Restaurant Recommendation System Based on User Preference

V Priyanka
M.Tech, AI & DS
Muthoot Institute of Technology and Science
Kochi, India
priyankarajith98@gmail.com

Ms.Sreenu G
Asst. Professor
Muthoot Institute of Technology and Science
Kochi, India
sreenug@mgits.ac.in

***Abstract*—Recommendations are integral parts of our life.They are important information filtering mechanisms that make our choices easier.Here a restaurant recommendation system is proposed which uses three types of recommendations.They are Content based,Location based and a hybrid based filtering method.Content based and hybrid based method provide recommendations for existing users and location based method provide restaurants for new users in the city.**

***Index Terms*—Recommendation,Content based,location based,Hybrid,LighFM**

I. Introduction

In the current world,there is massive amount of data around us.The quantity of information and resources used in our daily life are increasing day by day.There is an imperative need to filter information around us.A system that removes duplicate or undesired information from an information stream using a computerised processes is referred to as an information filtering system.Recommender systems or recommender engines provide solution to information excess problem by providing personalization of data.Recommendation systems are found in almost every industry and domain.This includes e-commerce,retail,media,banking,telecom,food,music,etc.

Recommendation systems predict what users want by assessing their actions which contains information on past preferences and are likely to foresee products or items that are more interesting to them.It is more personalized and unique.Recommendations are found everywhere.It can be in the form of recommending friends on social networking sites like Facebook,products based on your cart items in e-commerce sites like amazon, movie recommendations or even recommending songs.With the abundance of information available it can be difficult for customers to navigate through different items and recommendations help customers by making it simpler,less time consuming and more effective.

Recommendations are pivotal for companies as they improve click rates, revenues and other measures of success with successful recommendations.For examples for companies Amazon and Netflix it is important to know customer's and prospects' profiles, including demographics, geography.It is also important to deliver the right promotion, content, recommendation for a customer based on actions, preferences, and shared interests.Personalization provides several advantages for individuals.It simplifies consumers' lives since they only view content that is relevant to them.For businesses,it enhances customer experience by reducing clutter and showing only wanted content.Relevant product offers based on consumer choices can lead to more product awareness and, ultimately, greater product sales.

There are mainly 3 types of recommendation systems; Collaborative filtering,Content-based filtering and hybrid based filtering.Based on the preferences of many other users, collaborative filtering based recommendation system predicts what would be of interest to a person.For example if user A likes item X and user B likes item X and item Y then it is likely that user A likes item Y as well.The primary drawback to this approach is that the model's prediction for a particular user/item combination is the dot product of the respective embeddings. As a result, if an item is not observed during training, the system cannot generate an embedding for it and so cannot search the model with it.This problem is called the cold-start problem.Content based recommenders focuses on the items themselves and suggests additional products with comparable characteristics.The downside of this method is that it relies on item features only and not on user preferences.Hybrid filtering is the integration of the two methods and helps to to improve performance while minimising the downsides of each.Hybrid recommenders can help with existing problems of pure recommenders like cold start problems and sparsity.

In this paper a hybrid method,content based and a location based recommendation system is implemented.LightFM [8] is a hybrid matrix factorisation model represents users and items as linear combinations of the latent factors of their content features.In cold-start or sparse interaction data settings (using both user and item metadata), the model outperforms both collaborative and content-based models, and performs at least as well as a pure collaborative matrix factorisation model when interaction data is available.

II. Review of Literature

In [1] a latent factor model is proposed based on probabilistic matrix factorization and incorporates implicit feedback as complementary information.Both implicit and explicit factors are integrated with a matrix co-factorization algorithm.The explicit feedback matrix is divided into the product of the user latent factor matrix and item true latent factor matrix. The implicit feedback matrix is divided into the product of the user latent factor matrix and item display latent factor matrix. The explicit and implicit feedback matrices are decomposed into a shared subspace simultaneously.After that the latent factor vectors are jointly optimized using a gradient descent algorithm.

