Improved aspect based opinion mining From student feedback review

Mrs. Farhanath.K[1]*, Dr. R. Senthil Kumar[2] & Dr. T. Senthil Prakash[3]
PG Student[1], Assistant Professor[2], Head of Department[3]
Department of Computer science and Engineering,
Shree Venkateshwara hi tech Engineering College,Gobi, Tamil Nadu, India.
farhanath.ashiq@gmail.com , yoursrsk@gmail.com

Abstract

Student'sopinion is vital for scholarly institutions so as to assess teacher's faculty performance.Taking care of the subjective assessments of students proficiently while programmed report is a difficult assignment.To be sure, most associations manage quantitative input viably, though subjective feedback is either handled physically or disregarded inside and out. In this research work study proposes a prediction model based on the integrating LSTM with ANN was developed to contribute in the field of EDM.The first layer predicts the viewpoints inside the feedback and later indicates the direction (positive, negative, and impartial) of those anticipated aspects. The students' performances have learned during the training phase with the instances and its class labels. During this phase, the ANN model had trained by adjusting its initial random weights assigned to each node in the hidden layer to compute the appropriate output. After the initial setup the weights are modified with the help of the LSTM where each population of instances is assigned with each node of the ANN. The simulation result of this proposed model is discussed with three different metrics with other two techniques namely PNN and ANN.Sentiment analysis expects to decide the sentiment value, frequently on a positive–negative scale, for a given product or service dependent on a lot of literature surveys.As fine-grained data is more valuable than only a solitary generally speaking score, current aspect-based sentiment analysis techniques separate the sentiment and dole out sentiment scores to different part of the product or service referenced in the review. To the best of our knowledge, this study is an attempt that uses deep learning approach for performing aspect based sentiment analysis on students' feedback or item review for evaluating faculty teaching performance and item review level.

Keywords:Data mining, Data Prepossessing, Knowledge discovery in databases, Opinion mining and Association rule, Long short term memory, Artificial neural networks, Probabilistic neural networks, Educational data mining, Support vector machine , Aspect based sentiment Analysis, Conditional Random Field, Convolutional neural networks, Faculty Evaluation System.

1. Introduction

Over the most recent couple of decades, advanced education in India has deteriorated in terms of quality of academic environment, quality, and level of research and the general level of the products (students, research, and development) of the system. The fundamental objective of any higher educational institution is to improve the quality of managerial decisions and to impart quality education. Good prediction of students' success in higher learning institution is one way to reach the highest level of quality in higher education system. There are numerous prediction models available with a difference in approach. The researcher; however, there is no certainty if there are any predictors who can accurately determine whether a student will be an academic genius, a drop out, or an average performer reports student's performances. Existing mechanisms have just not been able to respond to the emerging demands be made by higher and professional education as required by the industry.

This research work is means to specify problems and prospects of quality education in universities as well as higher educational institutions, and gives a model for improving the quality of education. The model contains approaches to deal with limited financial and skilled personnel using Information Technologies (IT). DM application in the area of education is wide spread. Researchers have explored various applications of DM in education. Different authors have gone through a survey of the literature to understand the importance of DM in higher education. This research work has mostly focused on the DM application from the domain perspective. To analyze the importance of DM techniques, such as, prediction, association, and clustering to predict the academic performances of students in higher education, we tried to study the main attributes that may affect students' performances in the courses using temporal DM techniques have not been explore to the extent necessary. This is the inspiration for my work in DM for higher education. Higher educational institutions use automated computer programs/tools developed with different technologies, to predict the available grades in the college.

Data mining is the procedure of consequently finding valuable data in huge repositories. Data mining is an essential part of knowledge discovery in databases (KDD), which is the general procedure of changing raw data into useful information. The focal subject of intrusion detection using data mining approach is to identify the security infringement in information system. Data mining can process huge measure of information and it finds hidden and disregarded data.

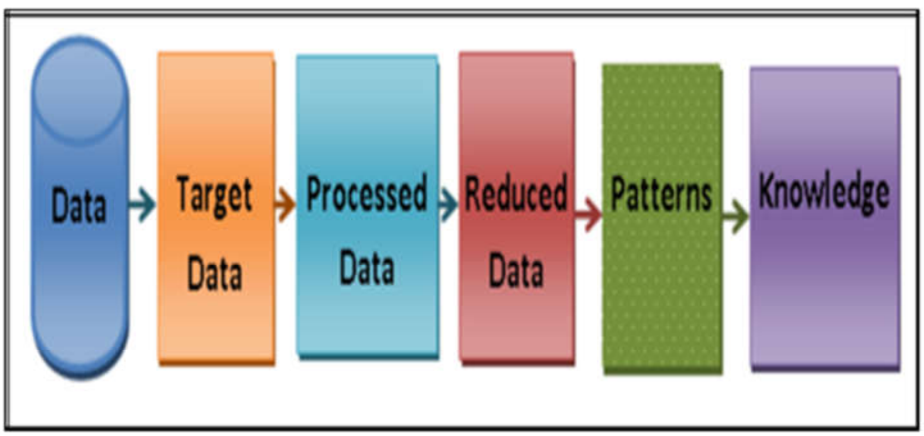

Figure 1.1: Steps in KDD Process

Data preprocessing is to change the raw input data into a suitable format for resulting analysis. Post processing step guarantees that only valid and valuable outcomes are joined into the decision support system. The difficulties that motivated the improvement of data mining are scalability, high dimensionality, heterogeneous and complex information, data ownership and distribution, and non- conventional analysis. Data mining tasks generally assembled into two types: isolating objects into gatherings (clustering i.e. descriptive) and assigning specific objects to these gatherings (classification i.e.predictive).

