Abstract
In this paper, we apply the data mining techniques on the students to find their psychological and learning perspectives. This approach will provide the faculty with a road map to tailor various programs for the intellectual, technical and psychological development of the students. Using student's data provides the university with the opportunity to make more informed strategic decisions. The challenge is of identifying useful information in vast student resources databases that will comprise not only of past and current academic performances but also the complex psychological attributes and learning behaviors. The information obtained by data mining on the behavior and potential of its students can be used towards enhancing the efficiency of students and thus enriching the teaching methodologies leading to increment in overall productivity. University that admits thousands of students and track their academic related information might find valuable information patterns contained within these databases to provide insights in such areas as performance enhancement, intellectual progress, psychological development and placements.

Index Terms: Data mining, Psychological aspects, Learning trends, Education industry, Personality.
Learning trends is the approach adopted by different individuals to look at same thing from different perspectives. It signifies the methodology and way of learning things. In order to explore & predict the cause & effect on performance of the various human variables like psychological traits & learning trends, lots of researchers have been trying to unfold the relationships between them.

Besides the value addition to the intellectual literature this area of research can prove to be of vast importance as it may be able to provide answers to various problems faced by Universities in recording the psychological inputs of their students & relating them with their learning styles so that they can adopt the methodologies in such a way that individual differences can be well harnessed. Through data mining they can create more flawless calculations & assessments.

This can be also used by various organizations in identifying the traits of their employees & assigning the right task in accordance to their potential competencies. Hence it can be used as a valuable tool for appraisal, training, career planning & development, etc.

LITERATURE REVIEW

The interest between personality traits and academic performance relationship has persisted throughout the 20th century. During this period, investigators have adopted several theoretical approaches to the topic, involving distinct conceptualizations of the relevant personality dimensions.

Early research efforts focused on the relation between academic performance and a broad personality trait termed persistence of motives (Webb, 1915). More recently, research has examined the relations between academic achievement and the personality dimensions proposed in Cattell’s (1973) and Eysenck’s (1970) models of personality structure.

The focus of this paper is with the most recent approach on personality traits and academic achievement, basically that is based on the Five-Factor Model of personality structure. In the past a strong & consistent association had been observed between conscientiousness & academic success. In addition a mild positive association is reported between openness to Experience & scholastic achievements in contrary to the mild negative association between Extraversion & scholastic achievements has been observed in Connor’s (2007).

David A. Kolb is a Professor of Organisation Behavior in the Weatherhead School of Management, Ohio, USA. Besides his work on experiential learning, he is well known for his contribution in the area of organizational behavior.

1 Kolb’s Learning Trends
The four learning trends according to the Kolb’s Model are:

1.1 Diversers (Reflective observer/Concrete experiencer)
Diversers take experiences and think deeply about them. They like to ask ‘why’, and will start from detail to constructively work up to the big picture. They prefer to watch rather than do, tending to gather information and use imagination to solve problem.

1.2 Convergers (Active experimenter/ Abstract conceptualization)
Convergers are the people who think about things and then try out their ideas to see if they work in practice. They like to ask ‘how’ about a situation, understanding attracted to new challenges and experiences, and to carrying out plans. They commonly act on ‘gut’ instinct rather than logical analysis.

1.3 Accommodators (Active experimenter/ Concrete experiencer)
Accommodators prefer the most hands-on approach & thus they have a strong liking for doing rather than thinking. They like to ask ‘what if?’ and ‘why not?’ to support their action-first approach. People with an Accommodating learning style like to work in teams to complete tasks.

1.4 Assimilators (Abstract conceptualizer/ Reflective observer)
Assimilators have the most cognitive approach, preferring to think than to act. They ask ‘What is there I can know?’ and like organized and structured understanding. People with an Assimilating learning style are more interested in
ideas and abstract concepts and less focused on people. People with this style are more attracted to logically sound theories than approaches based on practical value.

2 The Psychological Traits
The psychological traits give a detail insight into how the Psychology plays an important part in professional and intellectual development of students and how through data mining we can learn the trends and can predict the outcomes and can improve the overall performance of students. Data mining is a process of uncovering hidden knowledge from past amount of data.

