PREDICTING EMPLOYEE PERFORMANCE USING PERSONALIZED ANALYTICS

B.Kannan[1] | R. Kumerasan[2]
[1, 2](Final year, Department of Computer Science and Engineering, Anand Institute of Higher Technology, Chennai, India)

Abstract— *Evaluating employee's performance and undesirable behavior in professional environments is an important task. Employee's performance Professional details & performance is based upon various factors like personal details or demographic, Social, Performance details etc. The data mining techniques are more helpful in classifying professional database and help us in evaluating the performance and undesirable behavior of a employee. Data mining techniques are widely useful in professional data mining for analysis of employee data. In professional area data mining different data mining techniques like classification, clustering, association rule mining. The main goal of data mining process is to extract information from large amount of data and to translate raw data into meaning full information. These databases contain hidden information for improvement of employees' performance. The performance in higher professional in India is a turning point in the performances for all employees. This performance is influenced by many factors, therefore it is essential to develop predictive data mining model for employees' performance so as to identify the difference between high learners and slow learners employee.*

Keywords— Employee Performance; Dot Net; Data Mining

1. INTRODUCTION

There is a huge amount of data available in the Information Industry. This data is of no use until it is converted into useful information. It is necessary to analyze this huge amount of data and extract useful information from it.

Extraction of information is not the only process we need to perform; data mining also involves other processes such as Data Cleaning, Data Integration, Data Transformation, Data Mining, Pattern Evaluation and Data Presentation. Once all these processes are over, we would be able to use this information in many applications such as Fraud Detection, Market Analysis, Production Control, Science Exploration, etc.

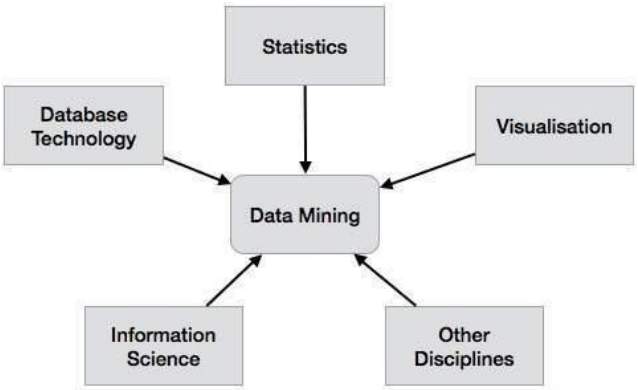
Data Mining is defined as extracting information from huge sets of data. In other words, we can say that data mining is the procedure of mining knowledge from data. The information or knowledge extracted so can be used for any of the following applications −

- Market Analysis
- Fraud Detection
- Customer Retention
- Production Control
- Science Exploration

Data mining deals with the kind of patterns that can be mined. On the basis of the kind of data to be mined, there are two categories of functions involved in Data Mining (1) Descriptive; (2) Classification and Prediction

A data mining system can be classified according to the following criteria

- Database Technology
- Statistics
- Machine Learning
- Information Science
- Visualization
- Other Disciplines



The main objective of higher professional institutions is to provide quality professional to their employees. One way to achieve highest level of quality in higher professional is by discovering employee's performance and undesirable behavior for employees who need special attention and allow the teacher to provide appropriate advising/counseling.

Database design is the process of producing a detailed data model of database. The logical data model contains all the needed logical and physical design choices and physical storage parameters needed to generate a design in a Data Definition Language, which can then be used to create a database. A fully attributed data model contains detailed attributes for each entity.

Datamining means predicting valuable data from raw data. In this system we predicting employee behavior and performance based on employee performance details, marks evaluation, attendance, social media interaction.so we have first loading sample or predefined datasets foe employees.in that we compare the employee database with sample dataset applying some datamining techniques such as Bayesian classification algorithm. This can be used to predict the behavior of employee performance.

