Chapter 4: Identification of Stakeholders and Classification of Learners

An important step towards establishing an e-learning environment is to identify the stakeholders, define their roles, responsibilities and respective positions in the organizational hierarchy. The roles can be viewed on an individual level or on a collaborative team basis, where the team comprises experts from various fields of specialization. The interaction among these stakeholders involves the flow of knowledge, from those who possess it to the ones who do not. This is referred to as Socialization phase in the KMeLS model. This chapter focuses on two aspects of Socialization, namely, identification of stakeholders, their roles and stakeholder interaction analysis.

4.1 Identification of Stakeholder and their Roles

An e-learning environment can be viewed as a three-tiered structure from the perspectives of various stakeholders working for the system, illustrated in Figure 4.1. The user view forms the topmost layer, followed by the facilitator view in the middle and the software developer view at the bottom. End users (Student, Instructor, Content Provider, and Administrator) lie in the topmost layer of the e-learning environment. The facilitators (managers or high authorities), the decision-making authorities of the organization, are positioned in the middle layer. Software developers responsible for implementation of the system according to the policies designed by the middle layer stakeholders lie in the bottom layer of the e-learning environment. They develop, implement and create the virtual environment for efficient and collaborative learning. The three views describing the roles, responsibilities and interaction among the stakeholders are discussed next.
4.1.1 User View

The user view forms the topmost layer. It is the interface between the user and the system. Student, instructor, content provider, and administrator are the most prominent end users of the system. The roles of these stakeholders are described as follows.

*Student:* A student evolves through the role of an information seeker to a learner and finally an examinee as shown in Figure 4.2. The first and foremost role of a student is that of an information seeker when she/he familiarizes with the system by gaining information about the course instructor or the institute. The student navigates around the system to select the best option for choosing an instructor, course etc. After enrollment, the student becomes an integral part of the system, thus acquiring the role of a learner. In later stages, the learner is evaluated to assess the depth of understanding acquired by her/him, and the role changes to that of an examinee. To
get a proper feedback from the system regarding her/his performance, the student again transforms into an information seeker.

*Instructor:* An instructor plays the role of a tutor, an evaluator/assessor and above all, a counselor as illustrated in Figure 4.3. The instructor takes over the role of a tutor while teaching and interacting with the student (learner) for the concerned subject. On course completion, the instructor turns into an evaluator to examine the knowledge gained by the student. Another important role played by the instructor is that of a counselor where the instructor delves into the social connectedness (Chang, 2004) of the student, essential for her/his holistic development.
In a traditional classroom based education system, the teacher is required to play all three roles simultaneously because of face-to-face interaction. She/he assesses the student continuously to keep an eye on her/his understanding capabilities. If there seems a decrement, the teacher needs to change the teaching style or pace to enhance the performance of the student. Side by side, the teacher must provide a psychological support to a student if she/he is unable to cope up with the stress and anxieties of the learning environment. Thus, the teacher takes over the capabilities of an evaluator, and an evaluator further takes over the capabilities of a counselor. Performing all these roles for different students at the same time can be demanding for a human instructor. The situation is more beneficial in the e-learning environment because it is an integrated environment that provides an option for the student to interact with all the experts in respective domains while accessing the system. The interaction between different roles of an instructor and a student is shown in Figure 4.3. In this system when a student enrolls as a learner, the instructor is present as a tutor. Being a tutor, the instructor uses the already designed and constructed content (by the content provider) from the system’s knowledge base and using either blended or synchronous approach, instructs the student in order to help her/him learn the subject.

Content provider: An e-content provider is the subject matter expert who puts the content in interesting and easily understandable format so that the learner grasps reliable and correct concepts. Chang (2004) establishes that contents designed by subject matter experts have a better impact on students rather than those designed by software experts who may not be familiar with the subject at all. It is, therefore, imperative that the software system for e-learning must be able to provide the basic assistance to the content provider through convenient and user friendly tools and interface.

While it is true that the e-learning tools and technologies, discussed in Chapter 2, can be applied effectively by a person or a team of non-experts, it is also true that trained and experienced Human Computer Interface (HCI) professionals will provide better results using the same techniques (Feldstein and Neal, 2006). Therefore, to design and develop a courseware, a content provider team should comprise subject
matter experts, subject instructors, software developers, graphic designers, multimedia specialists, usability consultants and pedagogy experts.

**Administrator:** The administrator holds the responsibility for managing the complete virtual system by allowing or restricting the access from any user. Depending on the size of institute or organization this role may be played by one person or a group of people. Another major responsibility of the administrator is to take care of all financial matters and legal issues.

Like any traditional institute or organization, an e-learning environment also comprises ‘behind the scene’ stakeholders who provide non academic support, by working in the physical part of the virtual system for the smooth function of the entire system. They are the indirect stakeholders defining the Facilitator and Software Developer views of the system respectively.

