DEVELOPING ADVANCED INSTRUCTIVE FRAMEWORK AND PERSONALIZED E-LEARNING PLATFORM BY EMPLOYING INFORMATION MODEL LINKED TO MBTI IN PREDICTING STUDENT PERSONALITY AND PREFERENCES: LEARNING MANAGEMENT SYSTEM (LMS)

Rahul Garg

Bachelor of Engineering (B.E.), ECE

Thapar Institute of Engineering & Technology, Patiala- 147004 (Punjab), India

ABSTRACT

E-learning turn into a primary factor in the advanced instructive framework. In the present assorted understudy populace,E-learning must perceive the distinctions in understudy identities to make the learning procedure progressively customized. The target of this investigation is to make an information model to recognize both the understudy identity type and the overwhelming inclination dependent on the Myers-Briggs Type Indicator (MBTI) hypothesis. The proposed model uses information from understudy commitment with the learning the board framework (Moodle) and the interpersonal organization, Facebook. The model enables understudies to end up mindful of their identity, which thus makes them increasingly productive in their examination propensities. The model additionally gives necessary data to instructors, outfitting them with a superior comprehension of every understudy's identity. With this information, instructors will be progressively fit for coordinating understudies with their separate learning styles. The proposed model has associated on precedent data assembled from the Business College at the German school in Cairo, Egypt (240 understudies). The model was tried to go at using ten data mining request computations which were Naïve Bayes, BayesNet, Kstar, Random woods, J48, OneR, JRIP, KNN/IBK, RandomTree, and Decision Table. The results exhibited that OneR had the best precision dimension of 97.40%, trailed by Random backwoods 93.23% and J48 at 92.19%.

I. INTRODUCTION

E-learning (stands for all forms of web-based learning) uses computer and computer networks to create, deliver, manage and support online learning courses. E-learning's importance has increased in our education system. It offers courses and degrees to students through Learning Management Systems (LMS) like Moodle and blackboard that can be accessed at any time or place. However, e-learning also has its limitations. One of these limitations is the traditional "one-size–fits-all" learning model, in which the same learning resources are provided to all students. This is why most of the studies in recent years are focused on developing personalized e-learning platforms to overcome the one-size- fits-all approach. Each student is unique in their individual preferences, learning styles, their progress in the learning process and their individual background

(IJAER) 2018, Vol. No. 16, Issue No. II, August e-ISSN: 2231-5152, p-ISSN: 2454-1796

knowledge of the course material. Personalized e-learning platforms seek to improve student satisfaction and motivation throughout thecourse. To achieve this result, methods and techniques from several scientific fields and application areas are used. The most well-known techniques are: Data Mining, Knowledge Discovery, User Modelling, Web Mining, User Profiling, Artificial Intelligence and Agent Technologies, etc. These techniques utilize data generated from student engagement on all forms of web-based learning systems to provide a high degree of personalized education.

A. Learning Style

People have different learning styles, through which their cognitive, affective, personal and psychological traits interact together to formulate the way learners perceive, and interact with the learning environment. In the traditional classroom a competent instructor is able to differentiate and adapt learning strategies to the existing needs of the students. One way of doing so is to identify the individual learning styles and preferences of the learners. The instructor is guided by their own experience in the traditional classroom. Similarly, e- learning can also be supported by various models intended to measure learning styles based on different theories that can be applied to the e-learning system. For example, Meyers-Briggs Type Indicator (MBTI), Kolb's "experiential learning theory", and others.

B. The Myers-Briggs Type Indicator

Carl Jung's hypothesis of Psychological Types (1921) thought about that people are either introverts or extraverts, and their habits come from these common natural psychological types. Individuals take in and process information in different ways, usually based on their personality traits. Variation in an individual's behaviour is due to the essential differences in the manner of the persons rather than using their perception and judgment. Individuals take in information either by sensing or by intuition, and they organize this information either by thinking about them or by feeling. Myers- Briggs used and applied the bases of Carl Jung's theory of Psychological Types (1921) to create their personality inventory (The Myers-Briggs Type indicator).

