CHAPTER 4

MATERIALS AND METHODS
4.0 MATERIALS AND METHODS

4.1 EXPERIMENTAL DESIGN

The Experimental Design consists of five phases as shown in Fig. 4.1. The proposed Multi-Faculty Model supports a larger audience, target groups, and multi-faculty environment with online collaborative features. The following are the important core phases considered for performing the evaluation and assessment of the e-Learning set-ups in colleges and universities.

1. Collect responses from students and faculties - a Self-Regulation Survey.
2. Performing Sentiment Analysis on Student Feedback.
3. Performing Web Usage Analysis using Google Analytics.
4. Collect Survey responses from students on Social Networking behavior.
5. Incorporate Multi-Faculty roles.
6. Comparison of e-Learning Frameworks

![Experimental Design Diagram]

**Figure 4.1: Experimental Design**

4.1.1 EXPLORATORY FACTOR ANALYSIS

In Phase I, it was planned to conduct a Survey to identify important self-regulated learning strategies among undergraduate students using the Self-
Regulation Questionnaire. This instrument was based on Brown et al. (1999). An Exploratory Factor Analysis will be done on the survey responses and the latent factors will be identified as an output of the Principal Component Analysis that will be done using the Software IBM SPSS Statistics. The outcome of this study is the identification of important factors that guide students to adopt self-regulated learning strategies.

4.1.2 SENTIMENT ANALYSIS

In Phase II, the student feedback received from the pilot website was analyzed to understand the sentiments and needs of the students who were using the website for their studies. The output of the Sentiment Analysis will be N-Grams and Sentiment Trees.

4.1.3 WEB USAGE ANALYSIS

In Phase III, the web analytic tool "Google Analytics" was used to analyze the visual reports on learner behavior like the metrics Sessions, Page Views, Duration of session etc., for critical evaluation.

4.1.4 SOCIAL NETWORK ANALYSIS

In Phase IV, a two-stage survey was conducted to understand, how the young university students, use the Social Networking Site LinkedIn and the responses were collected on the five points Likert Scale. An Exploratory Factor Analysis was conducted on the survey responses to identify the hidden factors associated with the measurement items in the data set. Subsequently, a theoretical model was constructed using Structural Equation Modeling principles, depicting the interrelationships between the latent constructs and measured variables constituting a measurement model and a structural model. Finally, a Confirmatory Factor Analysis was done to prove how well the model fits into the theory. The
software IBM SPSS, and AMOS 23 were used to perform the multivariate statistical analysis on the LinkedIn Survey response items.

4.1.5 DESIGN OF THEORETICAL MODEL

A theoretical model has been constructed based on all factors that contribute to enhancing the effectiveness of e-Learning Framework using Structural Equation Modeling Principles (SEM) as shown in Fig. 4.2. This model explains how this research work addresses the problems faced in e-Learning settings today by identifying the key factors that can enable student engagement and improvements in the learning outcome of online students. These factors help the stakeholders, the institutions, faculties, and students work in unison toward making the success of an e-Learning system. Based on the theoretical model represented in Fig. 4.2, the following hypotheses are framed and this study aims to test and prove the validity of the model.

H1: Sentiment Analysis of student opinions positively influences student empowerment.

H2: Instilling Self-Regulation in online students positively impact student empowerment.

H3: Social Networking by students has a positive impact on student empowerment.

H4: Faculty participation in Instruction Design has a significant positive impact on faculty empowerment.
H5: Web usage analysis has a positive impact on faculty empowerment.
H6: Multi-Faculty Model implementation has a positive impact on faculty empowerment.
H7: Student Empowerment has a positive impact on student engagement.
H8: Faculty empowerment has a positive impact on student engagement.
H9: Student engagement has a positive impact on student learning outcome.