### 4.1.2 Facilitator View

The facilitator view comprises the managers, facilitators (Hootstein, 2002) or high authorities who are the decision makers of the system. They are responsible for setting the guidelines and maintaining quality in the e-learning environment. Their responsibilities include:

- defining and forming the structure of the virtual and physical systems for the organization/institution by appointing other stakeholders.
- framing policies that specify the responsibilities and terms and conditions of contract with each user.
- maintaining quality in e-learning. More quality standards may be added here to the existing standards for quality in e-learning.
- designing the courseware, the content types (ways of describing the subject). The underlying theme and template of the design need to be uniform throughout so that the users get consistent look and feel for all subjects.
- providing full support in conveying user and system requirements to the software developers.
- continuous/regular involvement during system development to ensure quality.
- verification of the implemented working system to ensure its suitability to the requirement.
- provision of adequate training to the users for their quick and easy adaptability to the system.
- continuous monitoring and enhancements of the system to keep it updated and error free.

### 4.1.3 Software Developer View

Software developers lie in the last layer of the model and form the base of the system. They realize virtual environment by scaffolding quality learning with technology. They interact with facilitators, software developers and end users during various phases of software development to elicit their requirements and expectations with the prospective system. The work of the software developer is broadly divided into six phases.

**Analysis:** Software developers analyze the requirements set by the authorities. This involves extraction of information from the stakeholders required to personalize the user interface. The developer should be able to elicit both the expressed as well as abstract requirements of the stakeholders. This enhances the adaptivity of the e-learning system.

**Architectural designing (high and low level):** The system should be robust enough to solve the queries of the users, thus instilling in them, confidence about the credibility of the system. The information gathered from the analysis phase is converted to a technical architectural design. This includes designing user interfaces and knowledge databases, enhancing usability and reusability of components, solving security issues, providing online support, determining students’ learning path, providing required help to all stakeholders, designing evaluation system, student support system and complete integration of the overall system.

**Development:** The developers code the design using various programming...
languages. The design and development can be carried out in two ways: reuse components or designing from scratch. The latter of the two is more time consuming but can prove to be more effective whereas the former approach needs designing and development of interfaces to integrate the components. Learning through web requires proper planning and deployment strategies in place, therefore efficient use of web services and web communication is required during system development. Usability is also an important issue while developing a web-based system (Banati, Bedi and Grover, 2006). A user-friendly system enhances user affordance to the system and in turn boosts the confidence of the user on the system.

Testing: Regressive testing must be performed on the software developed for the e-learning environment as faults not only de-motivate the learners and instructors both, but may also lead to open security issues.

Implementation and Deployment: This includes putting everything in working order as required by the stakeholders. The system is uploaded on the web in ready to use condition.

Training: Sufficient hands-on training must be provided to users to get comfortable with the system. A complete online help for assisting and troubleshooting must be available to the users.

This chapter focuses on learners and instructors, the two key stakeholders in an e-learning system. Further, learners lie at the centre of the whole system having their future at stake. They enroll into the system with an underlying hope to gain knowledge and skills. Keeping these facts in mind, it would be beneficial to, somehow, predict the future performance of the learners, so that content is presented to them in a customized manner. The next section discusses the studies conducted by researchers in the field of learner-instructor interaction.

4.2 Related Work

Several research studies have been conducted, over the years, (Vrasidas and McIsaac, 1999; Moore, 2001; Jung et al, 2002; Wilson and Stacey, 2003; Koehler et al, 2004;
Wanstreet, 2006; Shackelford and Maxwell, 2012) that have emphasized on the value of interaction among learners and instructors for enhancing learner satisfaction in an online learning environment.

Moore (2001) rendered use of interactive medium pointless for instructors if they do not structure and encourage interaction. Yet having the opportunity for interaction is not enough; the effort must be seen as leading to a valuable outcome (Anderson, DuPlessis and Nickel, 2001).

Learner satisfaction would be improved if online instructors are aware of the expectations and desires of their students. In a study conducted by Dennen, Darabi and Smith (2007), learners rated items focused on communication needs and being treated as individuals as the most important. It was also reflected that the learners considered timeliness of feedback more significant in comparison to the extent of feedback.

The findings of a study conducted by Sher (2009) suggest the use of technology as a great facilitator. It not only enhances interaction among students, but also bridges the gap between instructor and students. He also opines that both student-student and student-instructor interactions are significant contributors to the level of student learning and satisfaction in a technology-mediated environment.

The results of Bernard et al (2009) confirm the importance of student–student, student-content and student-instructor interaction for student learning. Abrami et al (2011) discussed the importance of interaction among instructors and learners to understand proper structuring of knowledge tools. They also stressed on need of guidance to learners regarding the suitability and applicability of tools.

The research findings by Kiriakidis (2008) lay stress on the effects of interactions between students and the instructor. The data suggest that instructors should initiate, monitor, guide and frequently participate in online discussions. The study also points out that in online learning environment, teaching presence is created with frequent instructor discourse and social presence is created with frequent instructor and student discourse.
These studies highlight the significance of learner-instructor interaction in the field of learning, especially online learning environments. Thus, it can be summarized that an e-learning environment must provide opportunities of interaction among learners and instructors for inculcating critical thinking and fostering satisfaction among the learners.

The present study explores the realm of learner-instructor interaction further by taking into account, the preferences of the learners with respect to interaction and analyzes them to provide them content in appropriate way. The next section discusses the significance of interaction in an online learning environment and identifies important factors affecting the knowledge gain of a learner.

