies
Of Federal State Budgetary Educational Institution of Higher Education «Tula State Lev Tolstoy
Pedagogical University»
Tula, Russian Federation*

ANALYSIS OF EVALUATION OF RECOMMENDATIONS DEPENDING ON THE TYPE OF DATA USED TO TRAIN AI MODELS

Annotation:

This paper analyzes the assessment of recommendations by respondents and mathematical formulas depending on the type of data used to train artificial intelligence models - real (obtained during a survey of respondents) and synthetic, generated by software. To evaluate the recommendations, a survey was conducted among respondents who gave a rating for the recommendations received from the AI model. Accuracy formulas for AI models were also used to evaluate the quality of the recommendations provided. The rating data has been normalized to provide clearer comparison information.

Keywords:

Recommender systems, machine learning, recommendation technologies

Текст статьи:

Введение.

Рекомендательные системы используют различные технологии для создания рекомендаций для пользователей, при этом им необходимы данные для обучения модели ИИ. В данной статье будет проведено сравнение оценок качества рекомендаций, данные респондентами, в зависимости от типа данных, использованных для обучения – полученных от других респондентов (далее – реальные данные) и сгенерированные на основе шаблонов (далее – синтетические).

Данные, генерация синтетических данных, сфера использования модели ИИ

В сравнении используются модели ИИ, предназначенные для последующего создания рекомендательной системы для помощи в выборе курсов дополнительного образования. Данные, которые были получены от респондентов и сгенерированные программно, включают в себя баллы по различным дисциплинам науки, предпочтениям по курсам доп. образования, увлечениям, оценкам фильмов и сериалов разных жанров. На рекомендацию так же влияют данные о составе курсов доп. образования и их свойствах (смежные дисциплины, сфера науки и т.д.). Всего было использовано три модели ИИ для этой задачи: на основе библиотеки LightFM, на основе библиотеки Surprise SVD и на основе библиотеки TFRS, поскольку они подходят для использования в рекомендательной системе [1]. Генерация синтетических данных производится путем ввода значений, подходящих для одного из возможных типажей человека, которые имеют ярко выраженную предрасположенность к некоторым наукам [2].

Оценка точности рекомендаций на основе формул точности, применяемых для оценки в различных моделях искусственного интеллекта:

Для оценки точности рекомендаций возможно применение формул, которые оценивают метрики, использованные в ходе обучения модели ИИ. Для регрессионных моделей ИИ будут использованы следующие формулы:

1. Mean Squared Error (MSE): среднеквадратичное отклонение между значениями предсказания и истинного значения (1);
2. Mean Absolute Error (MAE): среднее абсолютное отклонение (2);
3. R^2 (коэффициент детерминации): мера соответствия предсказываемых данных фактическим (3).

Формулы для математической оценки точности в ходе обучения модели:

$$MSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (1)$$

где n – общее количество наблюдений;

y_i – истинное значение для наблюдения i (где i – порядковый номер наблюдения);

\hat{y}_i – предсказанное значения для наблюдения i (где i – порядковый номер наблюдения).

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (2)$$

где n – общее количество наблюдений;

y_i – истинное значение для наблюдения i (где i – порядковый номер наблюдения);

\hat{y}_i – предсказанное значения для наблюдения i (где i – порядковый номер наблюдения).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \quad (3)$$

где n – общее количество наблюдений;

y_i – истинное значение для наблюдения i (где i – порядковый номер наблюдения);

\hat{y}_i – предсказанное значения для наблюдения i (где i – порядковый номер наблюдения).

\bar{y} – среднее истинное значение, которое высчитывается по формуле (4).

$$\bar{y} = \frac{1}{n} \sum_{i=0}^n y_i, \quad (4)$$

где n – общее количество наблюдений;

y_i – истинное значение для наблюдения i (где i – порядковый номер наблюдения);

Для оценки точности рекомендаций, полученных в ходе работы модели ИИ, можно использовать следующие метрики:

1. Accuracy (5);
2. Precision (6);
3. F1 Score (7).

Формулы для оценки точности рекомендаций, полученных от модели ИИ:

$$\text{Accuracy} = \frac{\text{True positive of correct predictions}}{\text{Total number of predictions}}, \quad (5)$$

где *True positive of correct predictions* – количество истинных положительных предсказаний;

Total number of predictions – общее количество предсказаний.

