RESEARCH PROPOSAL

Research Topic: Developing Algorithms for High-Performance Real-Time News Recommendation System

Introduction

With the unexpectedly developing quantity of objects and information articles at the internet, recommender structures are one of the key technology to address the statistics overload and to help customers in locating statistics matching the their man or woman preferences. News and area-unique statistics portals are important understanding reasserts at the Web regularly accessed via way of means of hundreds of thousands of customers. In comparison to product recommender structures, information recommender structures have to address extra challenges, e.g. brief information article lifecycles, heterogonous user interests, strict time constraints, and context-established article relevance. Since information articles have best a brief time to live, recommender fashions ought to be constantly adapted, making sure that the suggestions are usually up-to-date, hampering the pre-computations of suggestions. In this paper we gift our framework for supplying real-time information suggestions. We talk the carried out algorithms optimized for the information area and gift an technique for estimating the recommender performance. Based on our evaluation we implement an agent-primarily based totally recommender system, aggregation numerous extraordinary recommender strategies. We study a context-conscious delegation strategy, permitting us to select the exceptional recommender set of rules for every request. The assessment indicates that the carried out framework outperforms conventional recommender methods and permits us to conform to the unique

residences of the taken into consideration information portals and advice requests [1].

Brief Background:

Recommendation task is a very common use case in every web and mobile application to make suggestions for interesting products or services to its users from a large space of possible options in order to enhance business and assist decision-making process. For example, 80% watched movies in Netflix and 60% video clicks in YouTube appear through recommendations [2,3]. Generally, historical data like user preferences, item features, user-item past interactions, temporal data and spatial data are used to produce the recommendation lists. Recommendation models can be classified into collaborate-filtering, content-based and hybrid which are based on the type of input data. These classical recommendation models are not efficient to generate recommendation lists due to dealing with the data sparsity and cold-start problems and balancing the recommendation qualities in terms of different evaluation matrices [4]. Moreover tags in social media like Facebook can be used to make personalized recommendations. Tags with noises, such as tag synonyms, semantic ambiguities and personal tags degrade the recommendation quality.

Literature Review

Liang et al [5] proposed an approach to eliminate the noise by determining the relevant tags of each item and the tag preferences of each user. However, since the historical input data are increasing exponentially, named Big Data, due to the growing of using internet over PCs, Tablets and Smart phones in real time, existing recommendation algorithms are tedious and time consuming to

make recommendations. Thus the Deep Neural Network (DNN), a multilayer percept network with many hidden layers whose weights are fully connected [6], has been incorporated in recommendation system to enrich the recommendation quality. A deep learning based video recommendation technique for YouTube was proposed by Covington et al [7]. Cheng et al. [8] approached an App recommender system for Google Play using wide and deep model. The way of building a news recommender system using Recurrent Neural Network (RNN) for Yahoo News was presented by Shumpei et al [9].

Objectives of the Research:

The data, inseparable part of every sector and function in the global economy, collected from internet are increasing rapidly, coming very quickly and heterogeneous in nature which refer to the Big Data but the existing data storage and processing tools are not able to handle that sort of data. The study on "Big Data" is one of recent trends in various industries, be it health, education, entertainment, science and technology, genetics or business operations. A review of the related literature reveals three pertinent problems in the area of big data on deep recommendation system -some of them have been addressed and some are only initiated. First problem is real-time recommendation system. Since data are not static in real-time application, that is data are coming continuously, the recommendation system ought to be incremented dynamically. Secondly, the computational complexity of high dimensional tensors and multimedia data sources to make recommendation is high. Final challenge is to balance the complexity and scalability of recommendation model with the exponential growth of features. Due to the nature of big data, scalability is a critical issue in every real-world application.

This research focuses on deep learning based algorithm which is the scalable news recommendation technique in big data analytics, where the principal consideration to select the best model is time complexity.

Research Methodology:

In this research, a very meticulous procedure has to be followed in order to achieve the intended outcomes. The procedure will include several steps befitting to the objectives of this research.

PHASE I: Literature Review

In this phase, the survey of the literature will be done to study the limitations of existing techniques of Recommendation System for big data.

PHASE II: Requirement Elicitation

Based on the literature review, the rational for the research is established in this phase. Then which big data tools will be used to simulate discovered recommendation techniques is decided in this phase.

PHASE III: Design, Implementation and Verification

Here, the proposed techniques will be modeled, implemented and verified whether these techniques are fruitful or not. The model will be validated analytically and the performance will be measured through simulation. In addition, the correctness of the algorithm will also be investigated and as a matter of fact, the proof of correctness will be performed based on mathematical analysis.

PHASE IV: Conclusions and Report Preparation

In phase IV, conclusion will be drawn and report will be prepared in the light of findings of the research.