4.3 Learner Interaction

The current online learning scenario stresses on the need of motivational, customized and student-friendly social learning environment. Although e-learning provides the flexibility of anywhere, anytime learning yet it increases the risk of dropouts during the course schedule. The course design should encourage the students to enroll, stay and complete the course in the stipulated time period. It is therefore imperative to understand the students and predict their capabilities. This task is intricate as not much data is available when a student enrolls in an online course. Also, the progress report of a learner, which reflects how much knowledge he/she has gained while pursuing a particular course, is available only on completion of the course. Thus, customizing a course as per the capabilities, right at the onset of course is a pertinent task.

This section lays emphasis on understanding the learning capabilities of a student before the commencement of a course to make the learning simpler and an enjoyable task. The proposed method predicts the knowledge gain of learners, on the basis of their present online interaction behaviour and regularity in attending online sessions.

The learners who are predicted to pass the course are termed as gainers and those who may not are referred to as non gainers respectively. In this study, special focus is on non gainers as they need special attention from the online community in order to
have their confidence boosted up. Following are the three important factors, affecting the knowledge gain of a learner, addressed in forthcoming sections:

a) online interaction behaviour
b) regularity in attending online sessions
c) preference of a learner

4.3.1 Online Behaviour of Learners in a Learning Community

Interaction among online learners is beneficial for knowledge enhancement but all the learners do not exhibit an active online behaviour. Simply participating in a discussion is different from being actively involved or engaged in a discussion. Modern researchers pay immense attention to peer to peer interaction that induces a constructive and critical thought process and leads to cooperative learning in the participants (Slavin, 1995). Waters and Gasson (2006) categorized learners participating in a learning network on the basis of their roles namely: initiators, facilitators, vicarious acknowledgers, contributors, complicators, knowledge-elicitors, closers and passive-learners. This classification highlights that the learners in all the categories except the passive ones are highly interactive or active learners. Thus, online learners can broadly be classified into active and passive learners (Sharma, Banati and Bedi, 2011a; Bedi, Banati and Sharma, 2010). The passive ones do not make any explicit contribution to debate. They read postings of other participants, construct their own meaning privately but do not project themselves.

The presented work classifies active learners into two categories, broadly on the basis of their frequency of interaction, that is, those who are highly interactive, who herein would be termed as strongly active and not so active ones, who herein would be termed as weakly active learners.

There may be several factors due to which passive and weakly active learners are coerced away from being participative like lack of time (due to family or job constraints), disability, lack of motivation to study, lack of good presentation skills, inability to understand medium of instruction (language, computer skills), incomprehensibility of the content, and so forth. Whatever be the reason of passivity
(or marginal activity), the needs of these learners go unnoticed because of their negligible contribution to the discussions and hence could not be catered to. This eventually hampers their own prospects of knowledge enhancement as shying away from active discussions obstructs the dynamic phenomenon of knowledge construction that occurs naturally when there is uninhibited interaction.

This work focuses on the behavior of strongly and weakly active learners. The passive learners do not show up in any of the online discussions and their contribution to the datasets required for analysis would be nil, hence their behaviour cannot be analyzed in this study. Postings of weakly active learners, though few in number, can still provide significant information required for interaction analysis. Correlating the additional information like their regularity in attending online sessions may throw light on their knowledge gain.

4.3.2 Online Attendance

Regular attendance in face to face classes is analogous to regular presence in the online forums as a part of e-learning course. The students, lacking motivation to attend the sessions regularly, lag behind their peers as they cannot keep themselves abreast of the concepts being discussed in the sessions. Participation and interaction in live environment affects the thought process of the participants more dynamically than attaining passive knowledge by going through all the postings later on.

4.3.3 Preference of Interaction

Active learners interact with both peers and instructor but they may have their preference of interacting more with one over the other. The number of messages posted to the peer/instructor is considered as metric to determine the preference of interaction of a student with respect to each of them. Thus, the students have been classified into:

  * Instructor-oriented: Instructor-oriented learners interact more with the instructor. The reason that keeps them away from sharing their thoughts with fellow learners may be the fear that they would be mocked at.
Peer-oriented: Peer-oriented learners shy away from the instructor and like to discuss their doubts with some of their fellow learners.

Learners belonging exclusively to each of these categories are at loss because instructor-oriented learners who seldom discuss things with peers fail to get their cooperation and also lose out on the opportunity to get peer group’s share of perspective on their queries. On the other hand, peer-oriented learners who hesitate interacting with instructor are devoid of the expert advice that an experienced instructor would give to solve their problems.

The next section proposes a methodology to classify a learner into a gainer or a non gainer, depending upon whether he is strongly or weakly active keeping in consideration their regularity, and further tries to look into their preferences of interaction.

4.4 Learner Classification using Naïve Bayes’ Technique

The analysis of learner preferences is important for customization of remedial plans being designed for them, so that they are able to interact with whom they are comfortable with and do not feel isolated in the online environment. Information is elicited out of the ties among learners and instructors by analyzing their behaviour and the roles they play in an online learning community. The data thus generated is then used to classify learners into strongly or weakly active non gainers using Naïve Bayes classifier (Figure 4.4).

