INTERNATIONAL JOURNAL OF
RECENT TRENDS IN MULTIDISCIPLINARY RESEARCH
July-August2021, Vol 01(01), 06-09

A Cuisine Based Recommender System Using k-NN and Map reduce Approach

KesavaDasu[1], K. VaddiKasulu[2], K. Raj Kamal[3], G. Pranith[4]
[1,2,3,4] *Dept. of. CSE, K L (Deemed to be University), Guntur, Andhra Pradesh, India.*

Article Type: Research

OPENACCESS

Article Citation: Kesavadasu[1], K. Vaddikasulu[2], K. Raj Kamal[3], G. Pranith[4], ACuisine BasedRecommenderSystemUsingK-NNandMapreduceApproach,International Journal Of Recent Trends In Multidisciplinary Research, April 2021, Vol 01(01), 06-09.

Received date: June 21, 2021

Accepteddate:July01, 2021

Published date :July 12, 2021

Abstract: In the present days, life can be made smarter, including the food we eat by taking a decision from the restaurant recommender systems. In this paper the authors proposed a restaurant recommender system based on the search of user cuisine.The top-k restaurants are identified along with the ratings of the bistros recommended. The recommendations are recuperated considering the tendency of the client cuisines which is an important category which inherently defines the other features and these features are considered to give a good service which is the novelty of this paper.Providing recommendations considering client cooking styles is the multifaceted nature of the problem. The well known k-Nearest Neighbor algorithm is implemented with the Map Reduce paradigm which can quickly process tremendous proportions of data. Its show is tested on bench marked dataset and the results are found to be successful.

Index Terms: Restaurant Recommender System, Nearest Neighbor approach, MapReduce, Cuisine based search.

1.Introduction

In this world of competition it is essential to be healthy and choose proper food to eat by making use of the improved resources that we have around us. The various search engines give a lot of diners that are suitableto us yet they don't focus in on food unequivocal search.In this paper we propose a restaurant recommender system where the user can get the restaurants based on his taste and he can also have a list of rest aurants that provide the food including the ratings of the bistros which give a better option for the client than choose.We classify the restaurants based on user express food by making use of the well known kNN algorithm. This also helps to a groupof users who have the same food habits to get good recommendations based on the similarity. This may be helpful to have good, happy and tasty get together where all the group of people have the save us our. With the use of internetand promotion in websites many restaurants data are available which has a amazing need of good restaurant recommendation system that filters and promotes good suggestion for tasty food.The proposed approach is an application of Machine learning based classification to classify the restaurant.

The food-diner client structure a kind of semantic network where multiple paths can be used to traverse different semantic associations like: perceiving food of same or different sort, customers of same age, taste, restaurants of same or different ratings which are nothing but various Metapaths in the Information Network.Information Networks can be classified as Homogeneous Information Networks where objects interact with near kind of associations and Heterogeneous Information Networks[1] where different kind of things communicate with various semantically interconnecting links. A heterogeneous network can contain a homogeneous network within it.

To expect the semantic relationship betweenthe objects in the Information Associations, it is gigantic to compute the similarity between the objects.ManySimilarityalgorithms are proposed to deal with the similarity between objects present in Information Networks.

With the augmentation of data computerization and the General Web we have heterogeneous Information Networks identified in every field like social media, medical databases data, paper-authornetwork etc. Also with heap of information flow in the association different traditional Data Mining strategies like gathering, classification,link prediction etc. are sensible and can be applied on the data present in the Information association. In the past few years Heterogeneous Information Networks has a widespread applications like finding associates on agreeable networks,link estimate, proposition of products, food and friends. Beside applying the Data Mining techniques, similarity measure can also be used tocompute the similarity between the articles present in the Information Network. Similarity measure is a verifiable measure used in Data Mining to handle the likeliness between the objects which talk with each other. The imperative variable in there commender systems is tocompute the similarity to compare two things that are closer to each other. In this paper the authors proposed rest aurantre commender system based on the client food and also provide there lated restaurants which are sensible to the cooking of the customer,the price of the food in a sorted order and also the distance from the force region of the customer.The paper is composed as follows: In Section 2 the related work is discussed, Portion 3 bright lights on the Methodology adopted,Section 4 show case experimental results, Section 5 concludes with future scope and research perspective.

2. Related Work

Recommender systems are more prominent with the increase being used of web and work on the machine learning principle of learning about the Neighbors for better solution. Recommender systems are classified as collaborative filtering models which are also called as Neighborhood models[2], these cond approach is content based filtering also called as data based models and the mix of the past and the later is the hybrid model.With the growing data it is a good option to make use of the Map Reduce paradigm for processing the data parallelly.In [3] the makers proposed the user based collaborative filtering recommender algorithm using Hardtop MapReduce. On the practically identical lines content based recommender structures is proposed in [4].The methodology proposed in [5] is a unique strategy of using Bhattacharya coefficient for collaborative filtering in sparse data. Chenyang Li and Kejing He [6] proposed an item based collaborative filtering using MapReduce.Though there are various MapReduce based recommenders proposed, applying kNN with Map Reduce is an ew idea for identifying the nearest Neighbors of the apparent attributes.In [7] Moon-hee Parketal proposed a restaurant recommender systems considering decision made by pack of people in adaptable environment using Bayesian Associations to model the preference of the users.

