



e-ISSN: 2319-8753 | p-ISSN: 2320-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 5, May 2021

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 7.512**

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)



# Implementation of App Recommendation System by Learning Users Niche from Social Media with User Input

Prof. Nandini Babbar<sup>1</sup>, Kaustubh Warankar<sup>2</sup>, Shubham Gadiwan<sup>3</sup>, Priyanka Shinde<sup>4</sup>,  
Pournima Sonawane<sup>5</sup>

Professor, Department of Computer Engineering, NBN Sinhgad School of Engineering, Pune, India<sup>1</sup>

U.G. Student, Department of Computer Engineering, NBN Sinhgad School of Engineering, Pune, India<sup>2</sup>

**ABSTRACT:** the popularity and development of mobile devices and mobile apps have dramatically changed human life. Due to the tremendous and still rapidly growing number of mobile apps, helping users find apps that satisfy their demands remains a difficult task. To address this problem, we propose a personalized mobile app recommender system based on the textual data of user social Media i.e., public accessible tweets, which can also reflect user's interest and make up for the sparsity of app usage data. Topic modelling techniques are applied to extract topics of user social media data, and the probability distributions of the topics are utilized to represent the features of the apps. Then, the user profile is constructed based on the user's interest to capture user preferences. Both the topic distributions of the apps and user preferences are taken into account to produce recommendation scores to generate recommendation lists for target users. We crawl real-world data sets from Twitter to evaluate the performance. The experimental results show that user social data i.e. tweet is effective for deriving the user interest, and the proposed app recommender system improves the performance of existing approaches.

**KEYWORDS:** App recommendation, social media, transfer learning, collaborative filtering.

## I. INTRODUCTION

The popularity and advancement of smartphones and mobile apps have brought considerable convenience to human life. A substantial number of mobile apps have been developed to assist users in various types of tasks, such as production, entertainment, and exercise. Millions of apps are available for downloading from the Apple App Store and Google Play Store. With such a tremendous and still rapidly growing number of mobile apps, users face difficulty in finding appropriate apps, and only a few apps have the chance to be exposed to users. Hence, helping users to find apps that meet their demands is a critical issue.

Recommender systems have been widely adopted by many online websites to help customers overcome the information overload problem and make their purchase decisions. Popular recommendation techniques, such as collaborative filtering (CF) and content-based filtering, can be utilized in app stores for users to search relevant apps. CF makes recommendations based on the assumption that customers with similar ratings and interests for some items are likely to have similar preferences for other items. Content-based filtering assumes that the users' interests are able to be represented by the content of items they have shown interest in, and those items that have content descriptions similar to the target user's favourite items are recommended.

In this paper, we capitalize on user reviews to understand the functionalities of apps from the users' perspective and leverage user reviews to develop an app recommender system. The topics hidden in the review texts can be a type of representation of the app features. We collect user reviews from app stores to perform topic modelling and represent each app as the probabilities of topic distributions. Then, user preference is inferred from the topic distribution of the user's consumed apps to construct the user profile. Since there are features the user will like or dislike, we further consider the sentiments related to topics, which helps to find favourable apps and avoid disliked apps. The topic distributions of apps and user preferences are both considered when producing recommendation scores of the relevant apps for the user to make personalized recommendations.

## II. RELATED WORK

Sr No	Paper Title	Authors	Publication Year	Algorithms	Conclusion
1	Personalized Mobile App Recommendation by Learning User's Interest from Social Media	Zhen Tu, Yong Li, Pan Hui, Li Su, Depeng Jin	IEEE TRANSACTIONS ON MOBILE COMPUTING MAY 2020	matrix factorization	In this paper, author demonstrate that it is possible to make personalized app usage estimation by learning user's app preference from the social media, i.e., public accessible tweets, which can also reflect user's interest and make up for the sparsity of app usage data. By proposing a novel generative model named IMCF+ to transfer user interest from rich tweet information to sparse app usage, authors achieve personalized app recommendations via learning the interest's correlation between apps and tweets.
2	DPLink: User Identity Linkage via Deep Neural Network From Heterogeneous Mobility Data	Jie Feng , Mingyang Zhang , Huandong Wang , Zeyu Yang , Chao Zhang, Yong Li , Depeng Jin	ACM 2019	LSTM	In this paper, author propose DPLink, an end-to-end deep learning based framework, to complete the user identity linkage task for heterogeneous mobility data collected from different services with different properties.
3	Smartphone app usage prediction using points of interest	DONGHAN YU, YONG LI, FENGLI XU, PENGYU ZHANG, VASSILIS KOSTAKOS	ACM 2018	matrix factorization techniques	In this paper author present the first population-level, city-scale analysis of application usage on smartphones. Using deep packet inspection at the network operator level, we obtained a geo-tagged dataset with more than 6 million unique devices that launched more than 10,000 unique applications across the city of Shanghai over one week. We develop a technique that leverages transfer learning to predict which applications are most popular and estimate the whole usage distribution based on the Point of Interest (POI) information of that particular location
4	Connecting Social Media to E-Commerce: Cold-Start Product Recommendation using Microblogging Information	Wayne Xin Zhao, Sui Li, Yulan He, Edward Y. Chang, Ji-Rong Wen, and Xiaoming Li	IEEE 2016	recurrent neural networks	Author proposes to use the linked users across social networking sites and e-commerce websites (users who have social networking accounts and have made purchases on e-commerce websites) as a bridge to map users' social networking features to another feature representation for product recommendation. In specific, we propose learning both users' and products' feature representations from data collected from e-commerce websites using recurrent neural networks.
5	Learning graph-based poi embedding for location-based recommendation	Min Xie† Hongzhi Yin‡ * Hao Wang† Fanjiang Xu† Weitong Chen‡ Sen Wang	ACM 2016	metric embedding method, graph-based embedding	Author propose a generic graph based embedding model, called GE, in this paper. GE jointly captures the sequential effect, geographical influence, temporal cyclic effect and semantic effect in a unified way by embedding the four corresponding relational graphs into a shared low dimensional space. Then, to support real-time recommendation, authors develop a novel time-decay method to dynamically compute the user's latest preferences based on the embedding of his/her checked-in POIs learnt in the latent space.

