International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8, Issue-1, May 2019

Machine Learning Finance: Application of Machine Learning in Collaborative Filtering Recommendation System for Financial Recommendations

Bhagirathi Nayak, Rajesh K. Ojha, P. S. Subbarao, VijayaBatth

Abstract: ***Prediction of stock market data have got a significant role in present scenario of finance. The various algorithms and models are used for forecasting of financial data. In this article we used application of recommender system. The recommender systems are one of the significant methodologies in machine learning technologies, which is using in current business scenario. This paper focuses on developing a stock market data recommender system using machine learning technique like k-Nearest Neighbors (k-NN) classification. Machine learning has become a widely operational tool in financial recommendation systems. Here the data considers the daily wise equity trading of Nifty 50 from National Stock Exchange (NSE) of 50 companies in 10 different sectors around 5986 days' transactions. We used k-Nearest Neighbors classification algorithm of deep learning technique to classify users based recommender system. We explore the traditional collaborative filtering with our methodology and also to compare with them. Our outcomes display the predictable algorithm is more precise than the existing algorithm, besides it is less time and robust than the existing methods.***

Keywords: Recommender System, Collaborative filtering, k-Nearest Neighbors, Classification, Equity Trading.

I. INTRODUCTION

The secondary markets are the trade sources of listed securities for the stock market or the equities market. The Nifty 50 equity trading in stock market continues to grow at an exponential rate. It becomes very difficult and time taking for all these traders to search for the right information. For the users of these trading it becomes increasingly difficult and time consuming to find the information they are looking for. Users can find the information with their interests by identifying a company trading with the help of user interface. Users may not always know their interests beforehand and they might change their interest with passing time. This may require them to change their selection regularly. **Recommender systems** provide identifying information by knowing the user's interests from the hints of their interaction with that trading. The purpose of this article to develop recommender systems model of finance which will reduce the difficulties of equity traders. Users always expect good recommendation due to the advances in recommender system. The bottleneck of the system is when the threshold of the trading is very low which makes in effective or poor quality of recommendation. For example, if a user would like to play a song on a music app and that app is not able to play that particular song then it is dumped by the user. It is the reason why app developing companies are putting a lots of effort to make their recommender systems more effective and accurate. What makes the challenge harder is the versatility in the user preferences and availability of choices. In addition to that several influential factors affect the mood of the user and brings more randomness in the behavior of the user. Here we are using collaborating filtering to predict the Nifty 50 equity of daily trade quantity. This article is designed NSE stock market data for collaborative filtering recommender system. The daily wise equity trading of 50 companies in 10 different sectors are considered and each sectors having top five companies.

II. LITERATURE REVIEW

We have gone through different research articles those reflect a comparative study on different machine learning techniques used for prediction of stock market. Atsalakis G. S., Valvani K. P. has done a survey on stock market prediction techniques [2]. Chuangguang huang and Jian Yin [4] described about the famous procedure of recommender system in collaborative filtering in his article. G. E. Kayakutlu, G. D TU [10] has introduced deep learning in fore casting the stock market and estimated its effectiveness on the stock price of Google. By Guang ping Zhuo, Jingyu Sun and Xueli Yu [5] "A Framework for Multi-Type Recommendations" treaties in the ground of web mining anxiety on certain shortcomings in collaborative filtering and also on multi type recommendation. Collaborative filtering (CF) is an operative method of recommender systems (RS) has been broadly used in online goods. Still, CF discomforts from some flaws such as difficulties with new users (cold start), data sparseness, and difficulty in noticing "malicious" or "untrustworthy" users and so on.