Comparability measure is used to evaluate the likeness between the objects and the ir interaction.Many similarity algorithms are proposed to analyze the Information Networks.The resemblance algorithms are useful for the prediction of links in the heterogeneous information networks and also to detect duplicate web documents.SimRank [8] is a similarity measure proposed based on random surfers 'model. It is essentially used for situating the web documents. PRank [9] is another closeness algorithm which includes headed together construction for calculating the similarity.PathSim[10]is a similarity measure which uses a commuting matrix.HeteSim [11] uses transition probability matrix.

3.Methodology

In this section we look at about changed phrasings like the Meta-Ways which are formed between the heterogeneous objects.Next weal so discuss the algorithm for calculating the similarity between the heterogeneous objects using Path Sim measure. A Meta-path is a semantic path that exists between the objects. For example for this food chain network which can be treated as a heterogeneous network we consider three different kinds of objects like user,restaurant and cuisine.Though these are of different kind, now they communicate with each other by the following relationships: user visits restaurant,restaurant offers cuisine,and cooking liked buyer

We have thought about the food enlightening record and endeavored to retrievethe resemblance scores for each client, the spots they visit andthe food they eat. There are numerous ways of managing calculatethe comparability yet we have picked the PathSim similarityscore as it shows properties like [i] symmetric property[ii]self most prominent property [iii] Balance of detectable quality. ThePath Simisa meta-path based approach to register the similarity. According to PathSim, a proximity matrix must be constructed between these objects to find how one object is like another.Instead of an adjacency matrix we build a driving grid which is proposed in [12].Butany Information Association is titanic of its sort as there are number of objects that interact with each other.

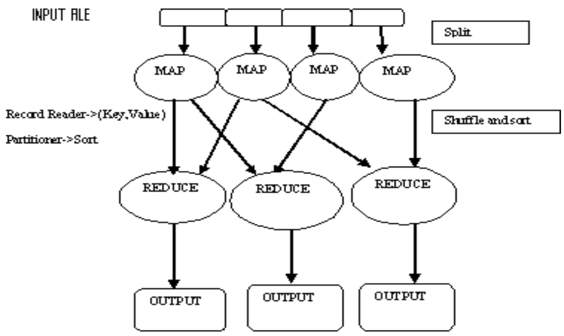

Fig1 Map Reduce Phases

To proceed with, the overall ratings file is given as an input to get the intriguing region id as the key which is visited by the set of clients as value.By considering each row a new Meta-path can be worked by using the matrix. From the first line of the MapReduce yield, it is seen that there are four clients (U1067,U1082,U1087,U1050)who visit the place id 132560.By traversing all the other places visited by these four users we get an commuting matrix between the users and places is depicted in figure-4.

Figure 5 depicts the commuting matrix between the user-id and the diner id.A '1' addresses the user visits the spot and an empty opening tends to that he did not visit the place.From this driving organization we can compute the similarity between two users using the formula
S(x:y)=2(Px->y:Px->yƐP)/|(Px->x:Px->xƐP)+|(Py->y:Py->yƐP)
|..Eqn(1)
This can be handled as follows for the above example:Similarity(U1067,U1082)=A/B=0.666
A=2[(1X1)+(1X0)+(1X1)+(1X1)+(1X0)+(0X1)+(0X1)]B=[(1+1+1+1+1)+(1+1+1+1+1)]
Similarity(U1067,U1087)=0.25Similarity(U1067,U1050)=0.166Similarity(U1082,U1087)=0.25Similarity(U1087,U1050)=0.4Similarity(U1082,U1050)=0.0571
Comparably we can more over handle the comparability between the restaurant and the cuisine it provides. After doing this weget a Meta-Path between the client restaurant cuisine.Thefigure 6 depicts the driving organization outlined between there staurantid and the cuisine offered in the restaurant.

Type-of-food	Mexican	Regional	pizzeria	Fast-food
place-id				
132630	1			
132560		1		
132732	1			
132733			1	
132663	1			
132594	1			
132740	1			
132608	1			
132609				1

Fig2:CommutingMatrixbetweencuisineandrestaurant-id.

Once we obtain the commuting matrix we can calculate the similarity between the heterogeneous objects using the Algorithm-1as follows:

Algorithmtofindthesimilarityof Meta-pathsbetweentheheterogeneous objects
Step 1:InputX Step 2: Construct the commuting matrix.Step3:forall ApplythePathSimsimilaritymeasureusingEqn1. $S(x:y)=2(P_{x\text{->}y}:P_{x\text{->}y}ƐP)/|(P_{x\text{->}x}:P_{x\text{->}x}ƐP)+|(P_{y\text{->}y}:P_{y\text{->}y}ƐP)|$

Algorithm-1:Tocomputethesimilaritybetweentheheterogeneousobjects.

