**BACKGROUND AND RELATED WORK**

Due to the rapid growth of information and communication technologies, the education environment has been enriched with those technologies and become more diversified.

In our days there are many systems that can teach the user on any subject and particular in programming. Those type of systems is called "E-Learning".

E-learning systems have become popular tools for teaching and learning. Advanced e-learning systems, such as WebCT (www.webct.com) and Cyber University of NSYSU have been developed recently that integrate a variety of functions. For example, these systems can be used to integrate instructional material (via audio, video, and text), e-mail, live chat sessions, online discussions, forums, quizzes and assignments, and the World Wide Web. With this kind of system, instructional delivery and communication between instructors and students can be conducted at the same time (synchronously) or at different times (asynchronously). Such systems provide a variety of instructional aids and communication methods and offer learners great flexibility as to the time and place of instruction. As a result, these e-learning systems may better accommodate the needs of learners who are geographically dispersed and have conflicting schedules.

**QuizJET: parameterized questions for programming language**

This system is kind of E-learning system. It has been designed to support web-based authoring, delivery, and evaluation of parameterized questions for programming language. QuizJET can be used for assessment and self-assessment of students’ knowledge of a broad range of a programming language topics from language basics to advanced concepts, such as polymorphism, inheritance, and exceptions. In a taxonomy of task types in computing, questions generated by QuizJET belong to the group of prediction tasks. These tasks are becoming increasingly popular in various computing-related courses. To a large extent, the nature of tasks generated by QuizJET follows the approach explored earlier in QuizPACK However, the switch of the domain from C to Java allowed QuizJET to generate questions of much larger complexity.

QuizJET offers a form-based online authoring interface for developing new quizzes and questions.

Graphical user interface, text, application

Description automatically generated

Fig 1 The presentation of a QuizJET question.

Graphical user interface, text

Description automatically generated

Fig 2 The evaluation results of a QuizJET question

**Classification and The Decision Tree Algorithm**

Our Web system needs to classify the user according to his level into an appropriate pool of questions, based on a short questionnaire that the user will be required to fill out in the first registration for the site. With the help of clustering algorithm and decision tree of data mining technique it is possible to classify each user accordance to his previous knowledge and to his level in real time.

Data clustering is a process of extracting previously unknown, valid, positional useful and hidden patterns from large data sets. The amount of data stored in educational databases is increasing rapidly. Clustering technique is most widely used technique for future prediction. The main goal of clustering is to partition students into homogeneous groups according to their characteristics and abilities.

1. **Data Clustering**

Data Clustering is unsupervised and statistical data analysis technique. It is used to classify the same data into a homogeneous group. It is used to operate on a large dataset to discover hidden pattern and relationship helps to make decision quickly and efficiently. In a word, Cluster analysis is used to segment a large set of data into subsets called clusters. Each cluster is a collection of data objects that are similar to one another are placed within the same cluster but are dissimilar to objects in other clusters.

1. **Data mining**

Data mining, also popularly known as Knowledge Discovery in Database, refers to extracting or “mining" knowledge from large amounts of data. Data mining techniques are used to operate on large volumes of data to discover hidden patterns and relationships helpful in decision making.

**Decision tree**

Decision Tree induction can be integrated with data warehousing techniques for data mining. A decision tree is a classification and prediction tool having a tree-like structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node (terminal node) holds a class label.

**The Algorithm**

The basic algorithm for decision tree induction is a greedy algorithm that constructs decision trees in a top-down recursive divide-and-conquer manner. Decision Tree Algorithm: generate a decision tree from the given training data :

1. Create a node N
2. 2 If samples are all of the same class, C then
3. Return N as a leaf node labeled with the class C;
4. If attribute-list is empty, then
5. Return N as a leaf node labeled with the most common class in samples.
6. Select test-attribute, the attribute among attribute-list with the highest information gain.
7. Label node N with test-attribute;
8. For each known value ai of test-attribute.
9. Grow a branch from node N for the condition test attribute = ai ;
10. Let Si be the set of samples for which test-attribute = ai;
11. If Si is empty, then
12. Attach a leaf labeled with the most common class in samples;
13. Else attach the node returned by generate-decision tree (Si,attribute-list-attribute);

**Related work:**

Today many systems are available for learning software principles.

These systems can be categorized according to the learning approaches they are based on:

**Write and solve:**

Based on this approach, a user is presented with challenging problem and he must solve it by writing code. After writing, he will receive feedback on whether the code he wrote meets the requirements and does in fact solve the challenge.

There are some systems that prepare a work schedule for several weeks in which the user encounters code questions at every level from the lower to the more advanced.

To practice and learn using these sites, the user must spend a lot of time solving the questions.

This approach is used by, for instance: LeetCode , HackerRank and CodeChef.

**Knowledge and memorization questionnaires:**

This approach involves the user being presented with a large number of questions and several option for answers. The user will need to mark the right answer.

As the user practice with this approach, he will encounter most questions in a short time and will be able to identify what the correct answer is after he marks it. Later he may encounter a question that he has already seen and answered incorrectly before, and chances are that he knew the answer this time.

This approach is used by, for instance: proprofs, quizizz, and study.com