Revised Manuscript Received on 30 May 2019.
***** Correspondence Author
Dr. Bhagirathi Nayak*, FCMS, Sri Sri University, Cuttack, Odisha, India
Mr. Rajesh K. Ojha, FCMS, Sri Sri University, Cuttack, Odisha, India.
Dr. P. S. Subbarao, FCMS, Sri Sri University, Cuttack, Odisha, India
Dr. VijayaBatth, Xavier University Bhubaneswar, Odisha, India

By Qian Wang, Xianhu Yuan, Min Sun [12] "Collaborative Filtering Recommendation Algorithm based on Hybrid User Model". Collaborative filtering confronts difficulties of adaptability and furthermore suggestion precision so the paper proposes a cross breed user model to expel some of its downsides. The recommender framework in light of this model not just holds the benefit of suggestion accuracy in memory-based strategy, additionally has the scalability in the same class as model-based technique. By Chuangguang Huang and Jian Yin [4] "Effective Association Clusters Filtering to Cold–Start Recommendations ".

This paper concentrates on the best way to defeat cold start issue in the conventional research of Recommendations System (RS). The famous procedure of RS is Collaborative Filtering (CF) while in genuine online RS, CF can't basically take care of cold start issue for the sparsity ratings dataset. The paper proposed a fresh effectually association clusters filtering (ACF) algorithm. It also tells an effective personalized collaborative filtering scheme for Network service recommendation.

III. COLLABRATIVE FILTERING

The CF (collaborative filtering hence forth to be referred as CF) is a method of creating recommendation using social intelligence or collective intelligence. The recommendation is based upon the clustering of people with common interests. Using applied mathematics, the analogous knowledge is discovered by analyzing the patterns among people of common interests and by evaluating the ratings provided by totally heterogeneous users or implicitly through observation of the activity of the different users within the system. This method is drastically different from the opposite and most ordinarily used content based filtering approach. In spite of only recommending things as a result of they are just like things a user has liked within the past, things are counseled supported different user's preferences. This information is matched with different users to seek out overlaps of common interests among users. CF is many times referred to as social recommendation, which filters information by using the ratings given by other people considered as social intelligence. For example, people want to watch a movie, might seek for critics from his friends or his social networks. Few friends those have interest in similar type of movies or has already watched the movies may share their suggestion and these critics may be a deciding factor whether to watch the movie or not.

A. Collaborative Filtering Methods

Recommended items liked by similar users, enable exploration of diverse content

Recommendation task

	I-1	I-2	I-3	I-4
U1	5	1	3	5
U2	?	?	?	2
U3	4	?	3	?

Set of user's U and a set of items I to be recommended to the users. Learn a function based on the past data that predicts utility of each item i ($\in$I) to each user u ($\in U$).

User-based k-Nearest Neighbors

	I-1	I-2	I-3	I-4
U1	5	1	3	5
U2	?	?	?	2
U3	4	?	3	?

Compute similarity of users, find k most similar users to user a. Recommended book not read by user a. Cosine similarity

$sim(a,b) = \frac{a \cdot b}{||a|| \cdot ||b||}$

This algorithm computes cosine or correlation similarity of rows (users) or columns (items) and recommends items that

k-Nearest Neighbors.

IV. K-NEAREST NEIGHBORS

K-NN is a very efficient classification algorithm which is extremely used in machine learning. It is a supervised learning algorithm hence it is more reliable also. K-NN is also exhaustively being used in pattern classification, clustering, prediction and intrusion detection systems. No hidden assumptions are to be made about the sample distribution (like the contemporary algorithms such as GMM, that presumes a Gaussian distribution on the sample) which defines it as a non-parametric algorithm and makes it perfectly applicable in solving a wide range of real life problems.

V. ALGORITHM

K-NN being a supervised learning algorithm stores the available information and tries to classify the new input by calculating their similarity with the existing cases by using distance functions as tools to measure similarity among patterns. Due its accuracy of classification and prediction it is still being used as pattern recognizer and statistical estimator in many applications.

Distance functions

Euclidean $\sqrt{\sum_{i=1}^{k}(x_i - y_i)^2}$

Manhattan $\sum_{i=1}^{k}|x_i - y_i|$

Minkowski $\left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q}$

It can be observed that the above three functions used to measure distance are only applicable with the variables those are of continuous in nature. These three functions do not fit well with variable of categorical type. When the dataset consists of both categorical and numerical variables then the standardization issue for the numerical variable within the range of 0 to 1 may come up.