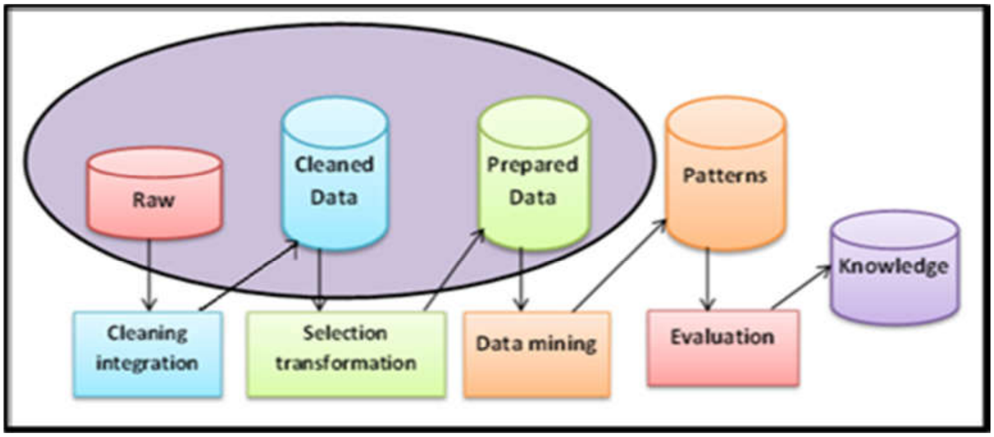

Figure 1.2: Data Pre-Processing in Data Mining (DM)

2. Review of Literature

System study is the primary stage of the system developed life cycle. This gives a clear picture of what actually the physical system is. The system study is done in two phases .In the first phase the preliminary survey of the system is done which helps in recognizing the scope of the system. The second phase of the system study is more detailed and in-depth study in which in the identifications of user's requirement and the limitation and problem of the present system are considered. After completing the system study, a system proposal is set up by the user.

Sindhu, S. Muhammad Daudpota, K. Badar, M. Bakhtyar, J. Baber and M. Nurunnabi, has proposed "Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation," It proposes a supervised aspect based opinion mining system based on two-layered LSTM model. The first layer predicts the aspects described within the feedback and later specifies the orientation (positive, negative, and neutral) of those predicted aspects [1]. M. Tubishat, N. Idris, and M. A. M. Abushariah evaluated ‘‘Implicit aspect extraction in sentiment analysis: Review, taxonomy, oppportunities, and open challenges,’’ Sentiment analysis is a text classification branch, which is defined as the process of extracting sentiment terms (i.e. feature/aspect, or opinion) and determining their opinion semantic orientation. At aspect level, aspect extraction is the core task for sentiment analysis which can either be implicit or explicit aspects. The growth of sentiment analysis has resulted in the emergence of various techniques for both explicit and implicit aspect extraction. However, majority of the research attempts targeted explicit aspect extraction, which indicates that there is a lack of research on implicit aspect extraction. This research provides a review of implicit aspect/ features extraction techniques from different perspectives. The first perspective is making a comparison analysis for the techniques available for implicit term extraction with a brief summary of each technique. The second perspective is classifying and comparing the performance, datasets, language used, and shortcomings of the available techniques [2].

N. Altrabsheh, M. Cocea, S. Fallahkhair, and K. Dhou proposed ‘‘Evaluation of the SA-E system for analysis of students,’’ It describes that Students' real-time feedback is acknowledged as an important source of information for teachers/lecturers to improve their teaching and address issues students may have, such as going deeper in some of the materials covered or providing more examples to understand an abstract concept. We developed the SA-E system for analyzing students' real-time feedback provided via social media, and, in this paper, we present the evaluation of this system in real settings with lecturers and students [3].M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub , Y. Jararweh, and B. Gupta, proposed a ‘‘Deep Recurrent neural network vs. support vector machine for aspect based sentiment analysis of Arabic hotels' reviews,’’ In this paper, state-of-the-art approaches based on supervised machine learning are presented to address the challenges of aspect-based sentiment analysis (ABSA) of Arabic Hotels' reviews. Two approaches of deep recurrent neural network (RNN) and support vector machine (SVM) are implemented and trained along with lexical, word, syntactic, morphological, and semantic features. The proposed approaches are evaluated using a reference dataset of Arabic Hotels' reviews [4].W. Wang, S. J. Pan, D. Dahlmeier, and X. Xiao proposed ''Coupled multi-layer attentions for co-extraction of aspect and opinion terms,'' In this paper describes the task of aspect and opinion terms co-extraction aims to explicitly extract aspect terms describing features of an entity and opinion terms expressing emotions from user-generated texts. To achieve this task, one effective approach is to exploit relations between aspect terms and opinion terms by parsing syntactic structure for each sentence. However, this approach requires expensive effort for parsing and highly depends on the quality of the parsing results. In this paper, we offer a novel deep learning model, named coupled multi-layer attentions. The proposed model provides an end-to-end solution and does not require any parsers or other linguistic resources for preprocessing. Specifically, the proposed model is a multi-layer attention network, where each layer consists of a couple of attentions with tensor operators. One attention is for extracting aspect terms, while the other is for extracting opinion terms. They are learned interactively to dually propagate information between aspect terms and opinion terms [5].

H. Xu, B. Liu, L. Shu, and P. S. Yu proposed ''Double embedding's and CNN-based sequence labeling for aspect extraction,'' This paper focuses on supervised aspect extraction using deep learning. Unlike other highly sophisticated supervised deep learning models, this paper proposes a novel and yet simple CNN model employing two types of pre-trained embedding's for aspect extraction: general-purpose embedding's and domain-specific embedding's. Without using any additional supervision, this model achieves surprisingly good results, outperforming state-of-the-art sophisticated existing methods. To our knowledge, this paper is the first to report such double embedding's based CNN model for aspect extraction and achieve very good results [6]. A. Cahyadi and M. L. Khodra, has proposed ''Aspect-based sentiment analysis using convolutional neural network and bidirectional long short-term memory' 'It describes that in order to improve performance of previous aspect-based sentiment analysis (ABSA) on restaurant reviews in Indonesian language, this paper adapts the research achieving the highest F1 at SemEval 2016. We use feed forward neural network with one-vs-all strategy for aspect category classification (Slot 1), Conditional Random Field (CRF) for opinion target expression extraction (Slot 2), and Convolutional Neural Network (CNN) for sentiment polarity classification (Slot 3). Aside from lexical features we also use additional features learned from neural networks. We train our model on 992 sentences and evaluate them on 382 sentences [7].

