Abstract

In our thesis, we focus on the cold start user problem in the field of recommender systems. We propose a mechanism called “User Rejuvenation” and combined it with the deep learning model Denoising Autoencoder to solve the cold start user problem. The “User Rejuvenation” first choose representative items for each group of users, after that we randomly set the dimensions corresponding to representative items in user vector to zero score with lower probability, and we randomly set the dimensions corresponding to non-representative items in user vector to zero score with higher probability. The main purpose for “User Rejuvenation” is to turn the non-cold start user vectors back to cold start user vectors for each group of users. After getting the group-specific cold start user vectors generated from “User Rejuvenation” mechanism, we can use them to train a Denoising Autoencoder model for the user group. When the training process is complete, the model will have capacity for restoring the cold start user vectors to non-cold start user vectors, and the recommendation is made through the predicted non-cold start user vectors.

Introduction

1. The importance of the recommendation system

With the explosive growth of information on the Internet, the dizzying selection of goods or services in various service platforms on the Internet has already put users in a dilemma that they cannot cope with. At this time, the recommendation system has become an indispensable existence that can effectively filter out a large amount of unnecessary information for users quickly and correctly, and help users focus on the information they are concerned about. In addition to allowing users to quickly and accurately obtain products that meet their needs, the recommendation system can also improve user satisfaction and increase user stickiness to the platform through accurate recommendations, which brings platform vendors considerable benefits. Therefore, applying the recommendation system to practice is prevalent in many e-commerce platforms (Amazon) or audio-visual streaming platform (Netflix), that help companies improve user satisfaction and increase platform revenue. Taking Netflix as an example, their chief product officer Hunt said that more than 80% of Netflix movie viewings are generated through their recommendation system. As Netflix’s Gomez-Uribe and Hunt mentioned in [5], “If we can increase the user’s stickiness on the platform by improving the quality of the Netflix recommendation system, then Netflix can reduce $1 billion in losses every year caused by customer churn.” The point of reducing losses by maintaining customer loyalty coincides with the point made by Forbes columnist Larry Myler, “For retailers, the cost of retaining existing customers is often much lower than the cost of seeking for new customers, and having a stable customer base can enable retailers to obtain higher customer lifetime value, as well as business revenue is much easier to predict.” Under the same circumstances, for the Netflix platform, it is easier to reduce customer churn than to develop new customers, because new customers often decide whether to join the platform after browsing less than 20 movie recommendations. As a result, Netflix can only predict the preferences of new customers with very little information and time. Compared with recommending videos that meet their expectations from past ratings information of existing customers, it is much more practical to improve the quality of the recommendation system. From the case of Netflix, we can realize that a well-designed recommendation system can bring benefits to both users and service platform providers.

1. The difficulty and importance of the cold start user problem

Generally speaking, when constructing a recommendation system, a commonly used method is “collaborative filtering recommendation”. The concept of collaborative filtering recommendation is using the user group with similar product preferences to the target user who you wants to recommend to predict the products that the target users may also prefer. For example, in the Amazon online bookstore, when a customer buys a book, the bottom of the website will display “customers who bought this book also bought”, this is a classic case of collaborative filtering recommendation. However, there is a problem with collaborative filtering recommendation. While recommending a product to a new user, because there is only a little rating information about the new user in the past, it is difficult for us to find the user group with similar preference to the new user, which makes it impossible to recommend using collaborative filtering recommendation. Such a problem is called “new user cold start problem”. In the past, various methods have been proposed to solve this problem, one of which is the “content-based method” [16]. The content-based method combines the user’s additional information, such as the user’s profile, age, or gender, etc., and we can replace the original user rating vector without rating information with the user representation vector constructed by these additional information. Then, the recommendation can be completed by calculating the vector similarity such as cosine similarity to find the products that the users who are highly similar to the new user vector like. However, the content-based recommendation method must be based on the premise that we can obtain additional information from users. But in practice, we may not be able to obtain this information. Therefore, there is another method to solve the cold start problem of new users, called “representative-based method” [8][4][12]. The core concept of this method is to find “representative items” in the recommendation system, and if we can obtain some ratings of representative items from new users, then we can indirectly infer the relationship between new users and other general items through the relationship between representative items and general items to make recommendations.

As a result, our goal is to propose a framework that can effectively solve the cold start problem of new users, so that we can recommend items that meet the needs of new users, which will not only enhance the adhesion of new users on the platform but also increase customer engagement on the platform. In this way, new users can gradually get rid of the cold start state and get more accurate item recommendations, and finally form a positive interaction cycle between the user and the platform.