The preference of interaction of weakly active non gainers is then used to develop strategies to motivate the learners and to increase the frequency and quality of their participation, thus improving their performance. The post experiment test is carried out to find out the knowledge gain of the learners after being exposed to:

a) the interactive method of learning

b) restructured course with customized content
To evaluate the amount of knowledge gain as an effect of the proposed measures, a paired t-test is applied to the pairs of students’ previous scores and their post test scores. A null hypothesis and an alternate hypothesis are designed for the same.

Null Hypothesis \((H_0)\): There is no difference in the mean values of the students’ previous and post test scores after being exposed to the new method of learning.

Alternate Hypothesis \((H_1)\): The new method of learning causes change in the level of knowledge of students, i.e., there is a difference in the mean values of the students’ previous and post test scores after being exposed to the new method of learning.

4.4.1 Naïve Bayes’ Classifier

Various machine learning algorithms that deal with supervised learning, such as SVM, Neural Nets, Naïve Bayes, C4.5, have been compared over varied data sets under similar experimental conditions. Amongst all these Naïve Bayes classifier has been found to be very popular due to its simplicity (Gelman, 1996; Lowd and Domingos, 2005; Zhang and Li, 2007), no requirement of tuning any parameter unlike in SVM and neural network based classifiers (Williams, Zander and Armitage, 2006) and exhibits classification accuracy at par with other classifiers (Pazzani, Muramatsu and Billsus, 1996; Domingos and Pazzani, 1997; Friedman, Geiger and Goldszmidt,
Adaptive Content Sequencing Incorporating Social Opinion in an e-Learning Environment

1997; Kuncheva and Hoare, 2008). More importantly, Naïve Bayes classifier works well with very little data as in the absence of any prior instances (Kelly and Tangney, 2005; Caruana and Niculescu-Mizil, 2006).

Bayesian classifiers provide an intuitive approach towards modeling classification problems. According to Bayes Theorem, the probability that an event $C$ occurs provided that $D$ has also occurred is:

$$p(C|D) = \frac{p(C) \cdot p(D|C)}{p(D)}$$

where, the probabilities $p(C)$ and $p(D)$ are the prior probabilities of occurrences of events $C$ and $D$ respectively without any additional information of occurrence of events $D$ and $C$ in the respective cases. The probability $p(D|C)$ defines the posterior or conditional probability of the occurrence of event $D$ supported by additional information of occurrence of event $C$.

Figure 4.5 represents a typical situation where $A_1$ and $A_2$ are two independent events. Let $B$ be another event that may co-occur with events $A_1$ and $A_2$. The conditional probability of occurrence of event $B$ is given by:

$$p(B) = p(A_1) \cdot p(B|A_1) + p(A_2) \cdot p(B|A_2)$$

Figure 4.5    Representation of Bayes’ Theorem

Naïve Bayes classifier is based on the Bayes Theorem. It assigns an object, described by a feature vector $A = \{a_1, a_2, ..., a_n\}$, to one of the predefined classes belonging to set $V$, where $V = \{v_1, v_2, ..., v_m\}$. The classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or
absence) of any other feature, given the class variable. Associating an object with a certain class being described by a set of independent features can be modeled efficiently using on the basis of prior and conditional probabilities (Stern, Beck and Woolf, 1999). Strong assumption of attribute independence leads to the fact that attribute can be learned separately, hence simplifying the learning process (McCallum and Nigam, 1998), though Naïve Bayes is still optimal for both complete independence and complete dependence between the features (Rish, 2001; Kuncheva et al, 2008). Given the attribute values, the Naive Bayes classifier selects the most likely class, $V_{NB}$, where,

$$V_{NB} = \text{argmax}_{v_j \in V} p(v_j) \prod_{i=1}^{n} p(a_i | v_j)$$

Conditional probability $p(a_i | v_j)$ is calculated using simple estimation method m-estimates:

$$p(a_i | v_j) = \frac{(n_c + m \cdot p)}{(n + m)}$$

where,

- $n$ = the number of training examples for which $V = v_j$
- $n_c$ = number of examples for which $V = v_j$ and $A = a_i$
- $p$ = a priori estimate for $p(a_i | v_j)$
- $m$ = the equivalent sample size

This technique is applied in the next section to predict the knowledge gain of learners on the basis of their discussion patterns.

### 4.4.2 Learner Classification

The learners are classified as gainers or non gainers on the basis of two independent characteristics: a) online interaction and b) attendance. The interaction pattern of these learners can be analyzed by mapping the probability of being strongly active or weakly active in terms of the number of messages posted by the learner in an online learning environment. The student being regular or irregular can be evaluated from the records of the login sessions.
The following notations are used here on:

\[ N \] = Total number of students involved in online discussion

\[ I_G \] = Event depicting that the learner belongs to class “gainer”

\[ I_{NG} \] = Event depicting that the learner belongs to class “non gainer”

\[ W \] = Event depicting that a student is a weakly active learner

\[ R \] = Event depicting that a learner is regular in attending online sessions

Let \( p(I_G) \) be the prior probability that the student in question is a gainer.

Let \( p(I_{NG}) \) be the prior probability that student in question is a non gainer.