L. C. Yu, C. W. Lee, H. I. Pan, C. Y. Chou, P. Y. Chao, Z. H. Chen, S. F. Tseng, C. L. Chan, and K. R. Lai, evaluated a ''Improving early prediction of academic failure using sentiment analysis on self-evaluated comments''. This study presents a model for the early identification of students who are likely to fail in an academic course. To enhance predictive accuracy, sentiment analysis is used to identify affective information from text-based self-evaluated comments written by students. Experimental results demonstrated that adding extracted sentiment information from student self-evaluations yields a significant improvement in early-stage prediction quality. The results also indicate the limited early-stage predictive value of structured data, such as homework completion, attendance, and exam grades, due to data sparseness at the beginning of the course. Thus, applying sentiment analysis to unstructured data (e.g., self-evaluation comments) can play an important role in improving the accuracy of early-stage predictions. The findings present educators with an opportunity to provide students with real-time feedback and support to help students become self-regulated learners. Using the exploring results for improvement in teaching and learning initiatives is important to maintain students' performances and the effectiveness of the learning process [8]. A. Kumar and R. Jain proposed ''Faculty evaluation system,''. An automatic system to analyze the textual feedbacks of faculty members teaching in any institute is proposed. The proposed system extracts all the important aspects from the feedbacks and then sentiment score of each aspect for each faculty is calculated using machine learning algorithms. The proposed system is flexible and versatile than the existing feedback evaluation system of teachers where students evaluate the teachers on predefined aspects decided by experienced and senior faculty and administrators. Our system, Faculty Evaluation System (FES) identifies strengths and weaknesses of teachers on all those aspects which are important to students. This information may also be used by higher authorities of the institute to form appropriate teams of faculty members for different academic and administrative activities of the institute [9].Y. Ding, J. Yu, and J. Jiang proposed that‘‘Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction," Opinion target extraction is a fundamental task in opinion mining. It describes that In recent years, neural network based supervised learning methods have achieved competitive performance on this task. However, as with any supervised learning method, neural network based methods for this task cannot work well when the training data comes from a different domain than the test data. On the other hand, some rule-based unsupervised methods have shown to be robust when applied to different domains. In this work, we use rule-based unsupervised methods to create auxiliary labels and use neural network models to learn a hidden representation that works well for different domains. When this hidden representation is used for opinion target extraction, we find that it can outperform a number of strong baselines with a large margin [10].

2.1Opinion Mining and Sentiment Analysis

An important part of our information-gathering behavior has always been to find out what other people think. With the growing availability and popularity of opinion-rich resources such as online review sites and personal blogs, new opportunities and challenges arise as people now can, and do, actively use information technologies to seek out and understand the opinions of others. The sudden eruption of activity in the area of opinion mining and sentiment analysis, which deals with the computational treatment of opinion, sentiment, and subjectivity in text, has thus occurred at least in part as a direct response to the surge of interest in new systems that deal directly with opinions as a first-class object. This survey covers techniques and approaches that promise to directly enable opinion-oriented information-seeking system.

2.2 Educational Data Mining and LSTM Model

Educational Data Mining (EDM) is an emerging area. It plays an important role for emergent methods for surveying the special types of data that come from educational environment, and using those methods to better understand student’s behavior, and the settings which they learn. EDM concentrates on algorithms and developing new tools for discovering data patterns. EDM develops methods and applies techniques from machine learning, statistics and DM to analyze data collected during teaching and learning.
LSTM model for aspect extraction and sentiment classification.

- The first layer arranges a review sentence in one of the six aspects including-Teaching Pedagogy, Behavior, Knowledge, Assessment, Experience, and General.
- The second LSTM layer predicts the sentiment direction (+ve, −ve or Neutral) communicated towards that specific aspect.

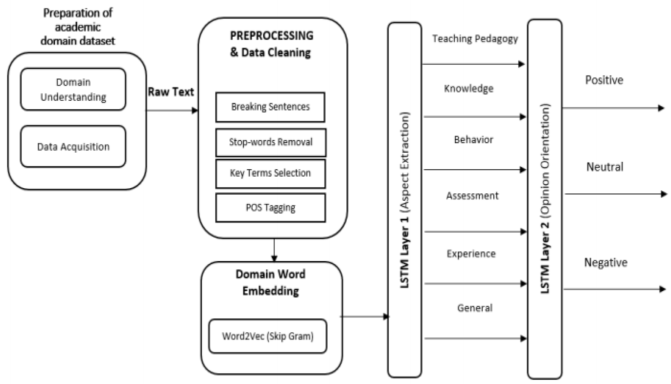

Fig: 2.2.1 LSTM Model

The working system of our proposed LSTM model that utilized in two stages of Aspect extraction and opinion orientation in a sequence naming task. So our proposed model categorizes a review in the predetermined classes by utilizing LSTM neural network. LSTM is one of the particular variants of RNN that conquers the issue of vanishing gradient problem.

It can store longer memory and better control on what to store and discard. Also, it has a gating component to control data stream inside the LSTM cell. These gates incorporates: input gate, forget gate and output gate denoted as fgt, igt and ogt individually. Every one of these gates are comprised of the sigmoid neural net layer and point wise multiplication activity. The yield of this sigmoid layer is somewhere in the range of 0 and 1. The current cell state of LSTM at timestamp t is signified as ct.

Our two layered LSTM model first procedures the input sentences where each word is indicated as a1, a2, a3, ..an. Next, we have utilized domain embedding layer in which input words are concatenated with the 100-dimensional word vectors bd and passed to the LSTM network layer.

$$\mathbf{cd} = \mathbf{bd} + \mathbf{at} (1)$$

While bd is the 100 dimensional word embedding vector, at is the vector of input words and their concatenation is represented by cd. At the point when the LSTM model gets the zd , it initially chooses what data to remove from previous output state h(t−1) so for this reason LSTM utilizes it's forget get as appeared in equation which finally outcomes a number somewhere in the range of 0 and 1.