Hamming Distance

$$D_H = \sum_{i=1}^{k} |x_i - y_i|$$

$$x = y \Rightarrow D = 0$$

$$x \neq y \Rightarrow D = 1$$

X	Y	Distance
Male	Male	0
Male	Female	1

MACHINE LEARNING IN FINANCE

Application of machine learning in finance proves the superiority of the technique, because finance is the most computationally intensive field where a large number of parameters are there to influence the decision making process. Machine learning techniques have become very popular in prediction and recommendation over the traditional methods. Both supervised and unsupervised models of machine learning are being used for financial prediction. These models are state based models, the econometric models or even the stochastic models are marred by the problems of over fitting, heuristics and poor out of sample results. Which is because, the financial domain is hugely complex and non-linear with a plethora of factors influencing each other. To solve this, if we look at the research done in Machine Learning in proven fields of investment recommendation, selection of mutual funds investments, prediction of future returns of the stock. In most of the cases Machine Learning had produced better prediction and classification than its counter statistical estimators due to its self learning and adaptability capability.

VI. DATA COLLECTION & ANALYSIS

Here we used the historical data set of Nifty 50 daily trade equity collect from National Stock Exchange (NSE) website. We have collected one year's daily wise data from selected companies of 10 different sector, like FMCG, Banking, IT, Pharmaceutical, Automobile, Construction, Finance, Power Industries, Steel and Telecom. Each sector having top five companies.

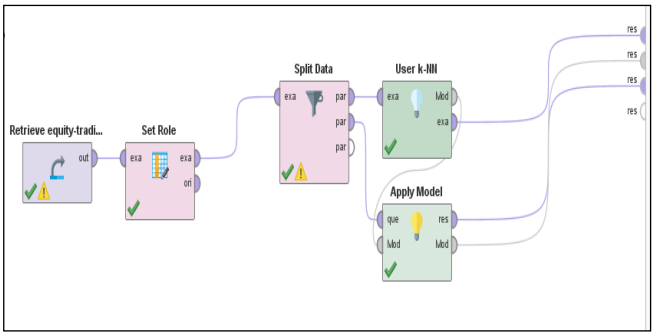
For Recommender Systems we need Ratings, which are on a rating of 1 to 10 (whole-star ratings only). We did filter the said data set and took percentage of daily trade quantity of Nifty 50 data to and converted it to rating by given weightage. We used days as user id and companies as item id. After that we did collaborative filtering using rows (users), columns (items) and recommends items as per rating using k-Nearest Neighbors algorithm. There is one file: 'equity_trading.csv' having more than 59 thousand data. Here we are using collaborative filtering for prediction of user ratings, which shows in figure 1.

Row No.	user_id	item_id	rating
1	DAY-1	ASHOKA LEYLAND	5
2	DAY-1	BAJAJ AUTO	7
3	DAY-1	HERO MOTORS	4
4	DAY-1	MARUTI	5
5	DAY-1	TATA MOTORS	4
6	DAY-1	AXIS BANK	4
7	DAY-1	ICICI BANK	4
8	DAY-1	KOTAK BANK	6
9	DAY-1	YES BANK	4
10	DAY-1	GAMMON INFRASTRUCTURE	6
11	DAY-1	HINDUSTAN	2
12	DAY-1	L & T	5
13	DAY-1	PUNJLLOYD	3
14	DAY-1	EDELWEISS	7
15	DAY-1	MAHINDRA FINANCE	4
16	DAY-1	MFSL	5
17	DAY-1	HINDUSTAN UNILIVER	7

Figure 1: The initial data set with rating

For collaborative filtering we use said file and create model using k-NN algorithm. We took data file and select for model and we did set role item_id for item information, user_id for user information and rating for label then splitting the total data in to two part in different ratio. These split data are created a model using k-NN algorithm, which shows in figure 2.