Based on the Jungian's psychological types, Myers-Briggs evaluates personality types and preferences through four aspects of personality:

1. Extraversion (E) or Introversion (I)
2. Sensing (S) or Intuition (N)
3. Thinking (T) or Feeling (F)
4. Judging (J) or Perceiving (P)

The Myers-Briggs Type Indicator (MBTI) reports anindividual's preferences on four scales, which are shown in Table below:

Preferences	Definition
Extraversion or Introversion	Where an individual gets his/her energy
Sensing or Intuition	The way an individual takes in information
Thinking or Feeling	The way an individual makes decisions
Judging or Perceiving	How an individual deals with the external world

The combined sets of these different preferences provide 16 distinctive identity types and are regularly symbolized by four letters to speak to an individual's movement on the four scales. For example, ESTP stands for Extroversion, Sensing, Thinking, and Perceiving, which show the four preferences of highest occurrence for this person. The MBTI assessment highlights the distinct nature of each learner's preference.

C. Dominant Preferences

In addition to the sixteen personality types, each personality type has one dominant preference (dominant process) which is used with the highest confidence. It routes our personality and outlines our motives and goals. It is followed in descending order by the auxiliary, tertiary, and the inferior process. All of the four processes are located in the middle two pairs of the preferences which are Sensing or Intuition and Thinking or Feeling. If an individual's personality type has extraverted judging preferences, their dominant preference will be a judging one. Conversely, an individual with introverted judging preferences will have their dominant preference as a perceiving one. If an individual's personality type has extraverted perceiving preferences, their dominant preference will be a perceiving one. Conversely, an individual with introverted perceiving preferences will have their dominant preference as a judging one. A more detailed explanation for the process of determining the dominant preference was described in. Using these combinations can reduce the number of sixteen personality types to four. This has made it more manageable to match learning styles to each student's dominant preference, plan teaching approaches and to monitor learning engagement. Extroverts (E) utilize their general inclination for the most part for the outside world and introverts (I) use their overwhelming propensity generally for the internal world. Table below shows the needs and capacities for every one of the 16 identity types.For example, the dominant preference found in ISFP, INFP, ESFJ and ENFJ is feeling. To simplify the analysis, we will define four MBTI classes as shown in table below.

Myers Briggs type	Dominant preferences
ISTJ, ISFJ, ESTP ,ESFP	S
INFJ, INTJ, ENFP, ENTP	N
ISFP, INFP, ESFJ, ENFJ	F
ISTP, INTP, ESTJ, ENTJ	T

D. E-learning and Social Network

Recently, social networks have become an important part of our daily lives. Social networks connect individuals and groups of similar interests. The most popular Social networks are Facebook, Twitter, LinkedIn, MySpace,and Google+.Facebook has more than 845 million active users, and is a standout amongst the most prevalent social networking among college studies. Students usually use these social networks for their social chatting, debating public issues and reviewing coursework related material with colleagues. E-learning uses the social network as one of its tools to provide and improve the communication in the learning process. Instructors are able to create closed groups for each class to improve the communication between the instructor and students and among the students themselves. We selected Facebook for the present study as it is the most popular form of social networking amongst the student population.

E. Educational Data Mining

Data mining (DM) is a method of discovering useful information from large amounts of data through extracting patterns that helps in solving problems that face different sectors of society. As of late, there has been an expanding enthusiasm for the utilization of DM to examine the training segment. Instructive Data Mining (EDM) is worried about creating strategies and dissecting instructive substance to empower a superior comprehension of understudies' identities and inclinations, just as improving educating and learning forms.

The advancement in educational technology accompanied by the massive use of the internet has created a new well- known approach known as e-learning or web-based education, which generates a vast amount of data about learners and learning. All this information provides a gold mine of educational data. EDM uses this massive amount of data to better understand learning and learners, hence, developing different computational techniques that associate data with theories that turn the practice into something that can benefit the learners. In the past few years, EDM has been an emerging and highly interesting area of research to many researches along with its related research areas like e-learning and LMS. E-learning helps in providing guidance to learners and LMS provides the platform of communication, collaboration, administration and reporting tools. The log files and data stored in the databases of these systems are being manipulated and tested using DM techniques.