In phase I, to calculate the prior probabilities \( p(I_G) \) and \( p(I_{NG}) \), the background knowledge of all the students (test scores in course last attended) is taken into account. The probability of finding a non gainer in a group is given by \( p(I_{NG}) \), evaluated as total number of non gainers divided by the total number of students. \( p(I_G) \) can be calculated in a similar way. Further, it is assumed that if the number of messages posted by a given student in the online forum is more than or equal to half of the maximum number of messages posted by any student, then he/she is classified as strongly active else weakly active. The student who scored above passing marks was termed as a gainer and the one scoring below the passing marks was termed as a non gainer. So, probability of finding non gainers and gainers, respectively, in a group is calculated as:

\[
p(I_{NG}) = \frac{\text{Total number of non gainers}}{N} \quad \text{4-1(a)}
\]

\[
p(I_G) = \frac{\text{Total number of gainers}}{N} \quad \text{4-1(b)}
\]

The probability that the student is a non gainer, given that he or she is a weakly active learner and regular, is calculated as follows:

\[
p(I_{NG}|W, R) = p(W|I_{NG}) * p(R|I_{NG}) * p(I_{NG}) \quad \text{4-2(a)}
\]

The probability that the student is a gainer, given that he or she is a weakly active
Identification of Stakeholders and Classification of Learners

learner and regular, is calculated as follows:

\[ p(I_G|W,R) = p(W|I_G) \cdot p(R|I_G) \cdot p(I_G) \quad 4-2(b) \]

The conditional probabilities for \( p(I_{NG}|W,NR) \), \( p(I_{NG}|S,R) \), \( p(I_{NG}|S,NR) \), \( p(I_G|W,NR) \), \( p(I_G|S,R) \) and \( p(I_G|S,NR) \) can be deduced on similar lines.

4.4.3 Obtaining Learner Preference

It is considered that in online learning discussions, an active learner may prefer to interact either with fellow learners more than the instructor or vice versa. This study focuses on the customization of content for non gainers to improve their level of understanding; hence two significant issues are addressed:

a) a non gainer preferring to interact with fellow learners
b) a non gainer preferring to interact with the instructor.

In phase II, the preference of a non gainer is evaluated during interaction depending upon the number of messages of each learner (non gainer) and to whom they are posted to. The terms used in for the same are as follows:

\( I_{Ti} \) = Event showing \( i^{th} \) non gainer interacting with the teacher (instructor)
\( I_{Fi} \) = Event showing \( i^{th} \) non gainer interacting with fellow students (peers)
\( M_{Fi} \) = No. of messages posted by the \( i^{th} \) non gainer to fellow students (peers)
\( M_{Ti} \) = No. of messages posted by the \( i^{th} \) non gainer to the teacher (instructor)

4.4.4 Methodology for Learner Classification

This section presents the two phased methodology to perform learner classification using Naïve Bayes’ classification.

Phase I:

1. Create a database of all students enrolled in an online course containing the following attributes:
   - Student Id
   - Number of messages posted by the student
2. Calculate all the prior and conditional probabilities as discussed in section 4.4.2.

3. Equations (2a) and (2b) give the required probabilities for the combination of weakly active but regular learners.

4. If \( p(I_{NG}|W, R) > p(I_{G}|W, R) \), then the student is classified as a weakly active but regular non gainer else a weakly active but regular gainer. Probabilities for other combinations can be computed similarly.

**Phase II:**

5. For every non gainer node, calculate the indegree/outdegree of each node in the following way:

   For every message,
   
   if the target node is a peer, then
   \[
   M_{F_i} = M_{F_i} + 1
   \]
   
   else if the target node is an instructor, then
   \[
   M_{T_i} = M_{T_i} + 1
   \]

6. Calculate the probability that the \( i^{th} \) non gainer interacts with instructor, as:

   \[
   p(I_{T_i}) = \frac{M_{T_i}}{M_{T_i} + M_{F_i}}
   \]

7. Calculate the probability that the \( i^{th} \) non gainer interacts with peers, as:

   \[
   p(I_{F_i}) = \frac{M_{F_i}}{M_{T_i} + M_{F_i}}
   \]

   If \( p(I_{F_i}) > p(I_{T_i}) \), then the \( i^{th} \) non gainer prefers peers over the instructor for interaction and vice versa.
4.5 Experimental Study

A study was conducted on an online group of students of B. Sc. (H) Computer Science I Year. The students discussed their doubts, queries, assignment problems, on this group. Messages relevant to the Paper 202 (Computer System Architecture) were chosen for the experiment. Out of 128 students enrolled in the group, 113 students willingly participated in the online group. Their attendance for every session was logged. The marks obtained by the students in the relevant subjects of the courses they attended last, were considered for their initial classification into gainer or non gainer categories. A post experiment test was also taken and scores were evaluated to find out the amount of knowledge gain. They were informed beforehand that this exercise (inclusive of the scores of the tests conducted in this study) would not affect their curricular assessment in any way. The instructor acted as initiator and moderator in the group. Ten questions were finally formulated as an outcome of discussions between the instructor and other subject experts. One question was posed at a time in the forum and the students were given five days’ time to explore the concept. One day was given to the students to discuss the concept online by posting messages. For the sake of simplicity, the discussion threads from the online discussion forum only were considered for analysis although the students were free to use any online means to interact. An interaction over a period of 8 weeks was analyzed in the present study.

