Online Application for Finding Localized Tutors Using KNN Algorithm

B Hari Babu[1], Dr K Venkataramana[2]

[1]*Student, Dept of Mca, Kmmips, Tirupati, Ap, India.*
[2]*Professor, Dept of Mca, Kmmips, Tirupati, Ap, India.*

Abstract- **This paper explains concerning java application for locating tutors employing a data processing technique. The increasing trend of personal or a bunch tuition has created it harder for a parent to search out an ideal tutor for his/her kid. The method of finding a coach is time consuming and even cash waste process. The increase in academic standards has lend to the requirement for added employment beside schooling. For the survival of today's cooperative world. Application for locating a decent tutor will facilitate a parent to search out tutors while not consulting any third party. The applying encourages folks and students to search out tutors in their near vicinities. The tutor application additionally considers the very fact that students want personal attention and steerage in studies. For this we tend to use KNN formula, to check tutor profiles. Beside the java platform we tend to user Graphical user interface (GUI), MYSQL is employed to store the tutor information. MYSQL formula classifies the tutor records and therefore helps in achieving most accuracy. The platform used for implementing data processing techniques. The general aim of datamining method is to extract information from an information set and rework it into a plain structure for any use. once tutor registers or signs up his/her information is keep in MYSQL. Then he will login in to the tutor portal. Parent will notice tutors supported 3 parameters, specifically space, board and preference. supported stipulated necessities, search results are retrieved from MYSQL and shall be flaunted to the user. This application really helps the oldsters for locating a decent tutor to his/her kids for his or her demand..**

Index Terms- **Java, Data Mining, MYSQL, KNN Algorithm, Classification, Tutoring.**

INTRODUCTION

Essentially, data mining is tied in with handling information and recognizing examples and patterns in that data, so you can choose or judge. Information mining standards have been around for a long time, at the same time, with the coming of enormous information, it is much more predominant. Enormous information caused a blast in the utilization of more broad information mining procedures, mostly on the grounds that the measure of the data is considerably bigger and in light of the fact that the data has a tendency to be more fluctuated and broad in its extremely nature and substance. With expansive informational indexes, it is never again enough to get generally basic and clear measurements out of the framework. With 30 or 40 million records of client data, realizing that two million of them live in one area isn't sufficient. You need to know whether those two million are a specific age gathering and their normal income, so you can focus on your client needs better. These business-driven necessities changed basic information recovery and insights into more mind-boggling information mining. The business issue drives an examination of the information that assembles a model to portray the data that at last prompts the making of the subsequent report. By the raising of the globalization there is a heavy competition between the students. As well As the parents can have the interest to provide a good education for their children. Today the society can expert more a knowledge from the students which means he has more technical skill along with that he/her can required communication skills. All the knowledge can be provided within the schooling is impassible. For this reason, parent can find the tutors. finding a tutor is o easy task. But finding a good tutor is important task. Sometimes it is difficult to find a good tutor. Due to this reason the parents and students can face many problems.

To overcome this problem, we can develop an online java application for finding a tutor according to their requirement. In this application all the details about

the tutors can be placed.the tutors can directly upload their details within the site. With this the students and parents can select their tutor with their requirements. There are many existing which can provide the details of tutors. But there are some drawbacks with these sites. It will take more time to search a tutor for our requirements. To overcome these problems, we can take a new online application. The intention behind creating the application is completely non-profit. The intension behind to develop this application is to provide a good platform between the students tutoress. Java application for finding tutors is to be used in the education center. Java application provides information regarding tuitions or coaching for students. It involves management of records, schedules, updates and search. This application is easy to use and understand. The main objective behind making this project is to enhance student and teacher relationship. Java application for finding tutors has been designed to make it convenient for students and teachers find each other. Students or parents can find tutors for online private tutoring. One can find tutors residing in the same vicinity as that of students. Students can find tutors for home tutoring or group tutoring. Tutors can list there profiles so that students can contact them. The application not only updates teacher information in terms of preferences of location, but also has options to choose specializations in their respective domain of education. The application makes an effort to ensure registration of skilled professionals and also encourages and attracts talented, highly enthusiastic and extremely passionate tutors to join the platform in order to contribute to the needs of students. This means the students are not only able to get best learning/training experience but also attain the support required to achieve their potential/goals. It allows individual teachers and institutes to create their profiles, add skills or experience etc. The application provides a demography with which parents can choose the best tutor by comparing the attributes and skill sets of the tutor. Using this application, a parent can easily find tutors for their child instead of asking any relative or friends. Parents would get their result instantly instead of waiting for days or waiting for someone else to respond. For finding a good tutor, education qualification, expertise in a subject/area, command over English language this application will help the tutors to