The description of the Psychological traits provided by the Big five model has been used for the analysis in this research study, the details of are mentioned as follows:

2.1 Extraversion
Extraversion is defined by pronounced engagement with the external world. Extraverts love to be with people, are full of energy, and often experience positive emotions. The individuals who mark high on this scale are having following traits: 1) Friendliness, 2) Gregariousness, 3) Assertiveness, 4) Activity Level, 5) Excitement seeking, 6) Cheerfulness.

2.2 Agreeableness
Agreeableness reflects individual differences in concern with cooperation and social harmony. Agreeable individual's value getting along with neurosis, but they differ in their degree of suffering and specific symptoms of distress. Those who score high on agreeableness are: 1) Considerate, 2) Friendly, 3) Generous, 4) Helpful, 5) Willing to compromise their interests with others.

2.3 Conscientiousness
Conscientiousness explains the way in which we control, regulate, and direct our impulses. Impulses are not basically bad occasionally when time constraints require a snap decision then acting on our first impulse can be an effective response. Impulsive individuals can be judged by others as colorful, fun-to-be-with, and wacky. The individuals who mark high on this scale are having the following traits: 1) Self-Efficacy, 2) Orderliness, 3) Dutifulness, 4) Achievement-Striving, 5) Cautiousness

2.4 Neuroticism
The term neurosis was used by Freud to describe a condition marked by distress, emotional suffering, and an inability to cope effectively with the routine demands of life. People who score high on Neuroticism may experience primarily one specific negative feeling such as: 1) Anxiety, 2) Anger, 3) Depression, 4) Immoderation, 5) Vulnerability, 6) Self Consciousness

2.5 Openness to Experience
Openness to Experience is considered as a dimension of cognitive style that distinguishes imaginative, creative people from down-to-earth, conventional people. Open people are intellectually curious and sensitive to beauty. Intellectuals usually score high on Openness to Experience; consequently, this factor has also been called Culture or Intellect. The people who mark high on this trait have the following attributes: 1) Imagination, 2) Artistic Interest, 3) Emotionality, 4) Adventurousness, 5) Intellect, 6) Liberalism.

In a research conducted on 308 undergraduates who had undergone the Five Factor Inventory Processes and provided their GPA came to a conclusion that conscientiousness and agreeableness, have a significant & positive relationship with all learning styles (synthesis analysis, methodical study, fact retention, and elaborative processing), whereas neuroticism has an inverse relationship with them all.

The Big Five together explained 14% of the variance in GPA, suggesting that personality traits make great contributions to academic performance. Furthermore, reflective learning styles (synthesis-analysis and elaborative processing) were able to mediate the relationship between openness and GPA (Komaraju, 2011).

As manual tracking was considered to be inappropriate hence an answer to it was searched in the form of Data mining. It was defined by Feelders et al. (2002) that data mining is a process of deriving information from large data sets through the use of algorithms and techniques.
drawn from the field of statistics, machine learning and data base management systems. Traditional data analysis methods often involve manual work and interpretation of data that is slow, expensive and highly subjective (Fayyad et al., 1996a).

The six important steps in the data mining process are: 1) Problem definition, 2) Acquisition of background knowledge, 3) Selection of data, 4) Pre-processing of data, 5) Analysis and interpretation 6) Reporting and use.

At each of these steps, we will look at important considerations as they relate to data mining in students databases comprising of psychological and learning aspects.

3 Challenges in Data mining based assessment
A domain expert involvement is important in interpreting the results of the data analysis. For example, a university might find a pattern indicating a direct relationship between agreeableness and 'accommodating' learning trend. The question then becomes, are some students agreeable because they are having higher 'accommodating' learning trends or they have 'accommodating' learning trend because they are agreeable. An expert in the area can take the relationship discovered and build upon it with additional information available in the university to help understand the cause and effect of the specific relationship identified.