The students were categorized into a future gainers or non gainers on the basis of two attributes, namely, a) interaction and b) regularity. Data mining software WEKA was used to evaluate the results. The 10-fold cross-validation scheme was chosen for the experiment. The database contained 113 instances of students. The snapshot of the database is shown in Table 4-1. The column headings of Table 4-1 and the attributes depicted by them are given as follows:

A: Student Id
B: Number of messages posted by the learners
C: Number of messages received by the learners
D: Prior classification of learners into strongly (S) or weakly (W) active learners
E: Attendance

F: Prior classification of learners into regular (R) or non regular (NR) students

G: Prior classification of learners into gainers (G) or non gainers (NG)

H: Scores obtained by the learners in the courses last attended

Table 4-1  A snapshot of the database used for classification

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>16</td>
<td>S</td>
<td>89</td>
<td>R</td>
<td>G</td>
<td>68</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>22</td>
<td>S</td>
<td>78</td>
<td>R</td>
<td>G</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>11</td>
<td>W</td>
<td>66</td>
<td>NR</td>
<td>NG</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>11</td>
<td>12</td>
<td>W</td>
<td>60</td>
<td>NR</td>
<td>NG</td>
<td>44</td>
</tr>
<tr>
<td>39</td>
<td>13</td>
<td>22</td>
<td>S</td>
<td>74</td>
<td>R</td>
<td>G</td>
<td>54</td>
</tr>
<tr>
<td>40</td>
<td>19</td>
<td>19</td>
<td>S</td>
<td>51</td>
<td>NR</td>
<td>NG</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>19</td>
<td>10</td>
<td>S</td>
<td>66</td>
<td>NR</td>
<td>G</td>
<td>64</td>
</tr>
<tr>
<td>100</td>
<td>11</td>
<td>20</td>
<td>W</td>
<td>69</td>
<td>R</td>
<td>G</td>
<td>55</td>
</tr>
<tr>
<td>101</td>
<td>10</td>
<td>4</td>
<td>W</td>
<td>36</td>
<td>NR</td>
<td>NG</td>
<td>35</td>
</tr>
<tr>
<td>102</td>
<td>6</td>
<td>3</td>
<td>W</td>
<td>57</td>
<td>NR</td>
<td>NG</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Initially the students were classified as gainers or non gainers on the basis of the marks obtained by them in the course last attended, as test scores for the present session were not available at the commencement of the course. Naïve Bayes classifier used this information along with the data about the regularity and interaction of the students to calculate prior probabilities, in addition to keeping it as a data for later comparisons. The classification results (Table 4-3) obtained using Naïve Bayes classifier show that 94 out of 113 students were classified correctly and 19 incorrectly. Additional inputs about the interaction and regularity therefore provide a clear picture with respect to the prediction of gainers v/s non-gainers as shown in Table 4-2 and Table 4-3.
### Table 4-2  Summary of the input data for Naïve Bayes Classification using WEKA

<table>
<thead>
<tr>
<th>Attribute</th>
<th>NG</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR</td>
<td>51.0</td>
<td>3.0</td>
</tr>
<tr>
<td>R</td>
<td>14.0</td>
<td>49.0</td>
</tr>
<tr>
<td>[total]</td>
<td>65.0</td>
<td>52.0</td>
</tr>
<tr>
<td>Interaction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>21.0</td>
<td>41.0</td>
</tr>
<tr>
<td>W</td>
<td>44.0</td>
<td>11.0</td>
</tr>
<tr>
<td>[total]</td>
<td>65.0</td>
<td>52.0</td>
</tr>
</tbody>
</table>

### Table 4-3  Summary of the results obtained by Naïve Bayes Classification

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Number of Instances</td>
<td>113</td>
<td></td>
</tr>
<tr>
<td>Correctly Classified Instances</td>
<td>94</td>
<td>83.1858 %</td>
</tr>
<tr>
<td>Incorrectly Classified Instances</td>
<td>19</td>
<td>16.8142 %</td>
</tr>
<tr>
<td>Kappa statistic</td>
<td>0.6641</td>
<td></td>
</tr>
<tr>
<td>Mean absolute error</td>
<td>0.202</td>
<td></td>
</tr>
<tr>
<td>Root mean squared error</td>
<td>0.3265</td>
<td></td>
</tr>
<tr>
<td>Relative absolute error</td>
<td>0.9311 %</td>
<td></td>
</tr>
<tr>
<td>Root relative squared error</td>
<td>65.7165 %</td>
<td></td>
</tr>
</tbody>
</table>

Results shown in Table 4-3 reflect that Naïve Bayes classifier provides an intuitive and satisfactorily accurate means of classification of students into gainers or non-gainers based upon other attributes, namely interaction and regularity. Kappa statistic suggests that the results are fairly agreeable. The content of the messages posted by the prospective non-gainers was then analyzed to judge how much they were able to comprehend about the given topic. On the basis of the classification of learners into gainers or non-gainers, the following groups were formed, based upon their interaction behavior (in light of their regularity in attending online sessions):
**Strongly active and Regular Non gainers:** Although the learners falling in this category posted a large number of messages, but most of those messages were far from being the contextual ones, hence did not add any significant gain in their knowledge.