In [2] the model uses the Variational autoencoder to extract the characteristics of users and items, so that this neural network model can better find the linear relationship between the user and the item and the nonlinear relationship.It combines neural collaborative filtering and variable automatic-encoder.The Variable automatic-encoder (VAE) is a variant of autoencoder. The main structures of the two encoders are similar, they are composed of a decoder and an encoder. The difference between the two is that the function of the VAE is to convert the input to a Gaussian distribution instead of generating a simple hidden factor vector. The VAE then randomly extracts the same signal from the input of the decoder from the Gaussian distribution described above to obtain an output that is consistent with the original input distribution.The NCF is composed of an input layer, an embedded layer, neural CF layer, and output layer. The output of one layer is used as the input to the next layer.

In [3], the paper proposes a model that combines localization, personalization and content-based recommendation in a dynamic environment.The system has two sections, one which has online activity, and the other which processes data offline. When the user is in motion, his geo-position changes notably, the system goes online and the recommendation module becomes active, retrieving nearby restaurants and ranking them, based on their properties, according to the scores generated offline. The offline part generally remains in a non-functional mode when the user is stationary The work of the offline system is to generate a user interest profile, using a Machine Learning algorithm, from the data set that keeps getting modified whenever the user checks-in a restaurant. There is autonomous switching between online and offline mode.Restaurants are classified as 'like' or 'dislike', depending on the taste of the user.Naive Bayes Classifier algorithm to recognize the factors that the user likes about a restaurant and to what degree does the person likes them.

In [4] the system was designed and implemented using a support vector machine for restaurant recommendation.User Interfaces for the customers and restaurant staff will be provided and customers can order food directly through the module without interacting with the waiter. Customer gives either 'explicit feedback' i.e. clear and confirmatory information given through actions like rating the product etc. or through the 'implicit feedback' i.e. user gives opinion about products rather than committing action. Different suggestions about the meal given to individuals so that they can choose the meal according to their choice. There is a facility to book a table in advance . By using the concept of Time -Series analysis system provides the availability of the table. so that customers can book their table in advance.This makes sure that the customer is sitting on the table and then automatically the device will send the information to the device.After that customer can place an order through the app on their mobile device. At the end, the customer can make the payment via phone. The SVM model after training will be deployed on the server which classifies the reviews given by the customers into two classes positive and negative and give recommendations of the food to the customers according to the preferences set by them. All the necessary statistical calculations are handled by the server.

In [5] a user-based collaborative filtering method is proposed for recommending a restaurants personally, based on ratings given by other users.A user rating similarity and user attribute similarity is used for finding the proximity between users.In the beginning a user-item matrix is build that is obtained from the questionnaire given to users by giving ratings of restaurants that users have visited. Then the user-item matrix is used to calculate similarity to determine the similarity between users.Similarity is calculated to look at active user similarity with other users, so that active-user have similarity with all users. The higher the similarity value, the more similar the preferences between the two users. otherwise, the smaller the similarity value, the preference of the two users is not similar.Similarity calculations are carried out in two stages, calculating the user similarity and calculating the user attribute similarity.Weighted coefficient is used to get result of these two similarities To get restaurant recommendations from other users, the similarity calculation is used to get similarity between users that can recommend restaurants according to our taste.

III. Proposed Method

The preprocessing steps consisting of cleaning the three yelp datasets,namely the user dataset,business dataset and the review dataset.A sentiment analysis is performed on the review dataset to generate a super score.Here three recommender systems are introduced specifically; Content based,location based and hybrid based model.

A. Sentiment analysis

Initially all datasets were cleaned separately and then merged together.Sentiment analysis were performed on reviews.The reviews were classified into three categories,positive,negative and neutral.Textblob was used to classify text.Textblob is a python library that perform basic NLP tasks.When a sentence is passed into Textblob it gives two outputs,polarity and subjectivity.Polarity is a float that lies in the range of [-1, 1] where 1 means positive statement and -1 means negative statement.The degree of personal opinion and factual information in a text is measured by subjectivity. Because of the text's greater subjectivity, it provides personal opinion rather than objective information.

VADER or Valence Aware Dictionary and Sentiment Reasoner is a lexicon and rule based sentiment analysis tool that is responsive to feelings expressed on social media.VADER employs a mix of techniques. A emotion lexicon is a collection of lexical properties that are classified as positive or negative depending on their semantic orientation. VADER not only tells us about the positivity and negativity scores, but also about the intensity of emotion and how positive or negative a sentiment is.The compound score is a metric that calculates the sum of all the lexicon ratings which have been normalized between -1 (most extreme negative) and +1 (most extreme positive).The super score is calculated by the sum of stars/ rating and the product of polarity and compound.This score represents the overall score of customers experience.