3. Objectives

- Student's feedback can help the lecturers understand their students learning behavior and improve teaching.
- Allowing the lecturer to have an overall summary of the student's opinion.
- It can be used to extract information from the student feedback about the teaching and learning methods adopted in an educational institute.
- Collecting feedback has numerous benefits for the lecturer and their students, such as improvement in teaching and understanding student's learning behavior and improve teaching.

- Student's feedback improves correspondence between the lecturer and the students.
- Feature based opinion mining manages extraction of the various features of the feedback of the student in an educational organization.
- It can be utilized to separate data from the student feedback about the teaching and learning techniques embraced in an educational institute.
- In sentiment classification, machine learning techniques have been utilized to classify each question as positive or negative.
- Improving the performances of teachers through the study of their specialization and expertise and the time of their service in the educational process, and to evaluate and determine courses for needy teachers for improving their performances.
- Improving the performances of students and improves the curriculum and what was reflect in the educational process for the betterment of students, to improve students.

4. Hypotheses:

4.1 Proposed Hybrid Model

Educational Data Mining (EDM) and Learning Systematic (LS) research have appeared as motivating areas of research, which are used for beneficial understanding from educational databases for many purposes such as predicting students' success. The ability to predict a student's performance can be beneficial for actions in modern educational systems. Here discussed the development of an evolutionary approach inspired both on the Long short term memory (LSTM) and the Artificial Neural Network (ANN). The traditional ANN lacks performance due to the poor modelling structure and the capability of assigning proper weights to each node under the hidden layer. This problem is overcome with the aid of LSTM optimization approach which produces appropriate fitness function evaluation in each iteration of the learning process. The performances gradually increase the accuracy of the prediction and classification more precisely.

The most powerful methodology to analyze useful information from the data warehouse is DM. The concept of EDM is rapidly growing in the field of education which focuses on any type of the educational institutions, while academic analyst is particularly related to institutional effectiveness and students performance issues. Main objective of this research work is to mine the information discovered from the student database for improving the students' performance.

4.2Artificial Neural Networks (ANNs)

A basic structure of ANN contains three types of layers that are interconnected and named as input layer, hidden layer and output layer. Each layer is composed of one or more nodes, denoted by the small circles as shown in figure 4.2.1. The link between the artificial neural nodes signifies the stream of information from node to the next node. In the figure the flow starts from the input node to the output but some other network types may also have feedback links. The input layer nodes are called as submissive because they do not change the data. They receive the single data from the external source and duplicate them to their many outputs. Hidden layers and output are referred as active nodes, that is, they change the value of data received by them.

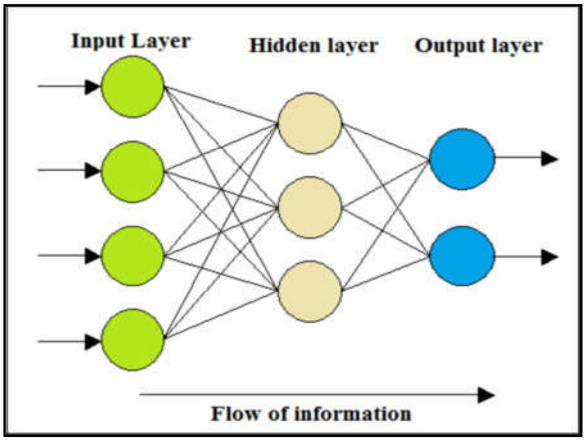

Figure 4.2.1: A Simple Artificial Neural Network (ANN)

All the hidden nodes are fully interconnected with the input nodes. The value received by the hidden node is multiplied by a set of predetermined numbers termed as weights. These weights are summed together to produce a single value as shown in figure 4.2.2 by the symbol□. Then a nonlinear mathematical function is applied to the value which controls the output of the nodes. This nonlinear function is represented in the S shaped curve and the function is termed as sigmoid. The input of the sigmoid function is between -∞ and +∞, whilst its output can merely be between the interval of 0 and 1.

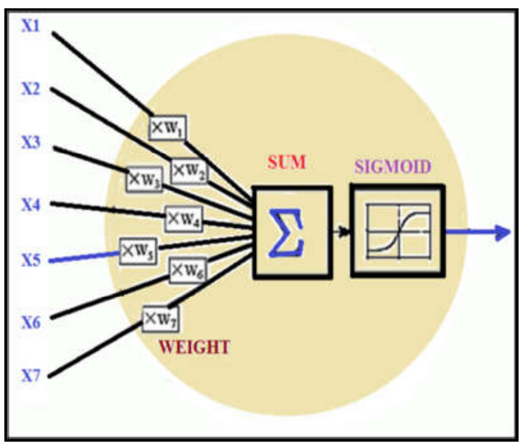

Figure 4.2.2: Neural Network Active Node in Hidden Layer and Output Layer

Like the hidden layer the same process is followed in the output layer's active node, and it also integrates and alters the data to generate the output. Generally neural networks can have number of layers and nodes on each layer. For analysis of detecting the target, the output layer contains only one single node. The value of the output node is threshold to produce a sign of negative for absence or positive sign for presence depending on the input data. Even though the function of processing unit looks very simple and there is no interesting inference in its process, the complete prospective and its influence is mainly when these nodes are interlinked or interconnected. The manner in which these artificial neurons are linked is termed as topology or architecture of an ANN.

4.3 Working Model of Artificial Neural Networks

ANN is a computational mechanism able to acquire, represent, and compute a mapping from multivariate space of information to another, given a set of data representing that mapping. Artificial neural network is trained by the use of a set of examples of associated input and output values. The purpose of an artificial neural network is to build a model of the data generating process so that the network can generalize and predict outputs from inputs that it has not previously seen.