Another challenge in mining data is dealing with the issues of missing or noisy data. Data quality may be insufficient if data is collected without any specific analysis in mind. This is especially true for psychological and learning trend information. Typically when student’s data is collected, the purpose is some kind of academic need such as evaluation of grades. The requirement of data for the required transaction is the only consideration in the type of data to collect. Future analysis needs and the value in the data collected is not usually considered. If the system administrator does not have control over data input then missing data may also be a problem. However, a data warehouse or data mart may help to prevent or systemized the handling of many of these problems. There are many types of algorithms in use in data mining. The choice of the algorithm depends on the intended use of the extracted knowledge.

The goals of data mining can be broken down into two main categories. Some applications seek to verify the hypothesis formulated by the user. The other main goal is the discovery or uncovering new patterns systematically (Fayyad et al., 1996). Within discovery, the data can be used to either predict future behavior or describe patterns in an understandable form. The following techniques have the potential to be applicable for data mining on student’s database:

Association Rule Mining (ARM) (Agrawal et al., 1993) is related to how items in a transactional database are grouped together. It is usually known as market basket analysis, because it can be likened to the analysis of items that are frequently put together in a basket by shoppers in a market. From a statistical point of view, it is a semiautomatic technique to discover correlations among a set of variables.

Clustering and classification is an example of a set of data mining techniques borrowed from classical statistical methods that can help describe patterns in information. Clustering seeks to identify a small set of exhaustive and mutual exclusive categories to describe the data that is present. This might be a useful application to student's data if you were trying to identify a certain set of students with consistent attributes. For example, a H.O.D may want to find out what the main categories of top performers are that students fall into with an eye towards tailoring various programs to the groups or further study of such groups. Single category may be more or less appropriate for one type of training program. A difficulty with clustering techniques is that no normative techniques are known that specify the correct number of clusters that should be formed. In addition, there exist many different logics that may be followed in forming the clusters. Therefore, the art of the analyst is critical.

In the same way classification (Han et al., 2011) is also a data mining technique that maps a data item into one of several pre-defined classes. Classification may be useful in a university to classify trends of movement through the semesters/trimester for certain sets of successful
students. A university is at an advantage when admitting new students if it can point out some realistic career paths for them. Being able to support those career paths with information reflecting student success can make this a strong resource for the university. Decision Tree Analysis (Kamber et al., 1997) also referred to as tree or hierarchical partitioning, is a somewhat related technique but follows a very different logic and can be rendered somewhat more automatic. Here, a variable is chosen first in such a way as to maximize the difference or contrast formed by splitting the data into two groups. One group comprises of all observations having a value higher on a certain value of the variable, such as the mean. Then the complement, namely those lower than that value, becomes the other group. Then each half can be subjected to successive further splits with possibly different variables becoming important to different halves. For example, students might first be split into two groups - above and below average CGPA (Cumulative Grade Point Aggregate). Then the statistics of the two groups can be compared and contrasted to gain insights about student’s analytical performance and reasoning abilities. A further split of the group having lower marks, say based on gender, may help prioritize those most likely to need special programs for enhancing analytical powers. Thus, clusters or categories can be formed by binary cuts, a kind of divide and conquer approach. In addition, the order of variables can be chosen differently to make the technique more flexible. For each group formed, summary statistics can be presented and compared. This technique is a rather pure form of data mining and can be performed in the absence of specific questions or issues. It might be applied as a way of seeking interesting questions about a very large data mart.

Regression and related models, also borrowed from classical statistics, permits estimating a linear function of independent variables that best explains or predicts a given dependent variable.

Exponential Smoothing (Holt et al., 2004) and details about future result prediction:

- The forecast for next period (i.e. \( t + 1 \)) is a weighted average of the actual and forecasted values of the time series in period \( t \).
- \( \hat{F}_{t+1} = w \times x_t + (1-w) \times \hat{F}_t \), where \( 0 < w < 1 \)
- Gives greater weight to recent data
- It is easy to update the forecasts as there is no need to re-estimate the equations
- When time trend is positive, forecasts are likely to be too low
- When time trend is negative, forecasts are likely to be too high

**OBJECTIVE OF STUDY**

The integration of psychology with academics has not been attempted using data mining but similar work has been done manually where it was concluded that conscientiousness & openness to experience are positively related to academic performance whereas extraversion is negatively related to academic performance, similarly gregariousness as negatively correlated to performance was concluded during the use of big model to analyze performance through self rating.