B. Content based recommender

A simple content based recommender based on similarity is being implemented.To produce these recommendations, the algorithm considers information such as the restaurant's description, cuisine, and attributes of restaurant.
Cosine similarity is a mathematical relationship used to measure the similarity between sentences.The cosine of angle between two sentences is used to measure or quantify how similar two vectors are.Here it is used to determine how to restaurants are similar to each other.The method returns a number ranging from -1 (completely opposing vectors) to 1 (the same vector). A value of 0 implies that the vectors have no correlation, whereas intermediate values imply intermediate levels of resemblance.
A method to compare restaurants based on sentiments like positive negative and neutral has also been incorporated for better choice of user.This way the user can compare restaurants and make choices based on it.
The major idea for this recommendation algorithms is that if a person loves one item, he or she will also enjoy another comparable to it. If a person has previously visited an Italian restaurant and enjoyed it, the recommendation system will suggest additional good Italian restaurants for that person to visit.The recommended restaurants along with their super score are being displayed.The super score is the sum of product of polarity and compound with the stars of each restaurant.

C. Location based Recommender

For a person who is new to a city it can be difficult to find recommendations for food.In this recommender, new users can get recommendations based on their current address.Using K-means clustering,it divides the restaurants into 10 clusters based on similarity.Based on the the cluster the user is,recommendations of restaurants are provided along with rate and review count.The number of clusters are determined by the elbow method.

D. Hybrid based Recommender

Hybrid recommenders make use of both content and collaborative filtering methods.The hybrid model used here is LightFM [8] .It is a hybrid matrix factorisation model that represents users and things as linear combinations of their

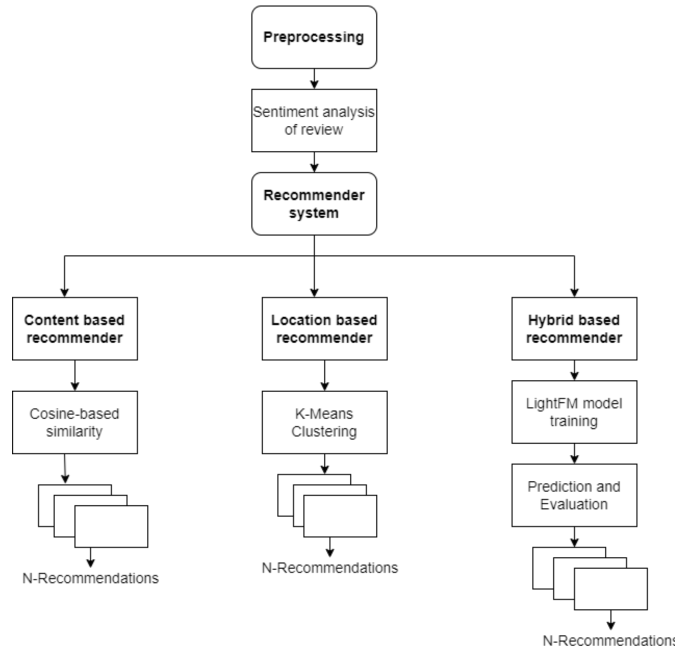

Fig. 1. Workflow

content characteristics' latent components. In cold-start or sparse interaction data settings the model outperforms both collaborative and content-based models, and performs at least as well as a pure collaborative matrix factorisation model when interaction data is available.
The model learns embeddings or latent representations in a high-dimensional space for users and items in such a way that user preferences over items are embedded. When these representations are multiplied together, they yield scores for each item for a certain user.Items with high scores are more likely to be attractive to the user.The user and item representations are defined in terms of feature representations.An embedding is estimated for each feature and these features are then combined to arrive at user and item representations.LightFM embeddings encode crucial semantic information about features and may be utilised for related recommendation tasks such as tag suggestions.
In our system it takes in a user id as input and gives recommendations on the basis pf known positives.Here known positives refer to the restaurants that he or she have liked previously.For example if a person has visited Chinese restaurants mostly, the recommendations will also consist of Chinese food.If there is not enough known positives it will return the most popular restaurant categories.There should be a minimum number of known positives to give recommendations to.