The most frequently used neural network method is the back propagation-learning algorithm. This is a learning algorithm of multi-layered neural network, which consists of an input layer, hidden layers, and an output layer as shown in fig 5.3.1 The hidden and output layer neurons process their inputs by multiplying each of their inputs by the corresponding weights, summing the product, then processing the sum using a nonlinear transfer function to produce a result. Artificial neural network "learns" by adjusting the weights between the neurons in response to the errors between actual output values and target output values. At the end of this training phase, the neural network represents a model, which should be able to predict a target value given an input value.
Formally, the input that a single node receives is weighted according to Equation (2).

$$net_j = \sum_i w_{ij} \cdot O_i \quad (2)$$

Where w_{ij} represents the weights between nodes i and node j, and O_i is the output from node j such as Equation (3).

$$O_j = f(net_j) \quad (3)$$

The function f is usually a non-linear sigmoid function that is applied to the weighted sum of inputs before the signal processes to the next layer. Advantage of the sigmoid function is that its derivative can be expressed in terms of the function itself such as Equation (4).

$$f'(net_j) = f(net_j)(1 - f(net_j)) \quad (4)$$

The Multi-layer Perceptron (MLP) can separate data that are non-linear because it is 'multi-layer', and it generally consists of three types of layers. The first layer is the input layer, where the nodes are the elements of a feature vector.The second type of layer is the internal or 'hidden' layer since it does not contain output unit.

The third type of layer is the output layer and this presents the output data. Each node in the network is interconnected to the nodes in both the preceding and following layers by connections. These connections have associated weights with them (Atkinson and Tatnall, 1997).

The error, E, for one input training pattern, t, is a function of the desired output vector, d, and the actual output vector, o, given by Equation (5).

$$E = \frac{1}{2}\sum_k (d_k - O_k) \quad (5)$$

The error back propagated through neural network and the error is minimized by changing the weight between layers. So, the weight can be expressed in Equation (6).

$$w_{ij}(n+1) = \eta(\delta_j . O_i) + \alpha \Delta w_{ij} \quad (6)$$

Where the learning rate parameter, j is ηδ an index of the rate of change of the error, and α is the momentum parameter. This process of feeding forward signals and back propagating the error is repeated iteratively until the error of the network as a whole is minimized or reaches an acceptable magnitude. Using the back-propagation, the weight of the each factor can be recognized and the weight can be used to weight of classification. Zhou (1999) described the method of determination of the weight using back-propagation.

From equation (3), the effect of an output oj a hidden layer node j on the output ok from an output layer node k can be represented by the partial derivative of ok with respect to oj such as Equation (5.6).

$$\frac{\partial o_k}{\partial o_j} = f'(net_k) \cdot \frac{\partial (net_k)}{\partial o_j} = f'(net_k) \cdot w_{jk} \qquad (7)$$

Equation (7) equation can produce values with both positive and negative signs. If only the magnitude of the effects is of interest, the importance of node j relative to another node j0 j0 in the hidden layer can be calculated as the ratio of the absolute values from the equation (7) equation.

$$\left|\frac{\partial o_k}{\partial o_j}\right| / \left|\frac{\partial o_k}{\partial o_{j0}}\right| = \frac{\left|f'(net_k) \cdot w_{jk}\right|}{\left|f'(net_k) \cdot w_{j0k}\right|} = \frac{\left|w_{jk}\right|}{\left|w_{j0k}\right|} \qquad (8)$$

Equation (8) shows that, with respect to a particular node in the output layer, the relative importance of a node in the hidden layer is proportional to the absolute value of the weight on its connection to the node in the output layer. When more than one node in the output layer is concerned, equation (8) cannot be used to compare the importance of two nodes in the hidden layer.

$$W_{j0k} = \frac{1}{J} \sum_{j=1}^{J} \left|w_{jk}\right| \qquad (9)$$

$$t_{jk} = \frac{\left|w_{jk}\right|}{\frac{1}{J} \cdot \sum_{j=1}^{J} \left|w_{jk}\right|} = \frac{J \cdot \left|w_{jk}\right|}{\sum_{j=1}^{J} \left|w_{jk}\right|} \qquad (10)$$

Therefore, with respect to node k, each node in the hidden layer has a value greater or smaller than one, depending on whether it is more or less important than the average, respectively. With respect to the same node, all the nodes in the hidden layer have a total importance such as equation (11),

$$\sum_{j=1}^{J} t_{jk} = J \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (11)$$

Consequently, with respect to all nodes in the output layer, the overall importance of node j can be calculated as equation (12).

$$t_j = \frac{1}{K} \cdot \sum_{j=1}^{K} t_{jk} \qquad (12)$$

Similar to Equation (12), with respect to node j in the hidden layer, the normalized importance of node j in the input layer can be defined as equation (13).

$$S_{ij} = \frac{\left|\omega_{ij}\right|}{\frac{1}{1} \cdot \sum_{i=1}^{I} \left|\omega_{ij}\right|} = I \cdot \frac{\left|\omega_{ij}\right|}{\sum_{i=1}^{I} \left|\omega_{ij}\right|} \qquad (13)$$

With respect to the hidden layer, the overall importance of node i, is given by equation (14)

$$S_i = \frac{J}{-} \cdot \sum_{j=1}^{J} s_{ij} \qquad (14)$$

Correspondingly, the overall importance of input node i with respect to output node k is given by Equation (15).

$$St_i = \frac{J}{-} \cdot \sum_{j=1}^{J} s_{ij} \cdot t_j \qquad (15)$$

4.4Probabilistic Neural Networks (PNN)

PNN is a feed forward neural network formed by Ibrahim. It is established based on Bayesian network and Kernel Fisher differentiate analysis. In a PNN, the actions are systematized into a multi-layered feed forward network. First layer is the input layer in which one neuron is existent for each autonomous variable. The next layer is the hidden layer. This layer comprises one neuron for each set of training data. It not only stores the values of the each predictor variables but also stores each neuron along with its target value. Next is the Pattern layer.

