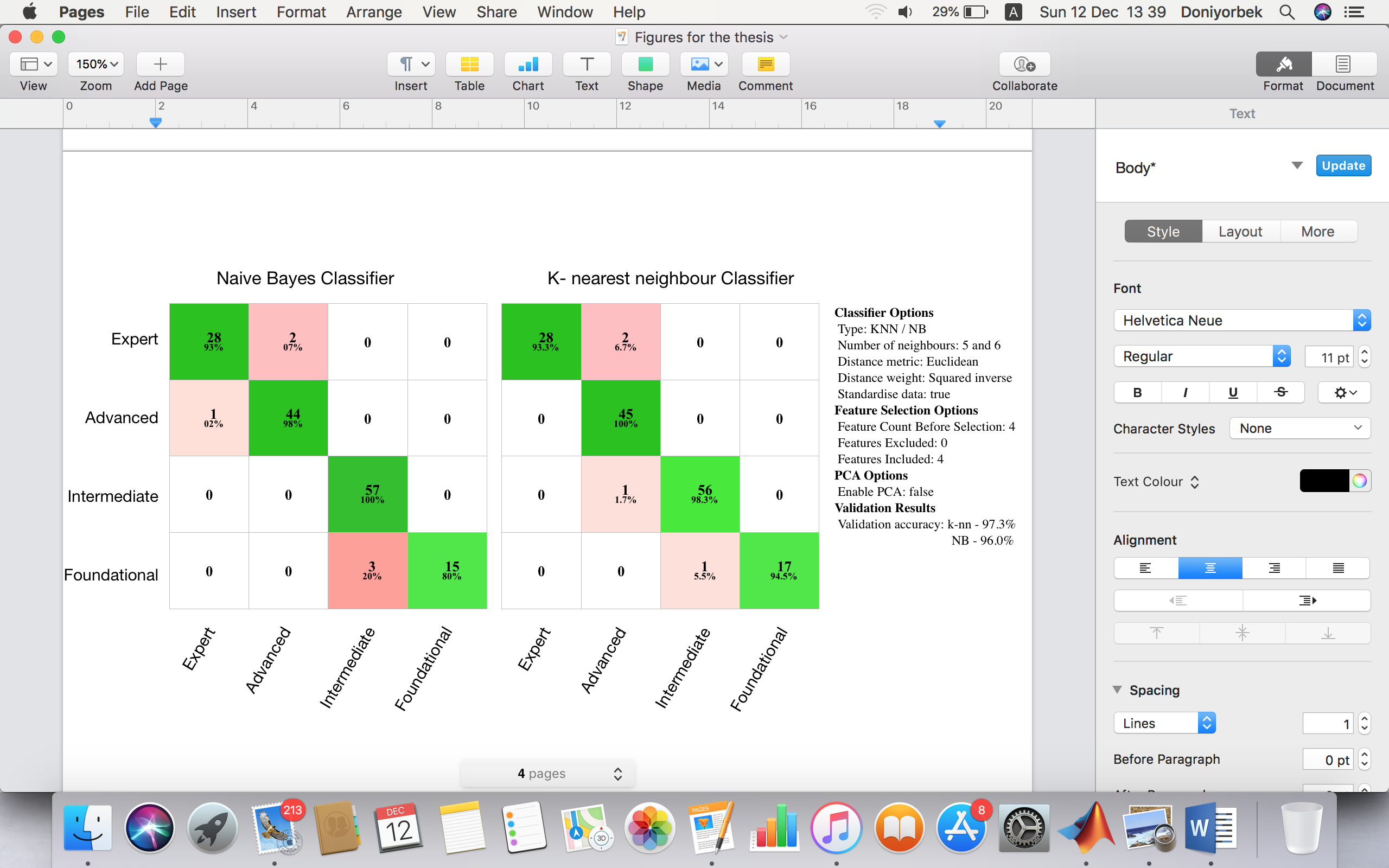
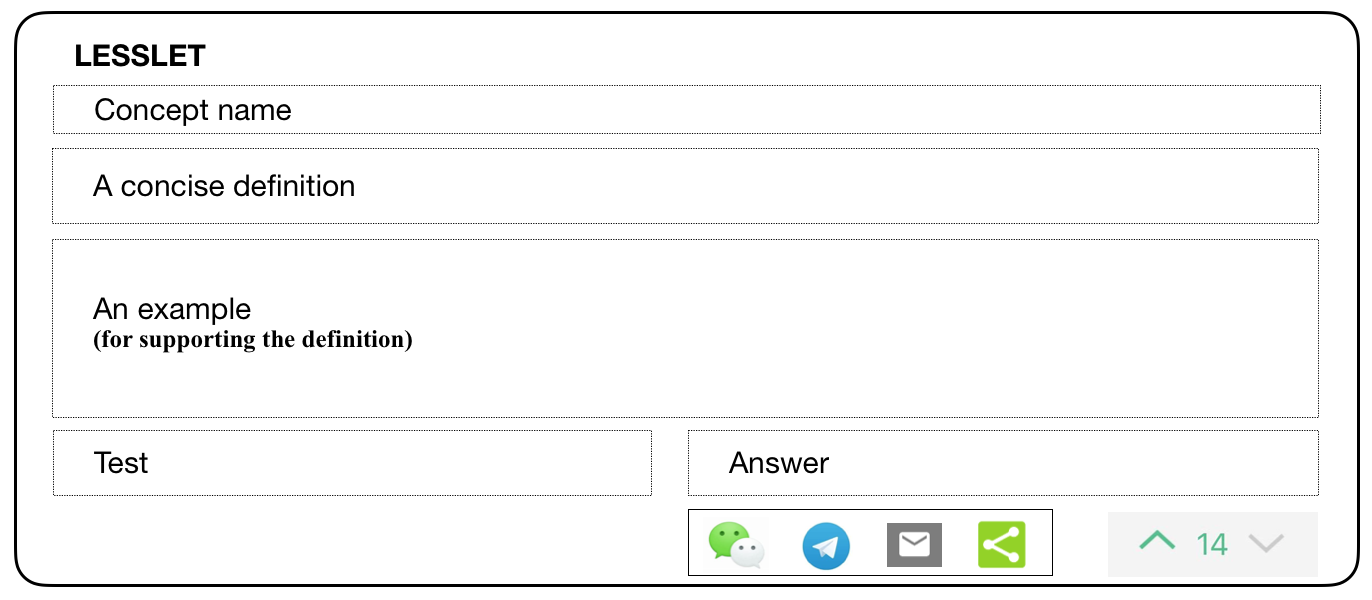
1. Lesslet and domain model
   1. Micro-lessons (Lesslet)