Figure 2: The Workflow of Collaborative Filtering Model using k-NN



The above item recommendation workflow, the Set Role operator is being used to assign different roles to the attributes appropriate for them. The user_id has been assigned the role of user identification and similarly the item_id has been assigned the role of item identification. Roles for the data attributes must be assigned even though they can ne named arbitrarily. Then the k-NN algorithm has to be trained by setting appropriate roles to the attributes using the training data set available. Here we have used Apply Model operator on the query set for recommendation of new items by applying our trained model. Before applying the model, we have assigned the user identification role to the query set. The Apply Model operator returns an example set containing the first n ranked recommendations for every user in a query set. After created the said model we checked the performance of the model, which shows in figure 3.

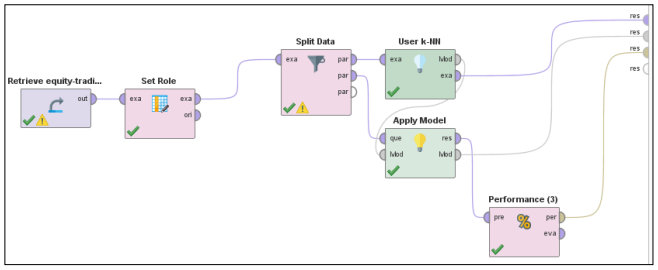


Figure 3: The Collaborative Filtering Model

The Performance operator has been used to calculate the error in recommendation. A set containing performance measure is returned as a performance vector by this operator. Both the Performance vector and the output dataset are shown in the figure 4 and 5 respectively.

VII. ROOT MEAN SQUARE ERROR (RMSE)

The RMSE of a prediction model with respect to the estimated variable X_{model} is defined as the square root of the mean squared error, where $X_{obs}$is observed values and X_{model} is modelled values at time/place *i*.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

PerformanceVector

```
PerformanceVector:
RMSE: 1.148
MAE: 0.896
NMAE: 0.224
```

Row No.	user_id	item_id	rating	predicti...
1	DAY-1	HDFC BANK	6	6.954
2	DAY-1	JAYPEE INFRATECH	5	7.393
3	DAY-1	BAJAJ FINANCE	4	4.003
4	DAY-1	PFC	5	3.889
5	Day-1	BRITANIA	6	5.504
6	DAY-1	MARICO	7	6.015
7	DAY-1	HCL	5	5.866
8	DAY-1	INFOSYS	8	6.419
9	DAY-1	SUN PHARMA	4	3.802
10	DAY-1	ZYDUSWELL	8	6.512
11	DAY-1	NTPC	7	6.715
12	DAY-1	TATA POWER	5	4.123
13	DAY-1	GAIL	4	5.149
14	DAY-1	SAIL	2	2.226
15	DAY-1	TATA STEEL	3	2.370
16	DAY-10	BAJAJ AUTO	6	5.126
17	DAY-10	HERO MOTORS	5	6.080