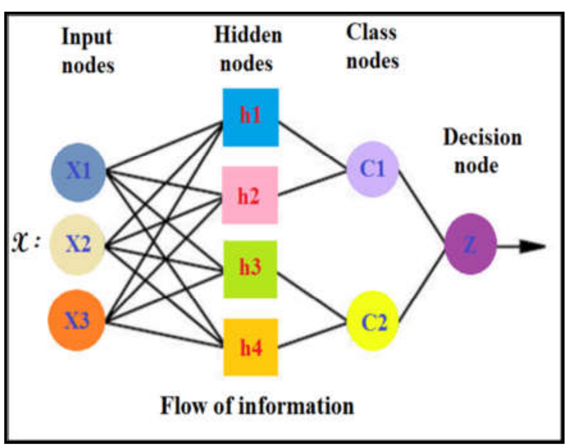

Figure 4.3.1: Structure of Probabilistic Neural Network (PNN)

4.5Design of the Proposed Evolutionary Technique

This section gives a brief note on the use of a hybrid approach ANN and LSTM termed as Hybrid approach for predicting students' performances.

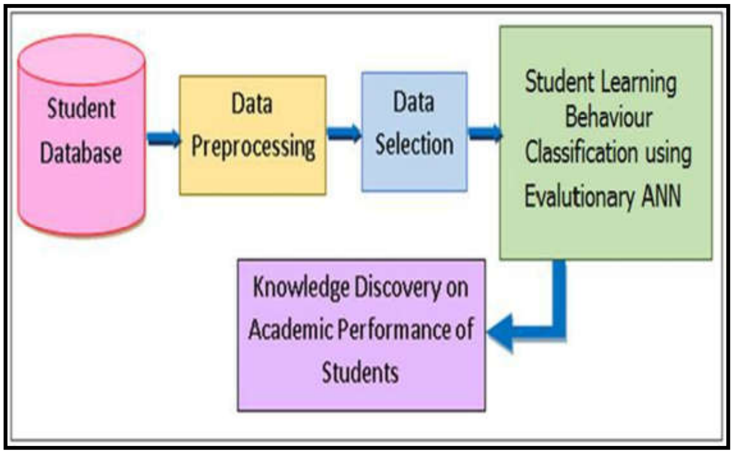

Figure 4.5.1: Student Learning Behaviour Classification using Hybrid Technique

A prediction model based on the integrating LSTM with ANN was developed to contribute in the field of EDM. The working principle of conventional ANNs and PNNs is illustrated in this chapter. The students' performances have learned during the training phase with the instances and its class labels. During this phase, the ANN model had trained by adjusting its initial random weights assigned to each node in the hidden layer to compute the appropriate output. After the initial setup the weights are modified with the help of the LSTM where each population of instances is assigned with each node of the ANN. The simulation result of this proposed model is discussed on the next chapter with three different metrics with other two techniques namely PNN and ANN.

4.6 Fast Algorithms for Mining Association Rules

This paper consider the problem of discovering association rules between items in a large database of sales transactions .It present two new algorithms for solving this problem that are fundamentally different from the known algorithms. Empirical evaluation shows that these algorithms outperform the known algorithms by factors ranging from three for small problems to more than an order of magnitude for large problems. We also show how the best features of the two proposed algorithms can be combined into a hybrid algorithm, called AprioriHybrid. Scale-up experiments show that AprioriHybrid scales linearly with the number of transactions. AprioriHybrid also has excellent scale-up properties with respect to the transaction size and the number of items in the database.

4.6.1 Hybrid Algorithm

The procedure hybrid approach can be described as follows:
step 1: Initialize parameters of LSTM model.
step 2:Define the structure of the BP neural network and generate genes.

step 3: Each gene corresponds to a neural network connection weight or threshold.
step 4: Choose fitness function, positive neutral and negative iteration time

Randomly initialize the fitness function and apply it to all the genes.
step 5: The best gene with highest fitness is stored and carried for any iteration
step 6:Evaluate each new fitness value
step 7: If the maximal iteration times or the fitness values are met stop the iteration, otherwise, the process repeated from step3.
step 8: Enchanting the weights which is optimized by LSTM as the initial parameters, the network makes independent learning
step 9:The learning phase of the Hybrid approach as follows:
step 10: Initialize the network. The network structure, predictable output, and learning rate are resolute conferring to the trial features. The LSTM is castoff to originate the optimal discrete elucidation for the early weight value and threshold of the BPN
step 11: Input the teaching sample and calculate the output of the ANN layers.
step 12: Calculate the learning error of the ANN
step 13: Spot-on the link weight values and thresholds of the hidden and output layers.
step 14: Evaluate error generated is satisfied or not and the anticipation necessities and whether the amount of reiterations has met its training boundary. If whichever state is met then the training stops. Else, learning iteration process carry on.

4.7 Advantages of Proposed System

- It provides better accuracy and efficiency than the other algorithms.
- To provide a better approach to teaching, learning and management process of the education system
- For the better decision making in the learning environment
- This system will be useful to improve the students learning and instructor's methods of delivery
- Scales to huge datasets
- It reduces the efforts and time people spend hunting for reliable reviews.
- People can view the opinions based on the features, one is interested in.
- People can get a cumulative count of positive and negative reviews based on the features.
- It gives an overall comprehensive compilation of opinions by fusing sentiments

5. Methodology

5.1 Reading Data and Preprocessing

The main module is the procedure of reading datasets and preprocessing. The proposed framework utilized Student feedback reviews of dataset. The dataset has a set of descriptions in a text file. The framework uses 300 dataset for the assessment. This module gathers those data's and stores into the database for further process.

Preprocessing is the procedure of elimination, which eliminates stop words from the dataset before processing. Stop words are regular language words which have very less meaning, such as "and", "the", "an", "a", and comparable words.

5.2 POS Tagging

The reviews are sent to the POS tagging module where POS tagger tag all the words of the sentences to their proper part of speech tag. POS tagging is a significant period of opinion mining, it is important to decide the features and opinion words from the reviews. POS tagging should be possible physically or with the assistance of POS tagger. Manual POS tagging of the reviews take bunches of time. Here, POS tagger is utilized to tag all the words of reviews.