2. RELATED WORKS

[1] Data mining is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from the professional context. This work is a survey of the specific application of data mining in learning management systems and a case study tutorial with the Moodle system.
[2] Though a number of prediction models have been developed, they either estimate a single model for all employees based on their past course performance and interactions with learning management systems (LMS), or estimate employee-specific models that do not take into account LMS interactions; thus, failing to exploit fine-grain information related to an employee's engagement.
[3] An enduring issue in higher professional is employee retention to successful graduation. To further this goal, we develop a system for the task of predicting employees' course grades for the next enrollment term in a traditional university setting.
[4] Data recorded while learners are interacting with Massive Open Online Courses (MOOC) platforms provide a unique opportunity to build predictive models that can help anticipate future behaviors and develop interventions. But since most of the useful predictive problems are defined for a real-time framework, using knowledge drawn from the past courses becomes crucial.
[5] The accurate estimation of employees' grades in future courses is important as it can inform the selection of next term's courses and create personalized degree pathways to facilitate successful and timely graduation. This paper presents future- course grade predictions methods based on sparse linear models and low-rank matrix factorizations that are specific to each course or employee-course tuple.

3. PROPOSED SYSTEM

The work aims to develop a trust model using data mining techniques which mines required information, so that the present professional system may adopt this as a strategic management tool.

The proposed system use professional data mining techniques to evaluate performance and identify undesirable behavior.

The approach may assist professional managers in supervising the development of employees at the end of each performance term, identifying the ones with difficulties to fulfill their requirements.

Figure 1: Architecture Diagram

Design is a multi- step that focuses on data structure software architecture, procedural details, algorithm etc. and interface between modules. The design process also translates the requirements into presentation of software that can be accessed for quality before coding begins. Computer software design change continuously as new methods; better analysis and border understanding evolved. Software design is at relatively early stage in its revolution. Therefore, software design methodology lacks the depth, flexibility and quantitative nature that are normally associated with more classical engineering disciplines. However, techniques for software designs do exit, criteria for design qualities are available and design notation can be applied.

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

➢What data should be given as input?

➢How the data should be arranged or coded?

➢The dialog to guide the operating personnel in providing input.

➢Methods for preparing input validations and steps to follow when error occur.

4. IMPLEMENTATION

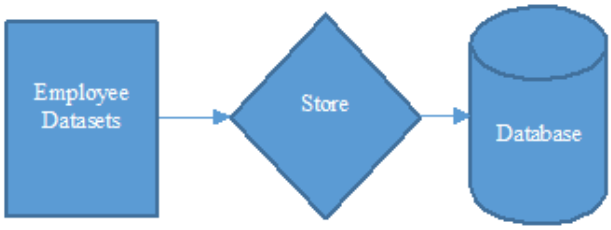
Bayes (1702-1761), who proposed the Bayes Theorem. Bayesian classification provides practical learning algorithms and prior knowledge and observed data can be combined.
Bayesian Classification provides a useful perspective for understanding and evaluating many learning algorithms. It calculates explicit probabilities for hypothesis and it is robust to noise in input data.

$$P(ai|vj) = Nc+M.P/N+M$$

N=no of Training Sample v=vj Nc=no of examples=vj and a=ai P= A priori estimate for P(ai|vj) M=equivalent sample size

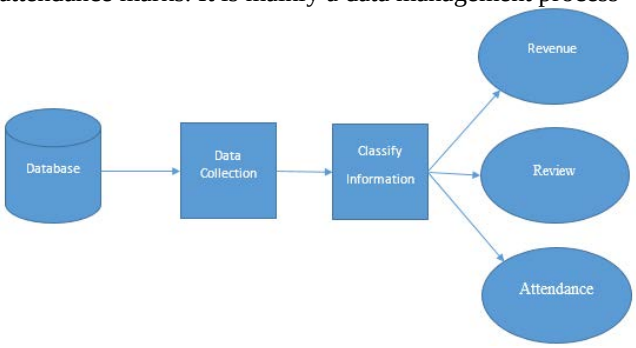
5. DATA COLLECTION

Data collection is the process of gathering and measuring information on targeted variables in an established systematic fashion, which then enables one to answer relevant questions and evaluate outcomes.
Collecting information about Employee's Demographic and Performance details. The data may contain employee's details of different subject marks in semester wise have been recorded and subjected to the data mining process.