4.2 EXPERIMENTAL PROTOTYPE-I

The initial experimental prototype was constructed based on the Single-Faculty model and its features are shown in Fig. 3.4. The technology used for the prototypical website was Google Site Web Application. This choice was selected as Google Sites is a popular and a free tool used for building websites easily and one can to design content and deploy the same in the website very fast in an
intuitive way. A faculty can upload the course material for the course he or she handles in Google Sites without writing any code. The interactive features of Google Sites help one to build a course website in an effective and efficient way. The functionality of the website can be enhanced further by adding other Google applications in Google Sites. Google Sites allows the integration of other powerful Google products like Google Docs, Google Forms, and Google Analytics. Once the course website is deployed and brought to use in practice, the next important question to be answered is whether the website satisfies the learners’ needs. The faculty who has designed the website is empowered if he or she is able to get feedback from students who have been using the website.

A Guestbook for feedback has been added to the website to collect the feedback on the actual usage from students using Google Forms. The information that has been collected as feedback from students is stored in spreadsheets. The data can be used later for analysis and decision making by the faculties. This data is useful and provide enough information to faculties regarding the students’ satisfaction, and usefulness, as well as any problems faced by them. By performing the sentiment analysis of student comments available in text form, faculties can easily judge the usefulness of the web content delivered to the students. Faculties can improve the website on a continual basis based on various inputs from student sentiments.

The next activity is to get the report on web usage statistics which gives the behavior patterns of website usage by students. This is achieved in practice by integrating the web analytic tool Google Analytics with Google Sites. Google Analytics automatically scans weblogs on a continual basis and generates many summaries of web usage statistics in visual form. Many useful metrics on site usage helps the faculties to know how the site performs. Some of the useful metrics are daily visitor count, time on site for each session, pages viewed for each
session etc. By periodically going through this information, faculties will be able to improve the website content so that it is more current and satisfies the learners’ needs.

The limitations of a single faculty model are its general nature, lack of support for multiple faculties, and student groups. The tools used are inadequate to know and monitor the progress of the individual learner. Also, this model is not scalable. It is difficult to include online collaborative features like forums Q&A and Assessment. However, some of the merits of the model are its ability in making the faculties as facilitators of learning, and their empowerment. Faculties can identify the academic needs of the students as well as recognize the voice of the students. This forms the macro level monitoring of students.

4.3 EXPERIMENTAL PROTOTYPE-II

The experimental prototype-II is based on the Multi-Faculty model and Fig. 4.3 shows how a Multi-Faculty Model is evolved.

The first major improvement in this model is the provision for faculty registration on the course website. The second improvement is the incorporation of scalability of students and staff that can be sustained. This the most important consideration for the Prototype-2. The third is the ability for meaningful and timely interventions. Weak students can be monitored by providing motivation through emails and chats. Measures can be taken to improve engagement and inhibit drop outs.

Fig 4.4 shows the features of the multi-faculty model which forms the basis for building the experimental prototype-II. The prototype-II has been constructed in order to implement experimental verification of impact of online learning on student learning outcomes. It also enables to study the impact of the usage of the
collaborative tools like Discussion forums, Question and Answering by the faculties and the students.

![Figure 4.3: Evolution of Multi-Faculty E-Learning Model](image)

Fig. 4.5 shows the architecture of the Hardware and Software components constituting the physical implementation of the prototype-II.
Figure 4.4: Experimental Prototype-II based on Multi-Faculty Model

Figure 4.5: Architecture of Prototype-2
The webserver was built using Apache Tomcat and the scripting languages used are JSP, Servlets, and PHP. The database which was used in the Prototype-2 was MySql.

![Web Application Development Environment](image.png)

**Figure 4.6: Web Application Development Environment**

Fig. 4.6 shows web application development environment which facilitates the web developers and admins to develop collaborative and interactive web applications that can promote facilitate Faculty-Student interaction and may result in enhancements in student learning outcome and retention.

### 4.4 STATISTICAL TOOLS USED

Experimental research entails the manipulation of an independent variable(s) and observing its impact on a dependent variable(s). The importance of experimental research is that it helps a researcher to explore and identify a cause and effect relationship between variables by manipulating the independent variable(s). An experiment may involve many variables that can be measured and some variables can be manipulated and some can be controlled. In an experiment, an independent variable can be manipulated in order to see its impact on
a dependent variable. An independent variable is also termed as a predictor variable and a dependent variable is also called an outcome variable.