Business metrics does a great job summarizing the past. Hence attempt to predict future customer responses has been done by researchers with the help of data mining & predictive analytics, but much work has not been done towards predicting the scholastic achievements by relating it with psychological traits.

The objective of this research paper is to conduct psychological and academic assessment of students using various data mining techniques in order to be able to establish a relationship between psychological traits, learning trends & academic performances. This paper attempts to establish the value of data mining based assessment instead of the traditional manual assessment systems.

a) Assessment of learning trends & relating it with the academic performance.

b) Exploring the relationship between psychology and academic performance.

c) Prediction of future results on the basis of CGPA during any degree course.

d) Relationship, if any, between psychology and learning trends.

**METHODOLOGY**

The created database comprises of the academic results, results of the psychological tests and the acquired learning trend of every student. The sample size was of
200 comprising of engineering students specializing in Computer science, and Information Technology from the Noida region. A standard questionnaire was prepared by the researchers for the psychological and the learning trend analysis of these students. These two questionnaires were distributed to all these students and the results were compiled in order to enter them in the database into the desired format and were integrated with the respective academic results. The dataset may give an impression of being light in initial glance but the entire prepared database is very rich as there are 94 columns in every tuple. This means that for doing the overall analysis off a particular student we have to consider all the 93 attributes corresponding to him which makes it holistic and unique. Besides this there are various other tables pertaining to departments, psychology, login, etc which are synchronized by having common identification fields.

ANALYSIS

In universities, traditional student data analysis methods often involve manual work and interpretation of data that is slow, expensive and highly subjective. For example, if a faculty is interested in analyzing the overall performance, they might have to extract data from several different sources such as personal and academic records involving noting of grades secured in courses under various departments. That data is then combined, reconciled and evaluated.

This process creates many opportunities for errors. As databases grows in size (in this case, due to inclusion of complex learning characteristics and psychological assessment in the students table), the traditional approach becomes more impractical and infeasible. Data mining has been used successfully in many functional areas such as finance and marketing. Its applications in Universities on student’s database provide an as yet unexplored opportunity to apply data mining techniques.

Though most applications offer opportunities for generating ad-hoc or standardized reports from specific sets of data, the relationships between the data sets are rarely explored. This type of relationship is discovered by data mining.

This paper also supports in the use of data mining based assessments for universities & organizations. It proves through the analysis that it can be a great way to use relational database systems where data is stored in separate files that can be linked by common elements such as name or identification number. The relational database provides organizations with the ability to keep a virtually limitless amount of data on employees and the important thing is that they also have psychological assessment in the database which can be used as background material for our project. It also allows organizations to access the data in a variety of ways. While these calculations are helpful to quantify the value of some psychological traits, the bottom-line impact of them is not always so clear.

It is conducted in three steps as mentioned below:

a) Academic assessment of students.

b) Psychological assessments.

c) Identification of Learning trends.

The Data Loading was done through SQL Loader which is a high-speed data loading utility that loads data from external files into tables in an Oracle database.

(REFER FIGURE 1 HERE)

(REFER FIGURE 2 HERE)

(REFER FIGURE 3 HERE)

The assumptions that was made prior to this initiative were validated by actual implementation models designed using Java (Front-end) and Oracle (Back-end) on the students database.

The first validation was found that Continuous progress in one field say computer related courses is very much common. Our results also established unique and unprecedented relationships between Learning trends and Psychology. For example students who scored high on ‘agreeableness’ (a psychological function) were found to lie in ‘accommodating’ Learning trend group. Similarly, the students who had higher creative and artistic abilities (i.e. Ranked higher in ‘Openness To Experience’, another Psychological indicator) had ‘diverging’ Learning trend. Other relations that we had assumed were between Psychology and academics. The
upward sloping graph (directly-proportional) graph was obtained between conscientiousness levels and the CGPA scored by students and the inverse was true for neuroticism levels.

With the classification process, the students were categorized into three groups on the basis of the scores they got in Psychological tests. They were categorized using decision tree algorithm into below average, average, above average. Classification of the psychological traits can greatly help the university and the students to tailor the programs for improving the strengths and the subsequent eradication of the possible weaknesses.