F. Data Classification

Classification is a form of supervised learning, where the computer is given labelled data and is asked to learn the relationship between the data and the labels. Data classification is normally done in two stages, the training stage and the testing stage. At the primary stage (likewise called the learning step or stage) a model that contains the diverse classes is produced through analysing a set of training instances. Each instance should have a place with a predefined class. In the second stage, the model is tried utilizing distinctive sets of data to have the capacity to decide the classification's exactness. If the model shows demonstrated precision, it very well may be utilized later on for arranging future data instances for which the class or information classification isn't known, thus, the model can be utilized as a data classifier in the decision-making process. Numerous techniques can be utilized for classification, for example, choice tree, Bayesian methods, decision tree, neural networks, and rule-based calculations. Below are three of the algorithms used in this research which have proved to achieve the highest accuracy:

- J48: is the standard Decision Tree induction algorithm for learning approach particularly pattern recognition and machine learning. J48 is based on the ID3 algorithm.

- OneR: is an accurate simple classification algorithm that creates one rule for each predictor in the data, and then matches it with the most accurate rules.

- Random Forest: is a group of classification or regression trees made from the random selection of samples of the training data. Prediction is made by grouping the forecast of the collection.

II. PROPOSED MODEL

A. The Data Model of Student Engagement from Adopting LMS and Social Networks Into E-learning

The primary goal of this examination is to enhance the quality of teaching and learning forms in the e-learning framework by methods for data and communication innovations.In this study, students had access to their online courses introduced over the LMS (Moodle). Some instructors introduce their courses on social network as a supplement in the learning process. due to the popularity of online social networks among college students. The data generated from the student's engagement with these tools was so vast and enriching. The data can help us create a model that predicts and classifies the student's personality and preferences using data mining tools, in which researchers can test our model with different classification methods.

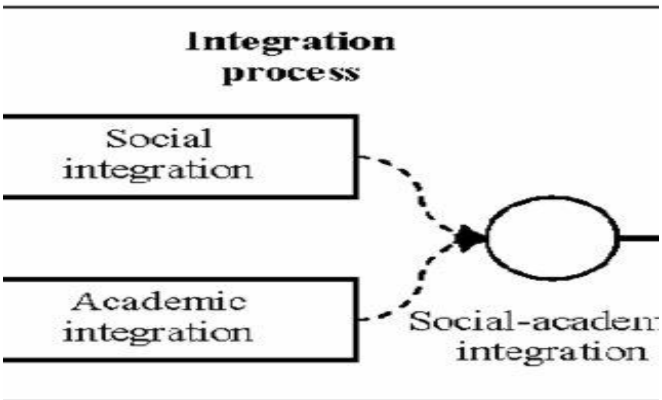

Fig 1: Integration Process

B. Data Collection

In this study, the information was gathered from the Business College of the German University in Cairo (GUC) that serves more than 1900 students in Egypt.The data was collected from the LMS (Moodle) provided by the college as well as Facebook groups created by the instructors for their students. The college provides fully online courses through LMS system (Moodle) which is an open-source LMS that help instructors create effective online learning communities. The data was collected for the academic year 2014-2015 from 240 students having 10 attributes as shown in Table III. Each of these students took about 6 courses per semester. Students were opening resources and doing activities related to these courses through Moodle and Facebook groups.

C. Data Pre-processing

Data Pre-processing is one of the important steps in DM. In this step, the dataset goes throw many stages as data cleaning, data integration, and data transformation. The gathered data was saved as Excel spreadsheets. The cleaning procedure is important so as to examine the information dependent on chosen classifier calculations, in which information with missing details are removed, incompatible data is remedied, exceptions are recognized, and copied information is evacuated.The data was represented by numbers and stored in the form of a CSV file so it can be imported to the data mining tool.

D. Modelling

The open source information data mining tool Waikato Environment for Knowledge Analysis (WEKA) was utilized for classification by giving inbuilt calculations that can be implemented on any dataset to help gauge the exactness of the resulting predictive model and visualize predictions, or the model itself.