**Strongly active but Non Regular Non gainers:** Learners in this category failed to attend the sessions regularly. As they were characterized as strongly active, this implies that they also posted a large number of messages, but as in the above case, majority of them were found to be irrelevant.

**Weakly active and Regular Non gainers:** This kind of learners needs special attention as despite of attending the online sessions regularly, they failed to make it to the gainer category.

**Weakly active but Non Regular Non gainers:** These learners did not attend the online sessions either due to lack of time or motivation. Irregularity combined with weak interaction hampered their prospects of getting a high score.

A student who is not able to achieve satisfactory results while pursuing an e-learning course may tend to drop out in due course. Invariably, the evaluation process begins late (usually at the end of a certain topic or course module) and it is too late to take the remedial step. Hence, measures must be taken, right at the beginning, to predict the learning curve of the learners and motivate them, if they seem to be future non gainers, in order to retain and usher them towards completion of the course. The following subsection discusses some remedial steps that may be adopted to achieve this goal.

### 4.6 Remedial Plans for the Learners

Each learner has his own learning curve and hence requires his own learning path adaptable to his needs. Further, different learners have different interaction patterns. Hence, there is a strong need for adopting various remedial steps based on proven instructional strategies at this juncture to let the slow gainers learn in ways that suit their learning styles. Over the years, researchers in the field of education have
conducted intensive research on devising instruction strategies for creating an effective e-learning environment. The instruction strategies employed in this chapter are not exhaustive, yet are adept in addressing the issues raised here.

The strength of strongly active learners lies in the fact that they are good in making their online presence felt. This trait can be tapped to steer these learners towards achieving their learning objectives by carefully selecting the right instructional strategies. Blending the challenge-based and problem-based methods of instruction with collaborative learning helps the strongly active learners in channelizing their socializing skills in a constructive manner (Savery and Duffy, 2001; Ramos and Espinosa 2003). Challenges/problems posed by the instructor give the learners a direction to think and brainstorming sessions let them bring their reflections to a common platform. Similarly, collaborative projects/assignments keep them engaged in learning activities till the accomplishment of their goals.

Weakly active learners, on the other hand, need instruction in a self-directed learning mode that lets them achieve their learning objectives in their own space. Knowles (1986) proposed a learning framework named ‘learning contract’ (Caffarella and Caffarella, 1986; Codde, 2006) that specifies— a) the learning objectives; b) learning resources and strategies; c) the target date for their accomplishment and d) evaluation of the learning activities. Instructor can help learners in designing their respective learning contracts. This also helps in cultivating a sense of responsibility in the learners. Codde (2006) sums up that the learning contracts encourage students to become active participants in their course of learning.

In case of non-regular learners (either strongly or weakly active), it is important to profile the causes of their abstinence from the learning sessions. Some of the prominent reasons of the learners keeping away from classes are, inability to use computer functionalities, inability to find the content, feeling of loneliness or being neglected (Rovai and Wighting, 2005; Nagel and Kotze, 2010). Once the reasons are known, the instructor can employ suitable policies and motivating strategies to foster learners’ regular involvement in the online learning activities. Introduction of attendance-based grading is a debatable policy. While on one hand, it makes some
learners more conscious of their presence and behavior in the class, on the other hand, it may increase discomfort in learners (Sleigh and Ritzer, 2001). Posting of regular messages could be another possible measure that may help the learners academically and enhance both the quantity and quality of interactions with their instructor (Hickerson and Giglio, 2009). Similar measures may encourage learners in attending the online sessions regularly.

Phase II (as discussed in Section 4.4.4) provides a mechanism to judge the level of comfort of the learners towards their peer group and instructor. The non-gainers are categorized into those who prefer to interact with their peers and those whose orientation seems more towards the instructor. The present study proposes certain approaches (category wise) that may be adopted to benefit both kinds of learners as follows:

**Peer-oriented Learners.** A conducive interaction platform should be provided, where they can interact with their friends (as it happens in face to face environment) without any fear of being mocked at. Instructor should play the role of a moderator/facilitator rather than a teacher. Online assignments/projects should be collaborative in nature such that they are able to enhance their knowledge through brainstorming sessions.

**Instructor-oriented Learners.** These learners may not be grouped together initially. Rather, the instructor should play the role of an initiator/expert/guide so that the learners receive a constructive feedback as and when required. The messages should be intended towards individual students to elicit their views. Such messages may be posted to individuals through e-mails or one to one chat forums. Once the process is initiated, the participants get motivated in due course and start interacting with each other. At that stage they may be introduced into the pre-defined groups.

The remedial plans A, B, C and D (Figure 4.6) reflect the summary of actions to be taken for various kinds of learners (as discussed above).
The combinations of these plans, for different categories of learners, that may help the learners strengthen their weak areas and subsequently improve their knowledge gain are illustrated in Table 4-4.