4.4.1 DESCRIPTIVE STATISTICS

Descriptive statistics is the type of statistics which is used to describe the dataset. It is used to describe the characteristics of data and to present the data in summary form. Descriptive statistics are generally used to determine if the sample is normally distributed. It is presented using frequency distributions, charts, and tables. It is generally reported as a measure of central tendency.

- **Central Tendency**
  - *Mean* – also known as the average
  - *Median* – the center most value of the given data set
  - *Mode* – The value which appears most frequently in the given data set

- **Statistical Dispersion**
  - *Range* – Range gives us the understanding of how spread out the given data is.
  - *Variance* – It gives us an understanding of how far the measurements are from the mean.
  - *Standard deviation* – Square root of variance is standard deviation, also the measurement of how far the data deviate from the mean

- **Measure of Asymmetry or flatness**
  - *Skewness* – It is the measure of the asymmetry of the distribution of a variable about its mean
  - *Kurtosis* – It is the measure of the “tailedness” of the distribution of a variable. It gives us the understanding of how closely the data is spread out.

A different measure of central tendency is used depending on what exactly one is trying to describe. Mean and median can only be associated with numerical data.
The mode can be associated both with nominal and numerical data. Descriptive statistics is extremely useful in examining the given data. The succinct reports with summarized information provided by Descriptive statistics help one to comprehend the data completely.

4.4.2 INFERENTIAL STATISTICS

The research study involves the collection of experimental data. Inferential statistics facilitates a researcher in making inferences from one’s data (A sample) that can be applied to a higher level (The population). Descriptive statistics summarize and present the data content. However, a researcher needs more powerful tools such as inferential statistics, in order to make inferences that extend beyond immediate data under consideration. One can arrive at how the population will look like by using inferential statistics that facilitates one to draw conclusions from the sample data. Hypothesis testing and point estimation are two concepts of inferential statistics that aid a researcher in making an inference on the characteristics of a population from the samples.

A frequent goal of collecting data in a research study is to allow inferences to be drawn about a population from a sample. Inferential statistics provide the basis for drawing conclusions that go beyond the observed data in such cases. A statistical inference of the likely occurrence of an observed effect is not attributable to chance alone. In contrast to descriptive statistics that simply presents a summary of one’s data, inferential statistics is capable of making important predictions about the world beyond the data. A general linear model forms the basis for most of the frequently reported inferential statistics. The important examples are t-tests, ANOVA (analysis of variance), linear regression analysis, and factor analysis. The concept of sampling and probability forms the foundations for Inferential Statistics. The problem to be overcome in conducting
research is that data are typically collected from a sample taken from a larger population of interest. Inferences drawn from the sample data cannot be conclusive and are not valid if the sample data is not found to be representative of an associated population.

4.4.3 ANALYSIS OF VARIANCE

ANOVA stands for Analysis of Variance. ANOVA represents an assemblage of the multivariate statistical methods that are used to infer whether there are noticeable variations exists between the means of three or more groups or variables in a population. ANOVA is performed by conducting an experimental analysis on sample data. For example, the one-way ANOVA can be performed to understand whether sales performance differed based on the different marketing strategies adopted, like the direct or online sale of a product. The customers are divided into two groups such as online and direct sales. However, in spite of the one-way ANOVA being an all-inclusive test statistic, one cannot clearly ascertain which particular group differs from another group. The ANOVA reports that the two groups were just different. However, a post hoc test reports the group-wise comparison of means and it can be inferred which group contributes much variance.

A one-way ANOVA using SPSS needs that the following six conditions are met:

- Condition #1: The dependent variable should be continuous and represents an interval or a ratio.
- Condition #2: The groups representing the independent variable should be disjoint.
- Condition #3: The observations made on each group are independent.
- Condition #4: Large outliers must be discarded.
- Condition #5: The data values of groups should be normally distributed.
- Condition #6: The homogeneity of variances has to be ensured.
4.4.4 STRUCTURAL EQUATION MODELING (SEM)

Structural equation modeling (SEM) is a multivariate statistical analytic method. SEM is used to analyze structural relationships between variables. This method consists of the combination factor analysis and multiple regression analysis. SEM analyzes the structural relationship between measured and latent variables. Researchers use this method because it uses a single analysis in estimating the multiple and interrelated dependencies. The variables used in SEM are of two types namely endogenous and exogenous. Endogenous variables represent dependent variables and exogenous variables represent the independent variables.