The *Lesslet* project was adopted from [26] as a crowdsourcing method to implement student generated content. As it is defined, lesslet is a mini-lesson (here it is defined as lesson letter or LessLet as a short) runs from a webpage as a simple program like a web-applet. The activity is handled as crowdsourcing approach in which the users of the system (teachers and learners) play the main role in creating and maintaining the lesslets. This phenomena, in literature, appears in a number of overlapping terms, such as - user generated content, peer-generated content [29], learning through teaching [26], peer assignment [30], micro-lesson [25], etc.

As its specifications, a lesslet can have five basic components: title, concise definition, an example “for supporting the definition”, a test “derived from previous two parts”, and an answer Fig.5.1. The most inspiring feature with the lesslet is that its self-contained feature. It is served to implement activities associated with problem solving, collaborating, assignment or collaborative assessment and self regulated learning. While the creator defining the concept, the user must think of a corresponding example and it supports the reinforcement learning. On the other hand, another user actively participates with a comfort feeling which the content, he is reading, is made by one of her peers. Besides, he/she learns with extra conscious with a feeling of being ready for taking its test and for other reaction components (e.g comments and upvote/downvote) and it supports a critical thinking of students. [30] puts this as “You solve, I learn”, in which he tries to use the potential of gifted students’ skills and brings up how they solve the problems in their different way of approach.