### Table 4-4 Proposed Combinations of Remedial Plans for Various Categories of Learners

<table>
<thead>
<tr>
<th>Learner Characteristics</th>
<th>Regular</th>
<th>Non-regular</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strongly Active</strong></td>
<td>Remedial Plan A</td>
<td>Remedial Plan A + Remedial Plan B</td>
</tr>
<tr>
<td></td>
<td>Remedial Plan C</td>
<td>Remedial Plan B + Remedial Plan C + Remedial Plan D</td>
</tr>
<tr>
<td><strong>Weakly Active</strong></td>
<td>Remedial Plan D</td>
<td>Remedial Plan D</td>
</tr>
</tbody>
</table>

Figure 4.6 Remedial Action Plans

[Inclusion of challenge-based/problem-based method of instruction (games, quizzes, etc.) to prevent deviation from topic]

- Collaborative assignments/projects to enhance knowledge through brainstorming sessions

[Remedial Plan C]

- Grouping of learners into peer-oriented and instructor-oriented groups and introduction of content accordingly
- Instructor to play the role of a:
  - moderator/facilitator for peer-oriented learners
  - initiator/expert/guide for instructor-oriented learners

[Remedial Plan D]

- Instructor to ensure self-directed learning and help learners in designing their respective learning contracts by chalkling out:
  - the learning objectives
  - learning resources and strategies
  - target date for activity accomplishment
  - evaluation scheme

[Figure 4.6 Remedial Action Plans]
The remedial plan combinations as proposed in Table 4-4 were introduced in the learning environment under study by suitably directing the students to e-resources which provided them with relevant learning material with respect to their requirements and preferences. A post-experiment test was taken by the students. A snapshot of the updated database reflecting post-test scores is given in Table 4-5. All column headings in Table 4-5 (except the last one) represent the same attributes as in Table 4-1. The last column heading, I, depicts the scores of the post-test taken by the learners in the courses they attended last.

**Table 4-5**  A snapshot of the updated database with post test scores

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>16</td>
<td>S</td>
<td>89</td>
<td>R</td>
<td>G</td>
<td>68</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>22</td>
<td>S</td>
<td>78</td>
<td>R</td>
<td>G</td>
<td>50</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>11</td>
<td>W</td>
<td>66</td>
<td>NR</td>
<td>NG</td>
<td>38</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>11</td>
<td>12</td>
<td>W</td>
<td>60</td>
<td>NR</td>
<td>NG</td>
<td>44</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>13</td>
<td>22</td>
<td>S</td>
<td>74</td>
<td>R</td>
<td>G</td>
<td>54</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>19</td>
<td>19</td>
<td>S</td>
<td>51</td>
<td>NR</td>
<td>NG</td>
<td>39</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>19</td>
<td>10</td>
<td>S</td>
<td>66</td>
<td>NR</td>
<td>G</td>
<td>64</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>11</td>
<td>20</td>
<td>W</td>
<td>69</td>
<td>R</td>
<td>G</td>
<td>55</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>10</td>
<td>4</td>
<td>W</td>
<td>36</td>
<td>NR</td>
<td>NG</td>
<td>35</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>6</td>
<td>3</td>
<td>W</td>
<td>57</td>
<td>NR</td>
<td>NG</td>
<td>43</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The paired t-test performed on the students’ previous scores and the scores of the post experiment test shows that the results are statistically significant (Table 4-6). The value of $\alpha$ was kept to be 0.05. A large, negative t (-8.69, rounded off), signifies that the experimental test scores (post test) were greater than the previous scores. Two-tailed P (the probability of observing a difference as large as or larger than observed, if the null hypothesis were true) being very small ($P<0.05$) reflects that the change in mean values are due to adopting the new learning approach and not due to
Identification of Stakeholders and Classification of Learners

coincidence of random sampling. The hypothesized value of the difference in means is zero and does not lie in the 95% confidence interval of the mean difference (-4.64 to -2.92). Hence the null hypothesis can be rejected. The test results indicate that the improvement in the post test scores of most of the learners and hence reflect their knowledge gain which may be attributed to the remedial measures adopted during the online sessions.

Table 4-6  Results for Paired t-Test: Paired Two Sample for Means

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Pearson Correlation</th>
<th>t Stat</th>
<th>P (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous Score</td>
<td>52.73</td>
<td>16.76</td>
<td>0.961333244</td>
<td>-8.6893</td>
<td>3.43618E-14</td>
</tr>
<tr>
<td>Post Test Score</td>
<td>56.514</td>
<td>15.82</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean of Difference between values of two groups: -3.78

95% Confidence Interval of this difference: -4.64 to -2.92

This chapter presented the first phase of KMeLS, that is, the Socialization phase. Socialization deals with identification of stakeholders playing key roles in the establishment of an e-learning environment. Interaction among the two major stakeholders, that is, the learners and the instructors was analyzed to elicit the preferences of the learners. This forms the first step towards content designing. However no content can ensure 100% success rate in terms of knowledge gain of the learners. Therefore, appropriate remedial plans were chalked out for the learners who could not gain effectively at the first go. These plans form a strong backbone of any e-learning content by helping in enhancing the knowledge levels for all kind of learners.

The identification of stakeholders helps in carrying out the second phase of KMeLS, that is, Externalization, where these stakeholders are requested to give their opinion on prioritization of learning content for the prospective learners of an online course. The Externalization phase is presented and discussed in the next chapter.