The Assumptions of possible relation between some critical psychological traits and the academic achievements were proved by using data mining algorithms in strategic areas. Some of the most important revelations were inverse relationship between CGPA and Neuroticism Levels of a student and a direct relation between Conscientiousness Levels and the academic achievements. These lessons can be further applied in similar related psychological traits for judging and predicting and possibly improving the academic achievements in the respective areas. The evaluation of continuous increment, decrement or variations in SGPA was analyzed and was used to predict the future results.

The exponential smoothing used in the predictive analysis of semester results was found out to be more or less accurate. The value of variable ‘w’ was taken from 0.1 to 0.9 and our findings proved that the percentage error for ‘w’=0.7 was minimum. The predictive analysis can be very useful for the university to enhance the productivity of the students by being able to chop off the problems before they arise and also by taking into account all the shortcomings of the students in the various branches of the institution. Data-Mining has an important role to play in this field and on the basis of the results we have obtained, the roadmap for future improvements is quite clear. The performance of the tool can be greatly improved by acquiring insights from the outcomes and employing them in similar assessments in vast and complex fields of Psychology and Learning styles.

The following graphs show the classification of student’s data depending upon extraversion psychological traits scores:

(REFER FIGURE 4 HERE)
(REFER FIGURE 5 HERE)
(REFER FIGURE 6 HERE)

INTERPRETATION

There is powerful relationship between Psychology and Learning Style of the student. Psychology affects the academic performance and explains why students perform well continuously in certain subjects. The results obtained from implemented procedure clearly validate our assumptions.

The interpretations of the above can be explained by the results shown in following graphs:

(REFER FIGURE 7 HERE)

The figure 7 shows the relationship between learning trend and psychological trait. There is positive relationship between Diverging (learning trend) and Openness to Experience (psychological trait). The student who is more inclined towards Openness to Experience will have more diverging trend. These are positively correlated with each other.

(REFER FIGURE 8 HERE)

The figure 8 shows the relationship between learning trend and psychological trait. There is positive relationship between Converging (learning trend) and Friendliness (psychological trait). The student who is having more friendliness behaviour will have more converging trend. These are positively correlated with each other.

(REFER FIGURE 9 HERE)

The figure 9 shows the relationship between learning trend and psychological trait. There is positive relationship between Agreeableness (learning trend) and accommodating (psychological trait). The student who is more accommodating will have more agreeableness
trend. These are positively correlated with each other.

(REFER FIGURE 10 HERE)

The figure 10 shows the relationship between psychological trait and performance of the student. The performance of the student is measured in CGPA which is the outcome of the examination results. There is negative relationship between Neuroticism (psychological trait) and CGPA (performance). The student who is having more neuroticism psychological trait will score less in CGPA of examination results. These are negatively correlated with each other.

(REFER FIGURE 11 HERE)

The figure 11 shows the relationship between psychological trait and performance of the student. The performance of the student is measured in CGPA which is the outcome of the examination results. There is positive relationship between Conscientiousness (psychological trait) and CGPA (performance). The student who is having more conscientiousness psychological trait will score more in terms of CGPA in examination results. These are positively correlated with each other.

Thus, the results show that psychology affects the academic performance of the students and the progress of organizations

CONCLUSION

The present study predicts that Learning trends do have a relationship with Psychology & thus can have an impact on performance. The inverse relation between neuroticism & CGPA is an indication to the above fact. The academic performance depends upon psychological behavior and thus has a scope of improvement.

The integrated analysis of learning trends and psychology has led to evolution of unique and unprecedented results which can be utilized by the university and students for enhancing the productivity. The interesting and useful results using the data mining algorithms can lead towards enhancing the capability of University in providing education and will also greatly help in tailoring the environment where students can grasp the concepts in the manner in which they are suited.

REFERENCES


Figure 1: The initial data content in Spreadsheet format

Figure 2: Snapshot of Psychological Assessment Form
Figure 3: Snapshot of Learning Assessment Form

Figure 4: Snapshot of the classification on a Psychological parameter Extraversion

Figure 5: Zoomed view of the text content mentioned in figure 4