Conclusion and Expected Outcome:

The World Wide Web is a stage that give admittance to online distributed news stories to the news per users from a large number of news sources far and wide. They are urged to communicate as perspectives, audits, sentiments, and evaluations, remarks, sharing this data by labeling, sharing and contributing new data [10]. The openness and accessibility of these offices make an enormous pool of data through news. Suggestions of pertinent data or news by separating or outline of web news and present brief news substance to the focused on news per users, draws in scientists to discover new techniques for connecting and suggesting related news stories absolutely. Proposal frameworks help to recommend the most significant things or data from a gigantic assortment of things or information in various application spaces at the web with insignificant endeavors of the client. News suggestion frameworks are explicit application space, where the goal is to prescribe the most pertinent news stories to the news per user dependent on some proposal draws near. News proposals offer explicit difficulties (specialized perspective and general methodology related difficulties) when contrasted with other application spaces [11]. For instance, news suggestion framework involves difficulties like tight reaction time, dynamic condition and ceaselessly changing assortment of news stories or clients, weighty burden tops at times of heavy traffic, and so forth [12].

In this proposed model, we will use reverse hierarchical clustering of an equivalent cluster to classify more distinct news. Unlike other search engines, we will apply three different methods of reverse Hierarchical clustering to seek out different news or information associated with an equivalent subject. What we find in pipilika or Google news portal that they assemble daily news and formulate cluster by collecting analogous news from the various sources of the newspaper. In this regard, they take the news cluster under consideration on daily news basis and don't cover news of previous days on the same issue. It from a user view, is more desired provided that they get to understand more diversified information of the same news since it creates monotony. And what we will do in this proposed model is that we put an efficient approach into operation for recommending news of an equivalent topic along with more diversified information where viewers can come across most of the unlike news on an equivalent agenda during a single cluster which also ensures previous couple of days' news. Three different recommendation approaches are based on reverse hierarchical clustering we employ to show the three different results although of these three methods hopefully will be successful and find our desire output, finding different news on one news agenda during a single cluster. However, it should be mentioned that our third method will be more effective, in terms of showing result, than the primary and second one [13].

References:

[1] Lommatzsch A. (2014) Real-Time News Recommendation Using Context-Aware Ensembles. In: de Rijke M. et al. (eds) Advances in Information Retrieval. ECIR 2014. Lecture Notes in Computer Science, vol 8416. Springer, Cham.

[2] James Davidson, Benjamin Liebald, Junning Liu, Palash Nandy, Taylor Van Vleet, Ullas Gargi, Sujoy Gupta, Yu He, Mike Lambert, Blake Livingston, and Dasarathi Sampath. 2010. The YouTube Video Recommendation System. In Proceedings of the Fourth ACM Conference on Recommender Systems (RecSys '10). ACM, New York, NY, USA, 293–296.

[3] Carlos A Gomez-Uribe and Neil Hunt. 2016. .e net.ix recommender system: Algorithms, business value, and innovation. ACM Transactions on Management Information Systems (TMIS) 6, 4 (2016), 13.

[4] Shuai Zhang, Lina Yao, Aixin Sun, "Deep Learning Based Recommender System: A Survey and New Perspectives", arXiv:1707.07435 [cs.IR], Version 5, 3 August 2017.

[5] Liang H, Xu Y, Li Y, Nayak R, Tao D, (2010) Connecting users and items with weighted tags for personalized item recommendations, Proceedings of the 21st ACM Conference on Hypertext and Hypermedia p51-60.

[6] Basiliyos Tilahun Betru, Charles Awono Onana and Bernabe Batchakui. Deep Learning Methods on Recommender System: A Survey of State-of-the-art. International Journal of Computer Applications 162(10):17-22, March 2017.

[7] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In Proceedings of the 10thACM Conference on Recommender Systems. ACM, 191–198.

[8] Heng-Tze Cheng, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson, Greg Corrado, Wei Chai, Mustafa Ispir, et al. 2016. Wide & deep learning for recommender systems. In Proceedings of the 1st Workshop on Deep Learning for Recommender Systems. ACM, 7–10.

[9] Shumpei Okura, Yukihiro Tagami, Shingo Ono, and Akira Tajima. 2017. Embedding-based News Recommendation for Millions of Users. In Proceedings of the 23th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM

[10] D. M. Durairaj and K. M. Kumar, ''News recommendation systems using Web mining: A study,'' Int. J. Eng. Trends Technol., vol. 12, no. 6, pp. 293–299, Sep. 2014.

[11] Ö. Özgöbek, J. A. Gulla, and R. C. Erdur, ''A survey on challenges and methods in news recommendation,'' in Proc. WEBIST (2), 2014, pp. 278–285.

[12] I. Verbitskiy, P. Probst, and A. Lommatzsch, ''Development and evaluation of a highly scalable news recommender system,'' in Proc. CLEF, 2015, pp. 1–5.

[13] Recommendation System for Bangla News Article with Anaphora Resolution –Kazi Wohiduzzaman and Sabir Ismail,4th International Conference on Electrical Engineering and Information & Communication Technology,2018.