coordinate and save their time by managing their schedule. The application considers location preference as the prime attribute in order to assign teachers and students in the same vicinity. The significant features of the application maintainability, where data regarding tutor and parent will be maintained by database, profiles uploaded can be modified based on requirement, responsive, reliable and user-friendly.

KNN ALGORITHM

In this system, we have different techniques, categorizations, machine learning. In pattern recognition and classification, the k-NN algorithm is a non-parametric method used for regression and classification. In both cases, the input consists of the k closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression. In k-NN classification, the output is a class membership. An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor. In k-NN regression, the output is the property value for the object. This value is the average of the values of its k nearest neighbors.

k-NN is a type of instance-based learning, or lazy learning, where the function is only approximated locally, and all computation is deferred until classification. The k-NN algorithm is among the simplest of all machine learning algorithms. Both for classification and regression, a useful technique can be to assign weight to the contributions of the neighbors, so that the nearer neighbors contribute more to the average than the more distant ones. For example, a common weighting scheme consists in giving each neighbor a weight of 1/d, where d is the distance to the neighbor.

The neighbors are taken from a set of labels for which the class (for k-NN classification) or the object property value (for k-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required. A peculiarity of the k-NN algorithm is that it is sensitive to the local structure of the data. The

algorithm is not to be confused with k-means, another popular machine learning technique.

K-NEAREST NEIGHBORS ALOGORITHM

A case is classified by a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a distance function. If K = 1, then the case is simply assigned to the class of its nearest neighbor. The neighbors are picked up from a set of objects or objects having same properties or value, this can be considered as a training dataset. The algorithm uses the Euclidean distance, which is a straight path connecting two points. Before applying KNN algorithm on a dataset, the dataset must be prepared, that means, the dataset's parameters must be scaled down to a normalized scale. Euclidean distance between points A and B is the length of the line segment connecting them. The formula for Euclidean distance is given in equation below:

$$D(A,B)=\sqrt{[(x1-x2)]^2+[(y2-y1)^2]}$$
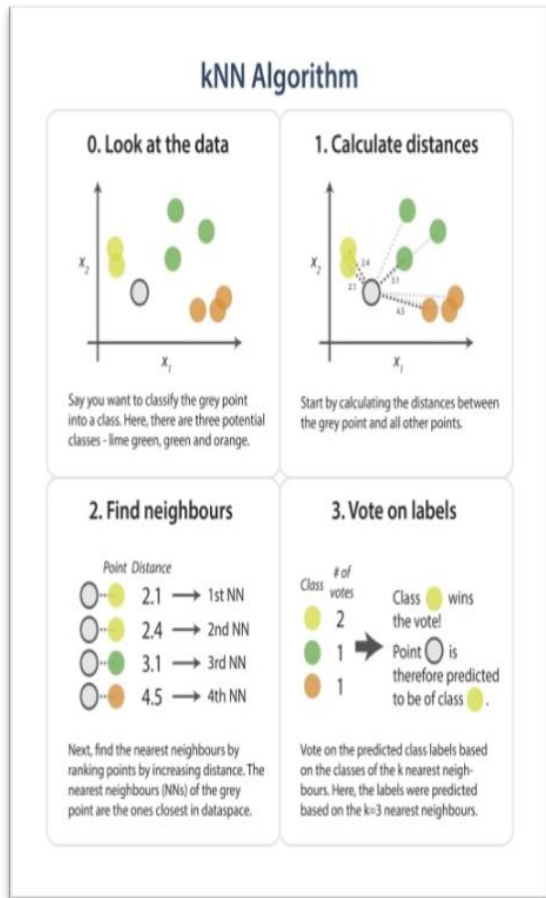
K nearest neighbors is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure (e.g., distance functions). KNN has been used in statistical estimation and pattern recognition already in the beginning of 1970's as a non-parametric technique.
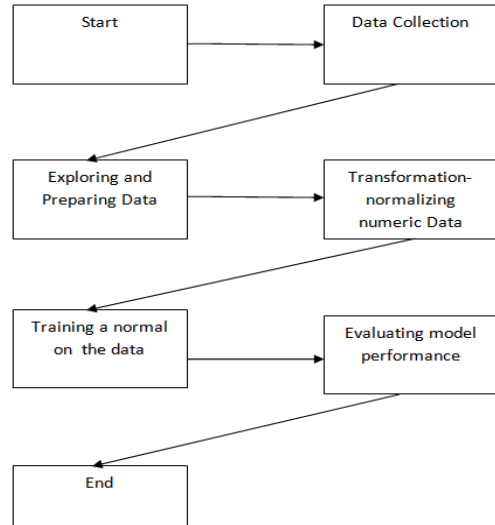
Flowchart of KNN algorithm