6. DATA CLASSIFICATION

Data classification is the process of sorting and categorizing employee datasets into various types, forms or any other distinct class. Data classification enables the separation and classification of employee datasets according to semester marks, assessment marks and attendance marks. It is mainly a data management process

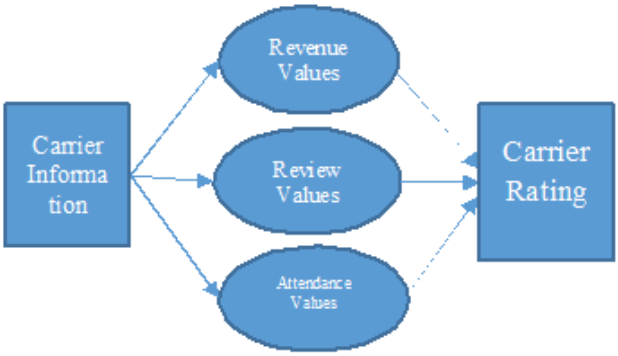


7. ORGANIZIND DATA

Organizing employee data into sensible groupings is one of the most. fundamental modes of understanding and learning. classification analysis is the formal. Bayesian classification algorithm and methods for grouping, or classifying, employee datasets.

We collected our college employee's real time data that describing the relationships between learning behaviour of employees and their performance.

Classification is a process which divides data into groups of similar objects. Data classification is the categorization of data for its most effective and efficient use. Using Bayesian classification algorithm to analysis the employee behaviour and predict the employee performance.



8. CONCLUSION AND FUTURE WORK

Sentiment and behaviour detection has a wide variety of applications in information systems, including classifying datasets. It is found that sentiment classifiers are severely dependent on employee behaviour. In this system Bayesian classification algorithm and it is used to classify the employee behavioural learning and employee performance. In this project, we were able to successfully implement the sentiment analyser tool for analyzing and classifying employee. In this system developed for predicting employee behaviour and performance based on comparing the predefined employee datasets. The performance of Bayesian classification algorithm implemented was satisfactory.

One of the most recent and biggest challenge that higher professional faces today is making employees skillfully employable. Many universities/institutes are not in position to guide their employees because of lack of information and assistance from their teaching-learning systems. To better administer and serve employee population, the universities/institutions need better assessment, analysis, and prediction tools. Considerable amount of work is done in analyzing and predicting performance, but all of these works are segregated. There is a clear need for unified approach. Other than performance attributes, there are large numbers of factors that play significant role in prediction, which includes noncognitive factors (set of behaviors, skills, attitudes). Suitable data mining techniques are required to measure, monitor and infer these factors for prediction. Thus enriching the input vector with qualitative values may increase the accuracy rate of prediction as well. Integrated Models/Frameworks are required for all the stakeholders of an Institution.

REFERENCES

[1] Nat'l Research Council, Building a Workforce for the Information Economy, Nat'l Academies Press, 2001.

[2] C. Romero, S. Ventura, and E. Garca,"Data Mining in Course Management Systems: Moodle Case Study and Tutorial," Computers & Professional, vol. 51, no. 1, 2008, pp. 368–384.

[3] L. Pappano, "The Year of the MOOC, "The New York Times, 2 Nov. 2012;

[4] Z. Pardos et al., "Adapting Bayesian Knowledge Tracing to a Massive Open Online Course in edX," Proc. 6th Int'l Conf. Professional Data Mining (EDM 13), 2013;

[5] A. Elbadrawy, R.S. Studham, and G. Karypis, "Collaborative Multi regression Models for Predicting Employees' Performance in Course Activities," Proc. 5th Int'l Conf. Learning Analytics and Knowledge (LAK 15), 2015, pp. 103–107.

[6] M. Sweeney, J. Lester, and H. Rangwala, "Next-Term Employee Grade Prediction," Proc. IEEE Int'l Conf. Big Data (Big Data 15), 2015, pp. 970–975.

[7] A. Polyzou and G. Karypis, "Grade Prediction with Course and Employee Speci c Models," to be published in Proc. 20th Pacic Asia Conf. Knowledge Discovery and Data Mining (PAKDD), 2016.

[8] S. Aud et al., The Condition of Professional 2013, NCES report no. 2013-037, Nat'l Center for Professional Statistics, US Department of Professional.