SEM consists of constructing two types of theoretical models, namely, a “Measurement Model” and a “Structural Model”.

4.4.4.1 MEASUREMENT MODEL

The measurement model is based on measured variables and it depicts how the variables represent the theory. The measurement model depicts how the constructs are represented by the observed variables. It represents the theory by specifying the observed variables for each construct. It ensures the evaluation of the construct validity. The observed variables are measured by the researcher. Fig. 4.7 shows how SEM represents the Structural Model and Measurement model. The variables used in SEM are called indicators and factors. These are also known as measured and latent variables. The factors represent the underlying constructs that have been discovered in the factor analysis. Fig. 4.8 shows the Measurement model.
The measurement model incorporates the following factor equations that specify the relationships of variables:

\[
\begin{align*}
x_1 &= \lambda_{11} \xi_1 + \delta_1 \\
x_2 &= \lambda_{21} \xi_1 + \delta_2 \\
x_3 &= \lambda_{31} \xi_1 + \delta_3 \\
x_4 &= \lambda_{42} \xi_2 + \delta_4 \\
x_5 &= \lambda_{52} \xi_2 + \delta_5 \\
x_6 &= \lambda_{62} \xi_2 + \delta_6 \\
\end{align*}
\]

\(\delta_i\) is the error variable which is the unique variable influencing \(x_i\). \(\lambda_{ij}\) is the loading of the indicator items.
4.4.4.2 STRUCTURAL MODEL

The Structural Model consists of constructs that represent the theory. It also depicts how constructs are dependent on each other. The following are the assumptions made in model construction:
1. Normal distribution of Multivariate.
2. Endogenous and exogenous variables exhibit a linear relationship.
3. Outliers are removed from Data.
4. Endogenous and exogenous variables exhibit cause and effect relationships.
5. Observed covariance must be true.
6. Samples can be in the range 200-400 sample size with 10-15 indicators (or 10-20 times the variables count).

4.4.4.3 SEM IMPLEMENTATION PHASES

1. The analysis begins with the definition of constructs which should be based on a theory. Pre-test has to be performed for evaluation of the item. The validity of the measurement model is ensured by performing a Confirmatory Factor Analysis.
2. Path analysis is adopted in constructing the measurement model by establishing proper relationships between exogenous and endogenous variables.
3. Model specification based on the expected empirical results.
4. Assessment of the validity of measurement model: CFA is used to assess the measurement model and the constructs' validity has to be ensured.
5. Specification of the structural model.
6. In the final step, a researcher validates the structural model’s validity. A model fits well if a chi-square test reports a low value. A number of Model Fit Indices are reported by the software. Model fit indices form two categories in
SEM, namely, incremental fit index (such as CFI, GFI, TLI, AGFI) and badness of fit index (such as RMR, RMSEA, SRMR). The best model fit can be ensured if any one of the above indices is within limits as per standard criteria.

4.4.5 STATISTICAL SOFTWARE USED

1) SPSS Statistics was used for the univariate ANOVA Analysis and principal component analysis.

2) IBM AMOS was used to perform the Confirmatory Factor Analysis

4.5 DATA COLLECTION AND DEMOGRAPHICS OF PARTICIPANTS

Annexure 1-J briefs how data was collected and the demographics of the participants for a) Self-Regulation Survey, b) Social Networking Survey, c) Multi-Faculty Model, d) Web Usage Pattern Analysis, and e) MOOC performance Dataset.

Chapter 4 presented the Experimental Design, development of Questionnaire and Instruments and Analysis tools used. The phases of experimental design are SRL Survey, and Exploratory Factor Analysis, Sentiment analysis techniques and tools used, and Web usage pattern analysis, Social Networking survey, and validation of the Multi-Faculty Model. This chapter also presented the overview of the statistical tools and the software packages used in the experimentation. The main statistical methods used in this research work are the ANOVA, Factor Analysis, and Structural Equation Modeling. The software packages used are IBM SPSS Statistics Version 24, and IBM SPSS AMOS.