The Figure-7 represents the Input file Xon which the first Map Reduce task is run to get the list of unique restaurant ids,the file will also provide you the list of user ids and the ratings given by them.

The estimation to make the key time frame out of MapReduce before data pre-processing is as follows:
Map Reduce paradigm is a parallel approach for processing monstrous data in the appropriated stage. We use Hardtop Distributed File System for the storage of the dataset and implement the Map Reduce which takes the<key,value>pair as input and gives the <key,value>pair as output.

Every record id in the document is given as a key and each line is given as data regard. After the MapReduce task is completed we get the placeid as keyand the list of users and their corresponding ratings as value.

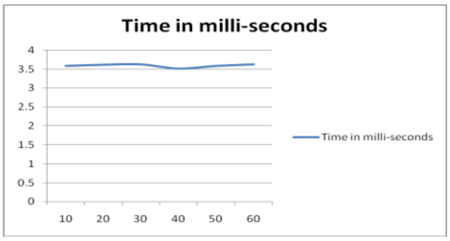

Fig- 2:Graphs howing no.of records and runtime

3.1Description of data Set

Estimation 2:To prepare the input file from raw dataset.

The chief place of the makers is to describe the restaurants using the prominent estimation kNN with the Map Reduce flavour. Using the kNN we perceive the nearest Neighbor to classify the bistros considering which cooking a huge part of its users are likely to eat in that restaurant.

k-NN algorithm is a distance based classification algorithm which gives out a class name of a dark case to its nearest Neighbor. Distance estimations like Euclidean, cosine and Man hattan distances can be used for each Neighbor if it is a relentless trade mark. The Euclidean distance is used inthe example of relentless factors any way with apparent variables we use the Hamming Distance. The food activity distance can be stated as dist(hamming)=|obji-objj|.- Eqn2.

For example if the client U1067 favors Mexican dish andU1082 in like manner incline towards Mexicanthen the Hamming Distance between U1067 and U1082 is zero. Else if U1082 prefers Chinese to Mexican it is one. The Estimation 3 as mentioned below does the classification.

4. Experimental Results

The assessment is coordinated on a lone center with 8 GBRAM,500GBHDD,intelcore i3 processor with a speed of 2.86 GHz. The typical runtime of the general large number of records in the fileturnedtobe3.59Milli-seconds. Hardtop is installed in pseudo distributed mode. The input is stored on to HDFS aa input file.The input file consists of the restaurant id as the first field.

The data set is taken from the UCI Machine Learning repository from the Data Folder with the URLhttps://archive.ics.uci.edu/ml/datasets/Bistro//consume data. The raw dataset contains 1161 instances. The first column gives the user_id these cond column provides the restaurant_id and three kinds of ratings are provided column three decides the general rating area five and six provide the food and service rating respectively. The cuisines offered by the restaurants are 59 but of which we classify the restaurant on the maximums avour of the users who visit the restaurant.

Fig-3:Famou scuisines offered by the number of restaurants

The figure-9 represents the graph in which we can understand that the Mexi can is the most famous cuisine offered by 239 bistros and the plan results also agree with this fact.

From the figure-10 the procedure adopted toget the restaurants organized is depicted. The Data pre-taking care of is done by perceiving which user_id will organize which cuisine and replaced with that particular food. The void values are killed and the data cleaning is done. The refined and processed file is given as input to these cond phase of MapReduce.

5.Conclusion

The bistro recommender structure considering food search is developed based on kNN based Map Reduce approach. This can be useful to individual users. The methodology proposed may be useful to any person who visits an unknown place and can be helpful to separate the best bistros of his or her choice. The usage of Map Reduce in our paper is a novel idea because huge records can be processed with few milli-seconds.The choice of k-NN makes use of the machine learning thought to perceive the nearest Neighbors. In the future, the paper can be extended to group users.

References

1. *C.C.Aggarwal,RecommenderSystems:TheTextbook,SpringerInternationalPublishingSwitzerland2016.*
2. *Zhi-Dan Zhao, Ming-Sheng Shang, "User-based Collaborative-FilteringRecommendation Algorithms on Hadoop" Third International ConferenceonKnowledgeDiscoveryandDataMining2010.*
3. *Anjali Gautam and PunamBedi, "Developing content-based recommendersystem using Hadoop Map Reduce", Journal of Intelligent & FuzzySystems32 (2017)2997–3008,2017.*
4. *BidyutKr.Patra,RaimoLaunonen,VilleOllikainen,SukumarNandi,"AnewsimilaritymeasureusingBhattacharyya coefficientforcollaborativefilteringinsparsedata",KNOSYS3090,12March2015.*
5. *ChenyangLiKejingHe"CBMR:AnoptimizedMapReduceforitem-basedcollaborativefilteringrecommendationalgorithmwithempirical analysis", Concurrency and computation Practice and experienceVolume29,Issue1025 May2017.*
6. *Moon-HeePark,Han-SaemPark,andSung-BaeCho"RestaurantRecommendation for Group of People in Mobile Environments UsingProbabilistic Multi-criteria Decision Making", APCHI 2008, LNCS 5068,pp.114–122,2008.*