**Fig. 1 A typical example for lesslet**

The great necessity for this kind of contents is the requirement for a massive learning materials. Another problem is to create a platform for these contents so that students can easily refer to them or find a learning content easily which addresses his/her learning needs and preferences. Creating so called platform is a dream for every researcher in technology enhanced learning and for every developer of ITSs. Later we may divide the problem into two groups as creating massive/diverse learning contents and maintaining the system to interact the learners with content and with other learners. The later problem may be addressed by adopting online social networking theories like Facebook.

* 1. Peer assessing

Pear assessment strategy in education is considered as one of the impacting aspects on student encouragement with comparing and reflecting by reviewing peers’ work [32]. Two types of peer assessment are generally identified: scoring (up-/down-voting) and leaving comments on peer’s learning tasks or activities. Student’s peer assessing helps them reflect on and improve their own performance by reacting on other students’ work [33].

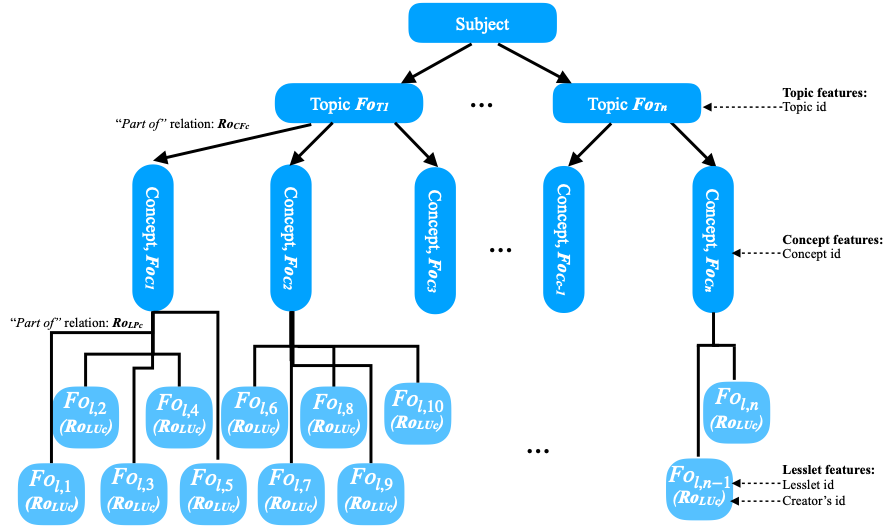
However, doubts may be found as learners’ reactions on peers may not reach as qualified as assessing them since their prior knowledge can cause to mis-evaluate. Nevertheless, [34] implemented a scoring instrument in which students produce their academic works (writing essays), than he asked peer students and teachers to evaluate by using the same scoring instrument. Their results yield strong degree of grading consistency between students and teachers. Furthermore, [35] researched formative and summative assessment activities as a positive judgment ability of learners. They have reported that students were capable of accurately assessing their peers on formative evaluation process.

**Table 1. A sample of a student’s score values and behavioural measures which may get on a topic with five concepts.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Lesslets for five concepts | Votes given by peers (upvote +, downvote -) | | | | | | Total |
| Lesslet 1 | + | - | + | + |  |  | 2 |
| Lesslet 2 |  |  |  |  |  |  |  |
| Lesslet 3 | + | - | + | + | + |  | 3 |
| Lesslet 4 | + | - | + | + | + | + | 4 |
| Lesslet 5 | - | + | - | - |  |  | -2 |
| Scoring behaviour - value in the interval [0,1] (1 is the highest) representing the fraction of score equal or less than -2 | | | | | | | 0.8 |
| Score - continuing number, representing the sum of scores for the topic | | | | | | | 7 |
| Fraction - value in the interval [0,1] representing the fraction of created lesslets (in this example there is no lesslet created for 2nd concept). | | | | | | | 0.8 |
| Visitor - number of visitors | | | | | | | 25 |
| User knowledge level for the topic: (based on expert opinion) “Advanced” | | | | | | | |