Attributes	Description
N_VisitedPages	Number of pages visited on the system by students for the whole academic semester (Pages inside Moodle)
Total_TimeSpent	Time spent visiting different pages, chatting and doing other activities
N_Comments	Number of comments written by students in comment blocks in different courses pages
N_Likes	Number of likes by students on posts in different courses pages
N_Shares	Number of shared posts by students in different courses pages
N_Posts	Number of posts written by a student in Blogs blocks in different courses pages
N_ChatSessions	The Number of chats the user joined
N_Early Assignment Submission	Number student Early Assignments Submissions during the semester by instructor observation
N_Late Assignment Submission	Number student Late Assignments Submissions during the semester by instructor observation
MBTI Type	student Myers-Briggs Type Indicator (MBTI) 16 types
Dominant Preferences	student Dominant preferences (class S, class F, class N, class T)

E. Proposed Model Flowchart

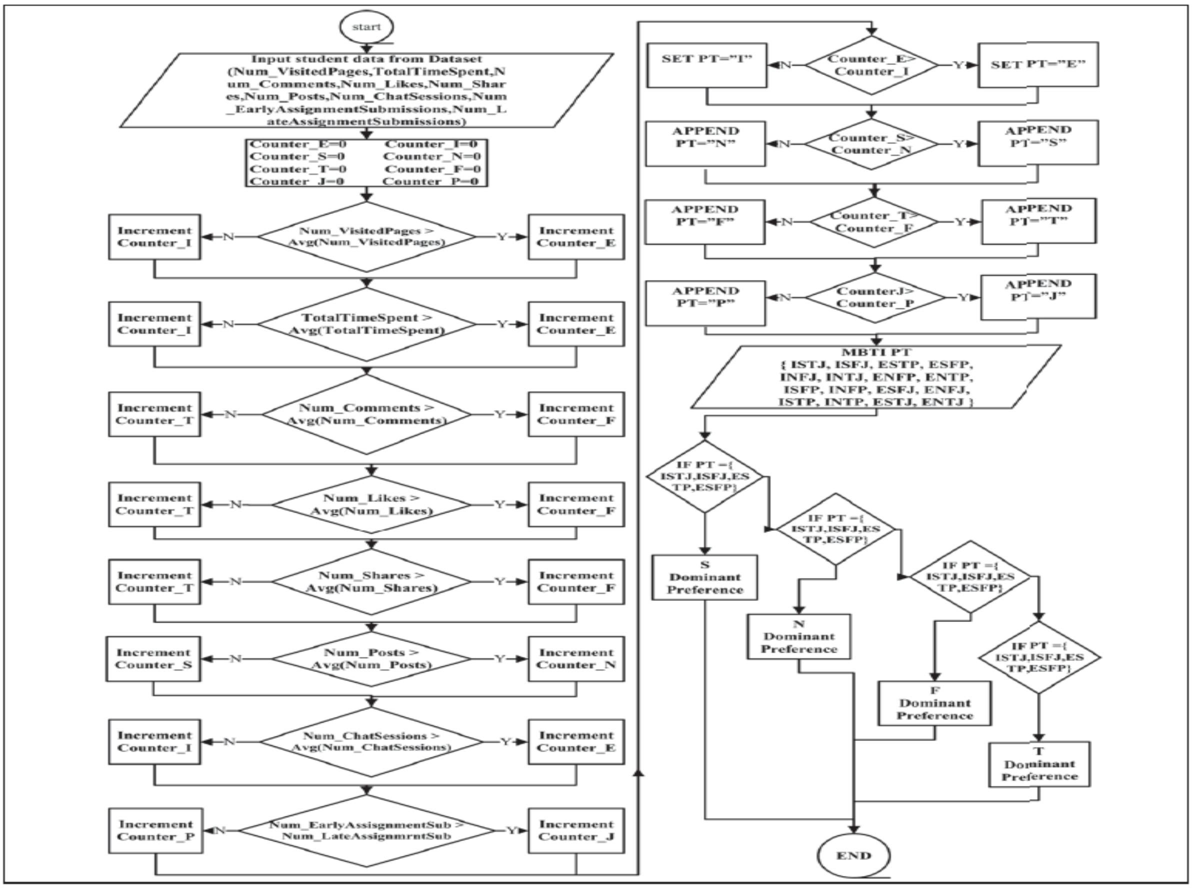

Fig 2: Flowchart

The flowchartshows how the data model predicts the student MBTI personality type (PT) and dominantpreference type (DP) via the behaviour of the student over the LMS (Moodle) and the social network (Facebook).

The first phase is to predict student PT; we defined and initialized our counter of 8 preferences measured in the MBTI model. Then each student behaviour measures one of the four preferences scales of the MBTI (i.e. Extraversion/Introversion, Sensing/Intuition, Thinking/Feeling, and Judging/Perceiving). In the first seven student behaviours (i.e. visited pages, time spent, comments, likes, shares, posts, and chat sessions). If the number of a certain behaviour is greater than the average number (avg) of this behaviour, the value of the corresponding preference counter will be incremented. The last two attributes (i.e. Number of Early Assignments Submissions, Number of Late Assignments Submissions) will be compared and the highest will increment the value of the corresponding preference counter. After all student behaviours have been measured, we will compare the counters to find the greater ofeach of the four Jungian psychological types preferences. The greater preference counter of each of the four Jungian psychological type's preferences will be represented by one letter in the student PT.