Figure 4: Result of Collaborative Filter

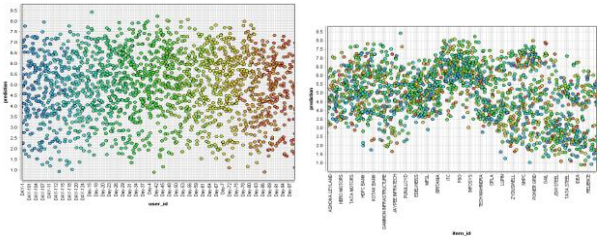


Figure 5: Scatter plot of Item and User prediction

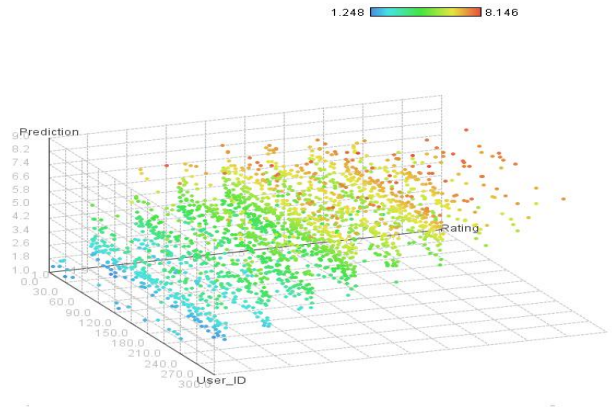


Figure 6: 3D plot User wise prediction

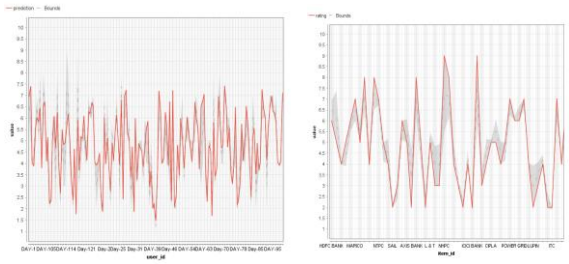


Figure 7: Line graph of Item and User Wise Prediction

VIII. CONCLUSION

Recommender systems occupied a key role in making predictions in the information and decision-overwhelmed world. It has not only changed the perspective of decision making by introducing group intelligence but also helped the implementers to maximize the financial gain by applying such social recommendation into e-commerce and finance. The goal of this paper is to introduce recommenders in finance of company's equity using a machine learning approach. We also presented here how recommender system can use in finance. Still now most of the articles are published either in movie or in ecommerce recommender systems. This article presented companies equity recommendation by applying our financial recommendation model, which may be utilized for recommending the different company's trade equity to the end user for understand the financial stock market. The result shows that the model can be used for large dataset by extending number of users.

REFERENCES

1. H. A. Abdou, W. M. Abdallah, J. Mulkeen, C. G. Ntim, & Y. Wang, "Prediction of financial strength ratings using machine learning and conventional techniques," *Investment Management and Financial Innovation*, 14(4), 2017, pp. 194-211.
2. G. S. Atsalakis, K. P. Valavanis, "Surveying stock market forecasting techniques, Soft computing method," *Expert Syst. Appl.*, 2009, 36:5932-5941.
3. A. Carzaniga, M. Rutherford, A. Wolf, "A Routing Scheme for Content-Based Networking," In *IEEE* INFOCOM 2004, vol. 2, pp. 918-928.