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}, \quad (6)$$

где *True positive of correct predictions* – количество истинных положительных предсказаний;

False positives – количество ложных положительных предсказаний.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (7)$$

где *Precision* – метрика, вычисляемая в формуле 6;

Recall – метрика, вычисляемая в формуле 8.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}, \quad (8)$$

где *True positives* – количество истинных положительных предсказаний;

False negatives – количество ложных отрицательных предсказаний.

Для нормализации полученных данных используется шкала от 0 до 5, где 5 – максимально возможный результат (метрика имеет максимально возможный результат), а 0 – минимально возможный результат (метрика имеет минимально возможный результат). В шкале так же используются дробные значения (например, 4.2).

Для оценки респондентами предлагается оценить рекомендацию в плане точности по шкале от 0 до 5, респонденты – обычные люди, которым было предложены рекомендации, полученные после обучения моделей ИИ на двух видах данных. В шкале так же используются дробные значения (например, 4.2).

Результаты оценки, сравнение показателей:

Результаты, полученные в ходе эксперимента, приведены в таблицах 1, 2, 3 и 4.

Таблица 1. Оценка эффективности моделей ИИ респондентами, реальные данные

| Параметр/Модель | LightFM | Surprise SVD | TFRS |
|-------------------------------------|---------|--------------|------|
| Точность рекомендаций, средний балл | 4.5 | 3.9 | 4.3 |
| Точность рекомендаций, низший балл | 3 | 3 | 4 |
| Точность рекомендаций, высший балл | 5 | 5 | 5 |

Таблица 2. Оценка эффективности моделей ИИ респондентами, синтетические данные

| Параметр/Модель | LightFM | Surprise SVD | TFRS |
|-------------------------------------|---------|--------------|------|
| Точность рекомендаций, средний балл | 4.1 | 3.7 | 3.8 |
| Точность рекомендаций, низший балл | 2 | 2 | 3 |
| Точность рекомендаций, высший балл | 5 | 5 | 5 |

Таблица 3. Оценка эффективности моделей ИИ программно, реальные данные

| Параметр/Модель | LightFM | Surprise SVD | TFRS |
|-------------------------------------|---------|--------------|------|
| Точность рекомендаций, средний балл | 4,6 | 4,6 | 4.5 |
| Точность рекомендаций, низший балл | 4 | 4 | 4 |
| Точность рекомендаций, высший балл | 5 | 5 | 5 |

Таблица 4. Оценка эффективности моделей ИИ программно, синтетические данные

| Параметр/Модель | LightFM | Surprise SVD | TFRS |
|-------------------------------------|---------|--------------|------|
| Точность рекомендаций, средний балл | 4.8 | 4.9 | 4.7 |
| Точность рекомендаций, низший балл | 4 | 4 | 4 |
| Точность рекомендаций, высший балл | 5 | 5 | 5 |

В ходе оценки эффективности каждой из моделей были сделаны следующие выводы:

1. Оценка, полученная от респондентов, отображает более высокий балл для данных, которые были собраны в результате опроса респондентов по сравнению с синтетическими данными;
2. Оценка, полученная в ходе программной оценки эффективности отображает выше для синтетических данных по сравнению с данными, полученными в ходе опроса респондентов;
3. Программный метод оценки эффективности имеет меньший разброс для всех трёх моделей, чем полученный в ходе опроса респондентов.

Заключение.

В результате оценки точности были сделаны выводы о том, что респонденты оценивают рекомендации, полученные в ходе работы модели ИИ на основе реальных данных выше, чем рекомендации, полученные от модели ИИ, обученной на синтетических данных.

При этом при программном методе оценки точности рекомендаций возникает обратная зависимость: для синтетических данных средний балл будет выше, чем для реальных данных. Таким образом, для рекомендаций, предназначенных для людей, рекомендуется обучать модель ИИ на реальных данных, и математическая оценка эффективности не всегда будет коррелировать с результатами, которые будут получены после опроса пользователей. Генерация синтетических данных может не учитывать скрытые факторы, связанные с психологией человека, поэтому для создания синтетических данных, при помощи которых можно обучить модель ИИ для достижения той же оценки респондентами, что и в случае обучения на реальных данных, потребуется проведение дополнительных исследований и экспериментов.

Использованные источники

1. Попкова А.А. СРАВНЕНИЕ МОДЕЛЕЙ / А.А. Попкова // Культура. Наука. Производство. 2023. № 11. С. 20-29.
2. Б.М. Теплов. Проблемы индивидуальных различий. - Москва : Издательство Акад. пед. наук РСФСР, 1961, 536 с. 2.1.