### III. OPEN ISSUES

Lot of work has been done in this field because of its extensive usage and applications. In this section, some of the approaches which have been implemented to achieve the same purpose are mentioned. These works are majorly differentiated by the algorithm for Mobile App Recommendation systems.

The popularity and advancement of smartphones and mobile apps have brought considerable convenience to human life. A substantial number of mobile apps have been developed to assist users in various types of tasks, such as production, entertainment, and exercise. Millions of apps are available for downloading from the Apple App Store and Google Play Store. With such a tremendous and still rapidly growing number of mobile apps, users face difficulty in finding appropriate apps, and only a few apps have the chance to be exposed to users. Hence, helping users to find apps that meet their demands is a critical issue.

### IV. PROPOSED SYSTEM

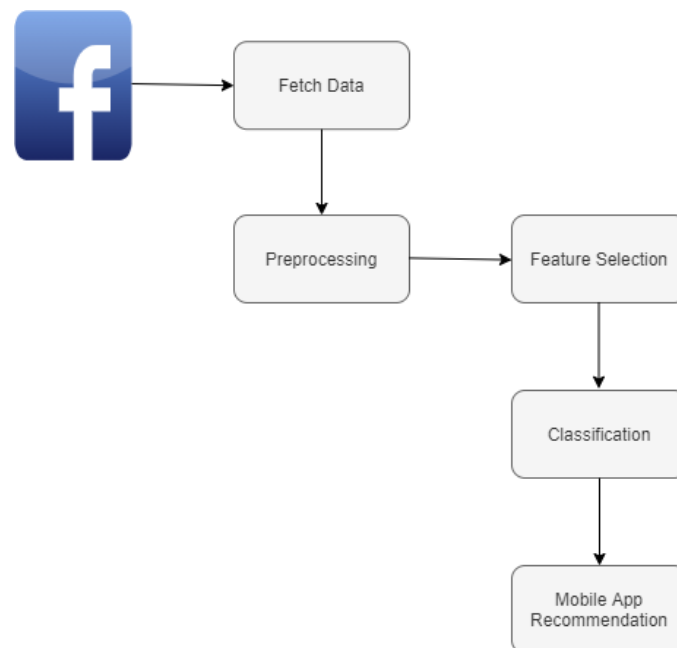


Figure. System Architecture

- Step 1: Facebook Data collection using Facebook streaming API and Facebook comments collection using online tools like extract comments.com, etc.
- Step 2: Data preprocessing like removal of duplicate posts, stop words.
- Step 3: Data normalization using porter stemming algorithm.
- Step 4: Data loading on Dataset.
- Step 5: Design and Implementation of program for per day frequency of posts.
- Step 6: Design and Implementation of program for feature selection.
- Step 7: Facebook post Classification.
- Step 8: Mobile app recommendations.

#### Algorithm:

Naïve Bayes:

Given training dataset D which consists of documents belonging to different class say Class A and Class B

Calculate the prior probability of class A=number of objects of class A/total number of objects

Calculate the prior probability of class B=number of objects of class B/total number of objects

Find NI, the total no of frequency of each class

Na=the total no of frequency of class A



$N_b$  = the total no of frequency of class B

Find conditional probability of keyword occurrence given a class:

$P(\text{value 1/Class A}) = \text{count}/n_i(A)$

$P(\text{value 1/Class B}) = \text{count}/n_i(B)$

$P(\text{value 2/Class A}) = \text{count}/n_i(A)$

$P(\text{value 2/Class B}) = \text{count}/n_i(B)$

.....

.....

.....

$P(\text{value } n/\text{Class B}) = \text{count}/n_i(B)$

Avoid zero frequency problems by applying uniform distribution

Classify Document C based on the probability  $p(C/W)$

Find  $P(A/W) = P(A) * P(\text{value 1/Class A}) * P(\text{value 2/Class A}) \dots P(\text{value } n/\text{Class A})$

Find  $P(B/W) = P(B) * P(\text{value 1/Class B}) * P(\text{value 2/Class B}) \dots P(\text{value } n/\text{Class B})$

Assign document to class that has higher probability.