* 1. Generic user knowledge modeling as an overlay model
     + 1. Domain model

User modeling is one of the important mechanisms on designing adaptive systems [36]. We used overlay model technique for learner’s knowledge modelling. A structured domain model was first constructed which consists of concepts and associated feature objects. In domain modelling the concepts can be as simple as a fact or as large as a chunk of knowledge elements. The term *learning object* was adopted as a synonym to the concept that is because it can denote a fragment of knowledge of any procedural knowledge [37]. For our object model **O,** {o1, o2, o3, …, on} refers to the application domain, ***n*** denotes the number of knowledge components and topics in the domain Fig 5.2 shows the generalised object model. Lesslets are other kind of learning objects associated to a certain object and they are user generated contents. It derives that ***c***∊***n*** ∀ ***oc***∊***O***. A domain model was structured with two sets ***Foc*** and ***Roc*** . The first one is the set of individual features ***Foc*** *{foc1, foc2, foc3, …, focn}* of learning objects. The second kind of objects are the set of correlated features ***Roc*** *{roc1,roc2,roc3, …,rocm}* of **oc** object. The topic will be marked as grasped if the user has managed to create lesslets for each concept associated to the topic and if those lesslets are met the criteria in a duration of given session time. To implement the content recommendation, the features of lesslets are exploited in machine learning classification techniques in the next section.



**Fig. 2 Generalised object model**

* + 1. Knowledge model

The data in knowledge model is the main source to which adaptation processes depend on. Knowledge model keeps some information about user’s knowledge level on a certain knowledge element in domain model. For this reason those two models are strongly connected to each other. Also, some features of knowledge model elements, other than knowledge level, resembles user’s personal, background, or user traits and these features are usually static or changes very little during the entire learning period.

Given a student’s knowledge data as a set of lesslets created for a topic see Table 5.1. In this example, the degree of scores are *advanced* according to the expert human judge, that the user’s this feature for the topic is considered to overlay higher than average of an expert knowledge level.

**Table 3 Representation of user features for learning object oc.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Set of ***Fuc***(user features for the set of ***Foc***object) | | | | |
| ***fuc1*** | ***fuc2*** | ***fuc3*** | ***…*** | ***fucn*** |

List of user features about domain model objects

* (SB) - scoring behaviour
* (ULS) - lesslet score
* *(LF)* - fraction of lesslets created
* (*LV*) - number of visitors
* *(UK)* - user’s knowledge level for the topic
  1. Applying user modelling approach to the knowledge classification methods

After modelling process done, for tracking, evaluating the user activities and for feeding and maintaining the user knowledge model, it is required to apply some of the reasoning approaches. ITS usually utilises machine learning or rule-based methods to provide adaptive user experiences. We have used two machine learning classifiers for this study. Machine learning methods are used to build models in order to make predictions based on the evidence(s) when the results or circumstances are uncertain. As algorithm identifies the patterns from dataset, it adaptively makes assumptions for new observations.

**Table 4 Training dataset samples for each class trained with k-nn and Naive Bayes classification.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***No*** | SB | ULS | LF | LV | UK (class) |
| ***1*** | 1 | 16 | 0.8 | 52 | Expert |
| ***2*** | 1 | 9 | 0.8 | 18 | Advanced |
| ***3*** | 0.7 | -6 | 0.7 | 9 | Intermediate |
| ***4*** | 0.4 | -11 | 0.5 | 8 | Foundational |

For every topic, there have specified a real world dataset, which obtained during previous teaching sessions, to train and test the two machine learning (k-nn and Naive Bayes) algorithms for knowledge classification. Four knowledge levels were specified as class names as Table 5.4 presents four examples from dataset - full dataset can be seen at Appendix B. The dataset consists of 50 samples of records distributed for each knowledge level which are examined carefully by an expert teacher. We could not have equal number of samples from all knowledge level (Expert, Advanced, Intermediate, Foundational). Since, among a group of students the ratio of Expert, Advanced, Intermediate or Foundational students is unfeasible. These two classification algorithms discussed below are commonly used in educational ITSs.

* + 1. A probabilistic method: development of Naïve Bayes knowledge classification

For Bayes classification, kernel smoother type is chosen, since the predictor values of feature class are in scalar type. The kernel density was set to unbound because the density of values in two column observations can extend over the whole real line Table 5.4.