IV. Result and Evaluation

For evaluation standard AUC or area under graph method is used.The AUC score shows good score and hence we can say that it gives good performance.

Metrics Used	Score
Hybrid train AUC	99.3
Hybrid test AUC	98.5

Fig. 2. AUC Score

```
User 1
length of known_positives: 8
Known positives:
 Urban Pantry | Do-It-Yourself Food, Restaurants, American (Traditional), Active Life, Grocery, Food,
 Ocean King Market | Food, Seafood Markets, International Grocery, Specialty Food, Ethnic Food
 Brian's Brew | Food, Coffee & Tea
 Take Five Café | Food, Coffee & Tea, Restaurants, Cafes
 Zaika Indian Contemporary Cuisine | Restaurants, Indian
 HOTLIPS Pizza - Hawthorne | Food, Beer, Wine & Spirits, Restaurants, Fast Food, Pizza, Gluten-Free
 Greater Goods Coffee Roasters | Coffee Roasteries, Cafes, Food, Coffee & Tea, Restaurants
 A Thai Basil | Food Stands, Food, Thai, Food Trucks, Bubble Tea, Restaurants
Recommended:
 Let's Roll Custom Sushi Bar | Restaurants, Canadian (New), Sushi Bars, American (New)
 Nana's Ice Cream Scoop Shop | Food, Ice Cream & Frozen Yogurt
 Brian's Brew | Food, Coffee & Tea
This one clicked
 Taqueria So Mexican | Restaurants, Mexican, Food Trucks, Food, Food Stands
 Jackson's Poultry | Food, Specialty Food, Farmers Market, Butcher, Meat Shops
k_p: 8
precicion at k : 0.2
```

Fig. 3. Output

V. CONCLUSION AND FUTURE SCOPE

Recommendations for existing users based on cuisine have been generated.The recommendation system helps new user sort restaurants based on location.Existing users can get restaurants similar to ones they prefer based on known interactions.LighFM model can be further explored for suggesting recommendations for new users for more personalization.

REFERENCES

[1] Chen, Shulong, and Yuxing Peng. ”Matrix factorization for recommendation with explicit and implicit feedback.” Knowledge-Based Systems 158 (2018):

[2] Liang, Dawen, et al. ”Variational autoencoders for collaborative filtering.” Proceedings of the 2018 world wide web conference. 2018.

[3] Gupta, Anant, and Kuldeep Singh. ”Location based personalized restaurant recommendation system for mobile environments.” 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI). IEEE, 2013.

[4] Pavate, Aruna, et al. ”Cuisine Recommendation, Classification and Review Analysis using Supervised Learning.” 2020 International Conference on Convergence to Digital World-Quo Vadis (ICCDW). IEEE, 2020

[5] Fakhri, Alif Azhar, Z. K. A. Baizal, and Erwin Budi Setiawan. ”Restaurant Recommender System Using User-Based Collaborative Filtering Approach: A Case Study at Bandung Raya Region.” Journal of Physics: Conference Series. Vol. 1192. No. 1. IOP Publishing, 2019

[6] Sahu, Himanshu, Neha Sharma, and Utkarsh Gupta. ”A New Framework for Collecting Implicit User Feedback for Movie and Video Recommender System.” Recent Trends in Communication, Computing, and Electronics. Springer, Singapore, 2019.

[7] Truong, Quoc-Tuan, Aghiles Salah, and Hady W. Lauw. ”Bilateral variational autoencoder for collaborative filtering.” Proceedings of the 14th ACM International Conference on Web Search and Data Mining. 2021.

[8] Kula, Maciej. ”Metadata embeddings for user and item cold-start recommendations.” arXiv preprint arXiv:1507.08439 (2015).

[9] Ladda, Aayush, et al. ”Recommendation System Using Hybrid Filtering.” (2021).

[10] Valerio, Dylan M., and Prospero C. Naval. ”PRTNets: Cold-Start Recommendations Using Pairwise Ranking and Transfer Networks.” Asian Conference on Intelligent Information and Database Systems. Springer, Cham, 2020.

[11] Lee, Sunhwan, Anca Chandra, and Divyesh Jadav. ”An empirical study on hybrid recommender system with implicit feedback.” Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Cham, 2019.