**V. RESULTS AND DISCUSSION**

**Dataset**

Facebook Dataset Facebook has billions of users and it is one of the most used website/app used in the world. A lot of data are generated on Facebook daily. Here, we are going to look at an anonymized dataset named pseudo-facebook.csv. We are interested in looking at the various trends and distribution of the parameters in our data.

**Mathematical Model**

Create the user’s profile in three separate but linked stadiums: post keywords, post named individuals and post distributions on topics. The user should convey any profile as

$$U = \langle F, G, E \rangle$$

Which corresponds to the three-stage model of posts? However, there are some differences between this and the profile of posts,

$$F = \langle f_1, w_1, t_1, f_2, w_2, t_2, \dots, f_n, w_n, t_n \rangle$$

$f_i$  represents keywords that the user has accessed from historical posts, and any entry consists of a representative phrase, weight, and the last time that the user has accessed it;

$$E = \langle e_1, w_1, t_1, e_2, w_2, t_2, \dots, e_n, w_n, t_n \rangle$$

$e_i$  Represents the named entities that have been compiled from the user’s historical posts and each entry consists of an approved authorization object, the weight and the last time the user has viewed them;

$$G = \langle g_1, w_1, t_1, g_2, w_2, t_2, \dots, g_n, w_n, t_n \rangle$$

$G$  reflects the distribution of subjects gathered from the users’ historical entries, each entry consists of a representative theme id, a corresponding weight and the last time it was viewed by users.

**Results**

The Experiments are done by personal computer with configuration: Intel (R) Core (TM) i3-2120 CPU @ 3.30GHz, 4GB memory, Windows 7, MySQL 5.1 backend database and JDK 1.8. The application uses web application tool for design code in Eclipse and execute on Tomcat server.

Positive (P) : Observation is positive.

Negative (N) : Observation is not positive.

True Positive (TP) : Observation is positive, and is predicted to be positive.

False Negative (FN) : Observation is positive, but is predicted negative.

True Negative (TN) : Observation is negative, and is predicted to be negative.

False Positive (FP) : Observation is negative, but is predicted positive.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1\text{-Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$



**Comparison Graph:**

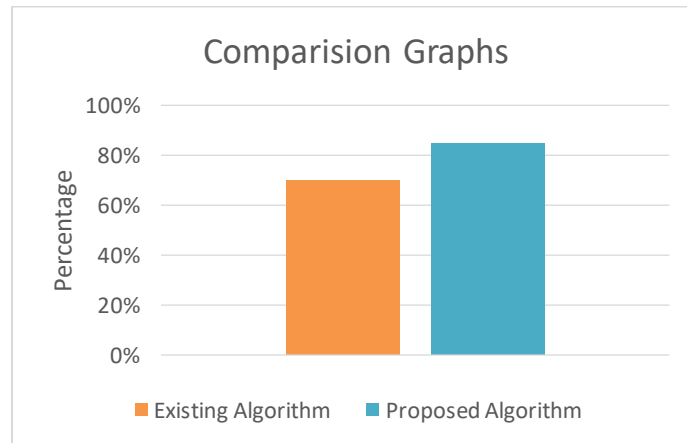


Fig 2. Graph

**Comparison Table:**

Sr.No	Existing Algorithm (TSVM)	Proposed Algorithm(NB)
1	65%	86%

Table 1.comparative result

**VI. CONCLUSION**

In this paper, we leverage the information of user social media to design a personalized app recommender system. The topic modeling approach is applied to extract topics from the huge amount of textual data in user social data to model the features of apps, and the user’s installed apps are capitalized to build the user profile to model user preferences. Both the user preferences and app features are taken into account to estimate the personalized app recommendation scores. Real-world data are utilized to perform experiments, and the experimental results show that the user reviews are effective for personalized app recommendations.

**REFERENCES**

1. Zhen Tu, Yong Li, Pan Hui, Li Su, Depeng Jin, "Personalized Mobile App Recommendation by Learning User’s Interest from Social Media" IEEE TRANSACTIONS ON MOBILE COMPUTING MAY 2020.
2. J. Feng, M. Zhang, H. Wang, Z. Yang, C. Zhang, Y. Li, and D. Jin, "Dplink: User identity linkage via deep neural network from heterogeneous mobility data," in The World Wide Web Conference. ACM, 2019, pp. 459–469
3. D. Yu, Y. Li, F. Xu, P. Zhang, and V. Kostakos, "Smartphone app usage prediction using points of interest," Proc. ACM UbiComp, vol. 1, no. 4, p. 174, 2018.
4. W. X. Zhao, S. Li, Y. He, E. Y. Chang, J.-R. Wen, and X. Li, "Connecting social media to e-commerce: Cold-start product recommendation using microblogging information," IEEE Transactions on Knowledge and Data Engineering, vol. 28, pp. 1147–1159, 2016.
5. M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, and S. Wang, "Learning graph-based poi embedding for location-based recommendation," in Proc. ACM CIKM. ACM, 2016, pp. 15–2



**Impact Factor:  
7.512**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details