Naive Bayes Classifier (NBC) is a probabilistic and supervised machine learning method. NBC is usually chosen for its simplicity and commonly used for natural language processing, user modeling, and medical diagnostics. NBC results are assumptions for the state in which the feature observations are given. And also, these features which affect to the state are independent, for that reason it is often referred as “feature independent model”.

Let {a1, a2, a3…, aq} be independent feature set ***A*** as predictors and ***C*** is the class which comes out as a predicted result. Then, the posterior probability can be calculated as:

(1)

where:

* is likelihood joint density of predictors given class ***C,***
* is the class prior probability
* is the joint density of predictors

We can rewrite the (1) equation with our predefined independent feature arguments ***Fuc (****SB, ULS, LF, LV****)*** given set of user features and predict (*UK*) the user’s knowledge degree or the class of hypothesis for NBK*.* The probability degree of a user’ knowledge for a topiccan be calculated from:

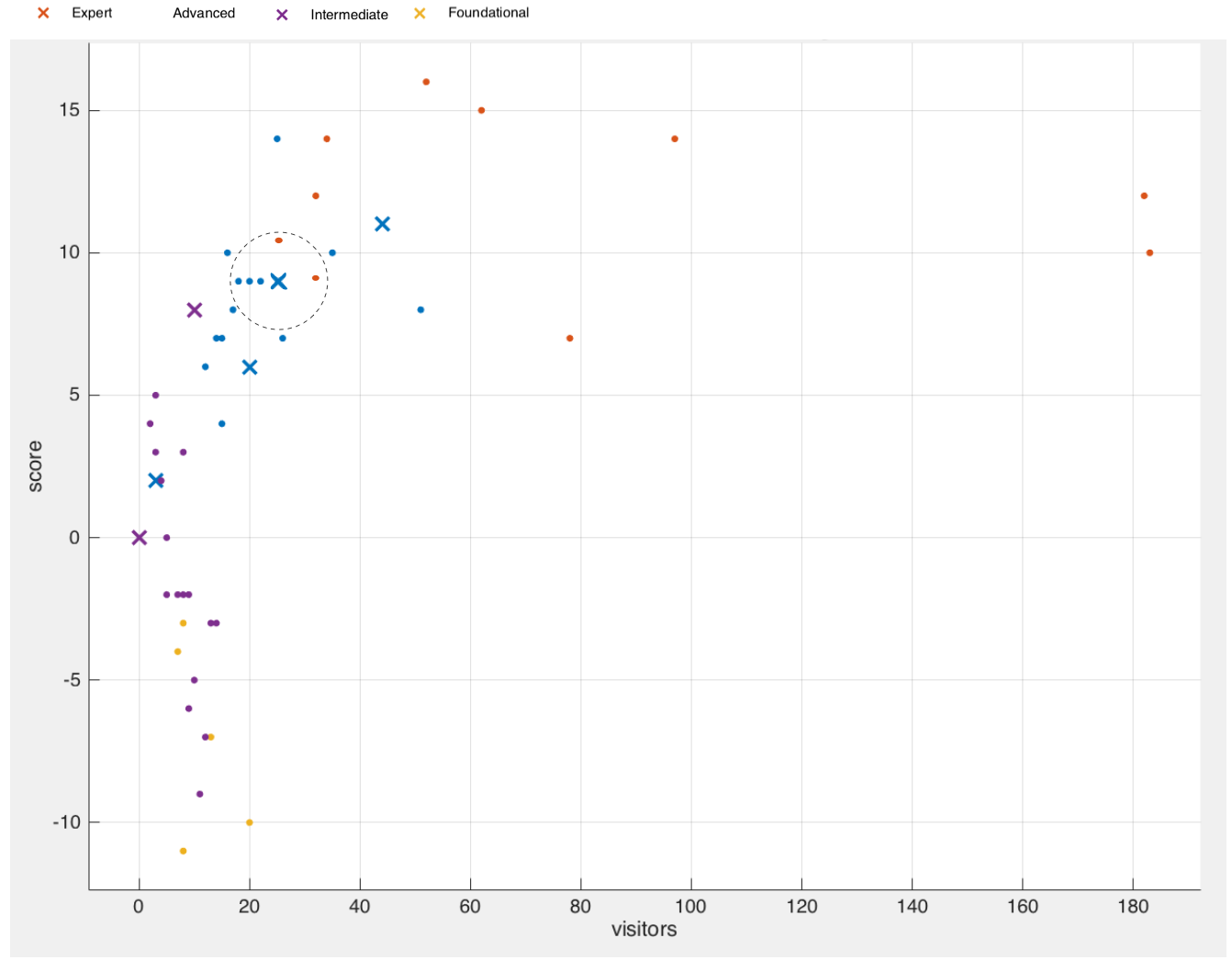
(2)