5.3 Aspect Extraction

All the features are take out from the reviews and keeping in a database then its corresponding opinion words are take outfrom these reviews with the assist of POS tagger. Then recurrent words are identified. Recurrent features are the “hot” features that individuals remark most about the given item. In any case, there are few features that only a small number of individuals discussed. These features can also be interesting to some potential clients and the producer of the item. Rare words are not considered for aspect extraction.

An item is an entity which can be subject, Experience, individual, association, or Department. It is related with a set of components or attributes, called aspects of the object. Each component may have its own set of aspects. An opinion is essentially a positive or negative view, attitude, or emotion about an object or an aspect of the object from an individual or an association. Given a collection of opinion texts on an item, the aspect extraction problem is to produce the aspects of the item from these documents.

It is simple to recognize a set of opinion words such as “good” and “bad,” etc. The next is that opinion words are usually associated with aspects under certain syntactic relations. We can determine a set of aspects in terms of syntactic relations. So also, syntactic clues can help to take out new aspects from the extracted aspects, and new opinion words from the extracted aspects. This propagation procedure proceeds until no more opinion words or aspects can be extracted.

5.4 Hybrid Model Detection and Classification

The polarity of the sentences is determined for each feature. Polarity is determined on the basis of majority of opinion words, if the number of positive words is more, then the polarity of the sentence is positive otherwise the polarity is negative and if the number of positive and negative words is equal then the sentence shows the neutral polarity. We now reach the step of predicting the orientation of an opinion sentence, i.e., positive or negative. In general, we use the dominant orientation of the opinion words in the sentence to determine the orientation of the sentence. That is, if positive/negative opinion prevails, the opinion sentence is regarded as a positive/negative one. The final result shows the total number of positive and negative aspects. Using this result, users can understand about the effectiveness of the Student feedback.

6. Result and Discussions

The records in the raw data set will then undergo data pre-processing following which, the various records will then be subjected to clustering. In this step, records having different attribute values will be grouped under various clusters wherein each cluster will be mapped.Onto a particular class label. The data set at this stage will be split into binary smaller data sets –the training data set and the testing data set. The training data set will be recycled for machine learning along with learning algorithm.

1. Performance Analysis of LSTM and Proposed

	Accuracy	Precision	Recall
LSTM	93	88	85
Proposed	94.2	89.3	87.1

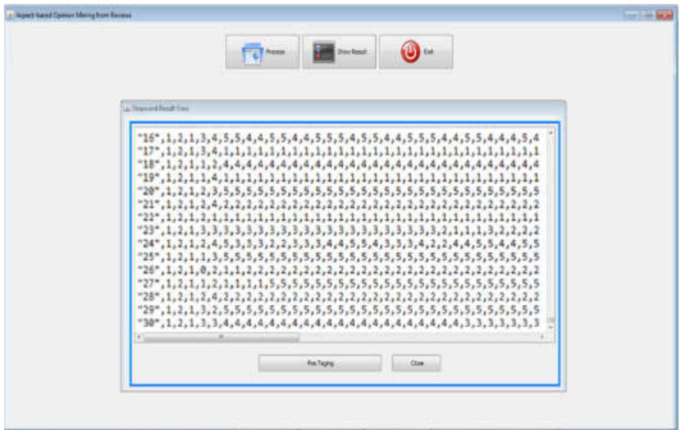

Figure 6.1: Aspect based opinion mining stream

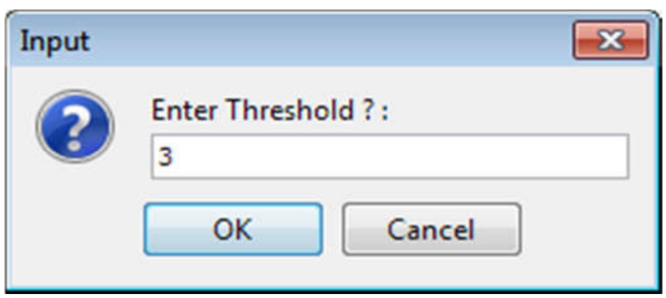

Figure 6.2: The threshold value declaration

Figure 6.2 illustrates the concept of threshold value which may vary from 0 to 5 according to the dataset. Since, the process is limited based on the threshold value. When the number of attributes was small, Conventional LSTM based classification methods have resulted in lower Re-call feature than the proposed HW-CNN method. From this experiment, a comparison of the F-measure determines the procedures revealed that the hybrid method is one of the best clustering algorithms. Generally, Recall also known as Sensitivity, which is defined as the ratio of the correctly positive labelled text by classification algorithm.

As shown in equation 5.1, the recall is defined with the ratio of True Positive (TP) and actual positive rate (TP+FP).

$$Recall = TP/\ (TP+FN) \qquad (16)$$

The sensitivity of the proposed HW-CNN resulted in the maximum Re-call feature. Since, the LSTM and hybrid HW-CNN method have the maximum sensitivity factor 93 and 94.2, respectively.

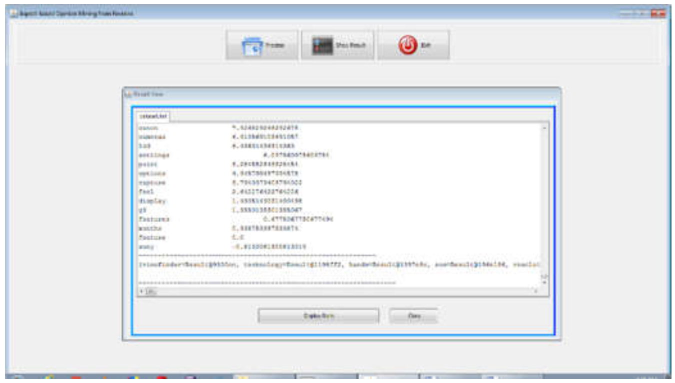

Figure 6.3: Review extraction

Figure 5.8 shows that the concept of review extraction in terms of features. Since, the concept is acceptable if we proceed with the hybrid classification algorithm. The features are extracted by using threshold value. Sensitivity and specificity are inversely proportional, meaning that as the sensitivity increases, the specificity decreases and vice versa.