The second phase is to predict student DP which is based on the student PT; every student personality type has one of 4 dominant preferences.

S = {ESFP, ISFJ, ESTP, ISTJ} N = {ENTP, INFJ, ENFP, INTJ } F = {ENFJ, ISFP, ESFJ, INFP} T = {ENTJ, ISTP,ESTJ, INTP } So, the dominant preference is assigned according to the predicted student PT. For example, if student PT is one found in either “ISTJ”, “ISFJ”, “ESTP”or ESFP, then the student DP is “Sensing”.

III. EXPERIMENTAL RESULTS

The experiment used 10 classification algorithms in order to determine the highest classification accuracy. The algorithms were Naïve Bayes, Kstar, J48, OneR, Random forest, KNN/IBK, RandomTree,BayesNet, JRIP, and Decision Table.

In order to measure the model performance, the dataset was split into two groups: the training set (20%) which was utilized to train the model, and the test set (80%) used to test the model in order to measure its accuracy, precision, recall, F-value, true positives (TP) rate, and false positives (FP) rate, which are defined by equations by means of a confusion matrix.

-Accuracy: is defined as the ratio of correctly classified instances which is equal to the sum of TP and TN to total number of instances.

$$\text{Accuracy}=\frac{(TP+TN)}{(TP+FN+FP+TN)} \quad \text{---(1)}$$

- Recall: is determined as absolute number of genuine positives divided by absolute number of genuine positives + absolute number of false negatives.

$$\text{Recall}=\frac{TP}{(TP+FN)} \quad \text{---(2)}$$

- Precision: is calculated as absolute number of genuine positives divided by absolute number of true positives + total number of false positives.

$$\text{Precision}=\frac{TP}{(TP+FP)} \quad \text{---(3)}$$

- True-positive rate (TP Rate) value: is the percentage of positive instances correctly classified.

$$\text{TP Rate}=\frac{TP}{(TP+FN)} \quad \text{---(4)}$$

- Positive rate (FP Rate) value: is the percentage of negative instances misclassified.

$$\text{FP Rate}=\frac{\text{FP}}{(\text{TN+FP})} \quad \text{---(5)}$$

- F-Measure: is a combination of recall and precision. It is also defined as harmonic mean of precision and recall.

$$\text{F-Measure}=\frac{2*\text{Recall}*\text{Precision}}{\text{Recall} + \text{Precision}} \quad \text{---(6)}$$

0.976, OneR and Random forest had the highest value of Recall at 0.974 and 0.933, respectively, and OneR had the highest value of F-Measure at 0.974.