The flowchart shown in figure. Shows the stepwise demonstration KNN Algorithm working. Start with choosing a data set, once done, shuffle the dataset just in case it is in serial order, which would render the whole algorithm pointless. After getting the desired shuffled dataset, apply normalization to scale the minimum and maximum values of the parameters. Then take a portion of the data as a training model and apply our algorithm to it. Finally, evaluate and analyze the accuracy of the prediction of the algorithm.

CONCLUSION

In this paper, we are use KNN algorithm. By this algorithm we can implement technical or non-technical people, parents or children's or tutor. With the help of this application parents can find tutor for their smartphone and unemployed tutor can register him/herself as an employed tutor.

REFERENCE

[1] Agrawal R, Shrikant R (1994) Fast algorithms for mining association rules. In: Proceedings of the 20th VLDB conference, pp 487–499

[2] Banerjee A, Merugu S, Dhillon I, Ghosh J (2005) Clustering with Bregman divergences. J Mach Learn Res 6:1705–1749

[3] Bezdek JC, Chuah SK, Leep D (1986) Generalized k-nearest neighbor rules. Fuzzy Sets Syst18(3):237–256. http://dx.doi.org/10.1016/0165-0114(86)90004-7

[4] Bloch DA, Olshen RA, Walker MG (2002) Risk estimation for classification trees. J Comput Graph Stat 11:263–288

[5] Bonchi F, Lucchese C (2006) on condensed representations of constrained frequent patterns. Knowl INF Syst 9(2):180–201

[6] Breiman L (1968) Probability theory. Addison-Wesley, Reading. Republished (1991) in Classics of mathematics. SIAM, Philadelphia

[7] Breiman L (1999) Prediction games and arcing classifiers. Neural Comput 11(7):1493–1517

[8] Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth, Belmont Brin S, Page L (1998) the anatomy of a large-scale hypertextual Web Search Sngine. Comput Networks 30(1–7):107–117

[9] Chen JR (2007) Making clustering in delay-vector space meaningful. Knowl Info Syst 11(3):369–385

[10] Cheung DW, Han J, Ng V, Wong CY (1996) Maintenance of discovered association rules in large databases: an incremental updating 10(3):265–294

[11] Cost S, Salzberg S (1993) A weighted nearest neighbor algorithm for learning with symbolic features. Mach Learn 10:57.78 (PEBLS: Parallel Exemplar-Based Learning System)

[12] Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inform Theory 13(1):21–27

[13] Dasarathy BV (Ed) (1991) nearest neighbor (NN) norms: NN pattern classification techniques. IEEE Computer Society Press

[14] DeVore L, Gyorfi L, Lugosi G (1996) A probabilistic theory of pattern recognition. Springer, New York. ISBN 0-387-94618-7

[15] Dhillon IS, Guan Y, Kulis B (2004) Kernel k-means: spectral clustering and normalized cuts. KDD 2004, pp 551–556