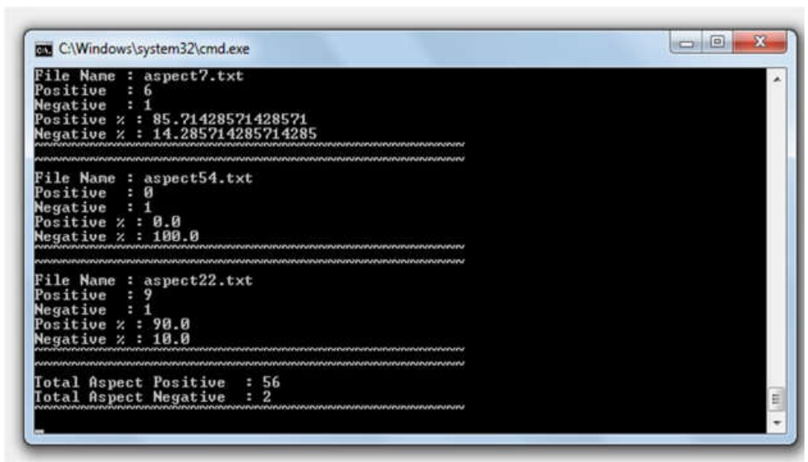

Figure 6.4Displaying the number of positive and negativeaspects

7. Limitations

The essential and the most pivotal advance of any data mining task is the collection or gathering of data. Inferable from the security and legal restrictions that real – time information of the students who were enrolled in a university possessed, we were not given the access to it. Consequently, we haveto locate an alternative method of reproducing or simulating real – time student data. We did as such with the help of some chose variables gathered through survey technique and taken the feedback from subject specialists which helped us create our own attributes which firmly resembled real – time student information. This information set contains twenty seven thoughtfully picked attributes for a thousand records or occasions. Such as a final result, having the facts prompted via our examination, gathering would have the option to see college students at danger early, and give better training for weak students. The future work may be at the way of hypothesis testing which notable shows that what are the exceptional forms of attributes have an effect on scholarly performance. After testing a few assumptions, some of most impacting factors may be distinguished and taken to anticipate the explanations for failure of the weak students.

7. Conclusion

The application of diverse data mining and machine learning knowledge of strategies inside the domain of training records mining offers which means facts that during flip allows inside the selection making method Though the significance of analyzing and predicting the reasons for the bad overall performance of the scholars has been addressed in other disciplines like psychology and sociology, its significance has been much less addressed in the vicinity of tutorial records mining. The major objective of these studies portraits is to discover the motives that make a contribution to the poor performance of college students in educational institutions, the ANN set of rules and LSTM based sentiment analysis for mining aspects relating to specified labels (Positive or Negative) or groupings of Student feedback reviews. This study may even work to pick out the ones students who wanted special interest to lessen failure degree and taking suitable motion for the following semester examination. The ANN proposed reduces the mistake prices and improves the overall performance. Experiments and consequences have proved our claims regarding prediction and accuracy.

7. References

[1]Sindhu, S. Muhammad Daudpota, K. Badar, M. Bakhtyar, J. Baber and M. Nurunnabi,"Aspect-Based Opinion Mining on Student's Feedback for Faculty Teaching Performance Evaluation",In IEEE Access, vol. 7,**(2019)**, pp. 108729-108741.

[2] M. Tubishat, N. Idris, and M. A. M. Abushariah, "Implicit aspect extraction in sentiment analysis: Review, taxonomy, opportunities, and open challenges", Inf. Process. Manage, vol. 54, no. 4,**(2018)**, pp. 545–563.

[3]N. Altrabsheh, M. Cocea, S. Fallahkhair, and K. Dhou,"Evaluation of the SA-E system for analysis of students",In Proc. IEEE 17th Int. Conf. Adv.Learn. Technol. (ICALT), Jul, **(2017),**pp. 60–61.

[4] M. Al-Smadi, O. Qawasmeh, M. Al-Ayyoub , Y. Jararweh, and B. Gupta,"Deep Recurrent neural network vs. support vector machine for aspectbased sentiment analysis of Arabic hotels' reviews", J. Comput. Sci., vol. 27,**(2018)**,pp. 386–393, Jul.

[5] W. Wang, S. J. Pan, D. Dahlmeier, and X. Xiao,"Coupled multi-layer attentions for co-extraction of aspect and opinion terms", In Proc. 31st AAAI Conf. Artif. Intell., Feb. 2017, **(2017)**, pp. 3316–3322.

[6] H. Xu, B. Liu, L. Shu, and P. S. Yu,"Double embedding's and CNN-based sequence labeling for aspect extraction", May **(2018)**, arXiv:1805.04601.

[7] A. Cahyadi and M. L. Khodra,"Aspect-based sentiment analysis using convolutional neural network and bidirectional long short-term memory",In Proc. 5th Int. Conf. Adv. Inform., Concept Theory Appl. (ICAICTA), **(2018)**,pp. 124–129.

[8]L. C. Yu, C. W. Lee, H. I. Pan, C. Y. Chou, P. Y. Chao, Z. H. Chen, S. F. Tseng, C. L. Chan, and K. R. Lai,"Improving early prediction of academic failure using sentiment analysis on self-evaluated comments", J. Comput. Assist. Learn., vol. 34, no. 4, **(2018)**,pp. 358–365.

[9]A. Kumar and R. Jain,"Faculty evaluation system", Procedia Comput. Sci., vol. 125, **(2018)**, pp. 533–541.

[10]Y. Ding, J. Yu, and J. Jiang, "Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction", In *Proc. 31st AAAIConf. Artif. Intell.*, Feb. **(2017)**, pp. 3436_3442.