Table 5.3 shows the feature of attributes and class hypothesis which affecting for students knowledge class. Attribute density was defined with the criteria given by subject instructors.

* + 1. A probabilistic method: development of k-NN knowledge classification

For knowledge classification, among the supervised learning methods, nearest neighbour (k-nn) classifier is widely used. This sub-section describes briefly on k-nn classification and its application to our particular study.

Categorising the new points according to the distance from the other points in (the world) training dataset is an effective classification method of k-nn. Suppose，we have six new observations and one of them has similar predictors with the data in Table 5.1 (25 visitors and 0.8 score) and they appear in the plot Fig 5.3 as “x” colours indicating their class names. We make the value of *k* equals to five. The size of dotted circle indicates k value which encompasses five other points and to make classification decision as “*Advanced*”.



**Fig. 3 Scatter plot of trains set for: k-Nearest Neighbour**

The three things need to be defined for classification are k-value, training dataset, choosing one of the distance metrics. The *Euclidean* (EU) and *Minkowski* (MI) distance metrics (3,4) are famously used metrics and I also used these two for evaluation. Euclidean is a distance measuring method used for detecting the length between two points in Cartesian coordinate system. Say, *x* and *y* are two points in a plain and their distance length can be converted into metric distance by means of Euclidean. The formula for IE can be written as:

(3)

The formula for MI can be written as:

(4)

Table 5.4 shows the domain’s feature (predictor) data of users and response data as class names representing the user’s knowledge level. Algorithm k-nn classifies the learner’s knowledge level (last column on the table) for the current topic depending on the feature values given on first four columns in the table. As it is shown in Table 5.4 there four target class distributions: *Expert, Advanced, Intermediate* and *Foundational*.

* + 1. Comparison of two given knowledge classification results

We have used Matlab (software version MATLAB\_R2015) classification learner APP which is Machine learning algorithm “Classification learner”. Classification accuracies of NBC and k-nn are discussed in this sub-section. For this purpose, dataset with 150 samples are used to train and validate the algorithms Table 5.5.

**Data preparation** - Data types in feature set are all double type and the class vector values are in sting categorical data type.

The best results with 150 observations for k-nn are 97,3% and 96,7% accuracy on EU and MI respectively. These results were obtained with k value set with {5 and 6}. The best results for NBC were 96%.

**Table 5 Comparing the performance of algorithms**

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm** | **k-NN** | | **NBC** |
| **EUk=5,6** | **MIk=5** |
| **Average number of misclassified observations** | **4** | **5** | **6** |
| **Classification accuracy (%)** | **97,3** | **96,7** | **96** |

During the experiment we have evaluated different options available to control the algorithm in order to obtain better results from classification methods. The distribution names of BNC have to be specified because of the predictors were all not similar, for example some of them were comma-separated pair values. When we train BN classifier model using Gaussian distribution for all predictor distributions the output of confusion matrix was 86,4% accuracy. Major misclassification area has occurred in predicting the “*Foundation*” as “*Intermediate*” knowledge level. Since the misclassification of these two level of knowledge were not as significant as to consider for recommendation process, because in this adaptive learning system the content recommendation have more consideration to upper levels of knowledge. Then, we used comparison of cross validation and model generalisation error rate got better from 0.14 to 0.10. In addition, after opting the *kernel* distribution with smoother type, we could improve the confusion matrix to about 96% accuracy.