4. C. Huang, J. Yin. "Effective association clusters filtering to cold-start recommendations." In *Seventh International Conference on Fuzzy Systems and Knowledge Discovery IEE*, vol. 5, 2010, pp. 2461-2464.
5. G. Zhuo, J. Sun, and X. Yu. "A framework for multi-type recommendations." In *Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE*, vol. 3, 2011, pp. 1884-1887.
6. E. Guresen, G. Kayakutlu, and T. U. Daim. "Using artificial neural network models in stock market index prediction." *Expert Systems with Applications* 38, no. 8: 2011 10389-10397.8
7. K. L. Maglaughlin, and D. H. Sonnenwald. "User perspectives on relevance criteria: A comparison among relevant, partially relevant, and not-relevant judgments." *Journal of the American Society for Information Science and Technology* 53, no. 5, 2002: 327-342.
8. J. K. S. Liew, and B. Mayster. "Forecasting etfs with machine learning algorithms." *The Journal of Alternative Investments* 20, no. 3 (2017): 58-78.
9. D. Olson, and C. Mossman. "Neural network forecasts of Canadian stock returns using accounting ratios." *International Journal of Forecasting* 19, no. 3 2003: 453-465.
10. R. K. Ojha, and B. Nayak. "Application of Machine Learning in Collaborative Filtering Recommender Systems." *International Journal of Engineering & Technology*7, no. 4.38 2018: 213-215.
11. R. D. McLean, and J. Pontiff. "Does academic research destroy stock return predictability?" *The Journal of Finance* 71, no. 1 2016: 5-32.
12. Q. Wang, Xianhu Yuan, and Min Sun. "Collaborative filtering recommendation algorithm based on hybrid user model." In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 4, pp. 1985-1990. IEEE, 2010.
13. I. Rose, R. Murty, P. R. Pietzuch, J. Ledlie, M. Roussopoulos, and M. Welsh. "Cobra: Content-based Filtering and Aggregation of Blogs and RSS Feeds." In *NSDI*. 2007.
14. J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez. "Recommender systems survey." *Knowledge-based systems* 46 2013: 109-132.
15. F. Ricci, L. Rokach, and B. Shapira. "*Introduction to recommender systems handbook*." In *Recommender systems handbook*, pp. 1-35. Springer, Boston, MA, 2011.
16. P. Resnick, and H. R. Varian. "Recommender systems." *Communications of the ACM* 40, no. 3 1997: 56-59.
17. M. Deshpande, and G. Karypis. "Item-based top-n recommendation algorithms." *ACM Transactions on Information Systems (TOIS)* 22, no. 1 2004: 143-177.
18. J. S. Breese, D. Heckerman, and C. Kadie. "Empirical analysis of predictive algorithms for collaborative filtering." In *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., 1998, pp. 43-52.
19. N. Miller, J. A. Konstan, and J. Riedl. "PocketLens: Toward a personal recommender system." *ACM Transactions on Information Systems (TOIS)* 22, no. 3 2004: 437-476.
20. Huang, and J. Yin. "Effective association clusters filtering to cold-start recommendations." In *Seventh International Conference on Fuzzy Systems and Knowledge Discovery*, IEEE, 2010 vol. 5, pp. 2461-2464.
21. B. Nayak, and C.A. V. Batth. "Association Rule Mining using Apriori algorithm for work-related beliefs of Generation X and Generation Y." *International Journal of scientific research and management (IJSRM)* 3, no. 7 2015: 3324-3330.
22. Y. Jiang, J. Liu, M. Tang, and X. Liu. "An effective web service recommendation method based on personalized collaborative filtering." In *IEEE International Conference on Web Services*, 2011, pp. 211-218.

AUTHORS PROFILE

Dr. Bhagirathi Nayak, is result oriented professional with an experience of 25 years from IIT Kharagpur in the areas of Computer Science & Engineering and Business Management, Database Designer and Developer, Teaching, Academic Administration and Pedagogical Activities. He did his Ph.D. in Computer Science. Dr. Nayak presently associated with Sri Sri University, Cuttack, Odisha as an Associate Professor and Head, Information & Communication Technology.

Rajesh Kumar Ojha, PhD scholar, Sri Sri University, M. Tech, B. E in Computer Science & Engineering. Working as Associate professor and HOD, Department of Master in Computer Application, Gandhi Institute for Technology (GIFT), Bhubaneswar, Odisha.
Mob: 8763822485 Email: mailtorko@gmail.com

Prof. (Dr.) Srinivas Subbarao Pasumarti. He has over 23 years of experience in teaching, research and academic administration. He is an alumnus of IIM Ahmedabad. He did his Ph.D. in Entrepreneurship from Nagpur University and Post-Doctoral from Berhampur University on "FDI in Indian Tourism Sector. He received the Best Teacher Award from JNTU, Kakinada for the years 2009-2013. Very recently he received the AIMS-IRMA Outstanding Management Teacher award from the Director, IRMA.

Dr. Vijaya Batth is a faculty member in the Accounting Area and holds triple qualifications of being a Chartered Accountant, Company Secretary and Bachelor in Law. She has about sixteen years of experience in Industry, Practice and Academics. Her Doctoral Research was on Corporate Governance in Indian Public Sector Undertakings with regard to the position of Independent Directors. She has actively participated in various seminars and conferences held across the country and contributed articles in reputed publications. Dr. Batth presently associated with Xavier University, Bhubaneswar as an Assistant Professor.