APPLICATION OF NATURAL LANGUAGE PROCESSING IN CLASSIFYING EXAMINATION QUESTIONS ACCORDING TO REVISED BLOOM’S TAXONOMY

Kierven A. Villaruela , Dr. Leonardo L. Alejandrob

a Adamson University, Philippines, b Adamson University, Philippines

[kavillaruel.mscsics@gmail.com](mailto:kavillaruel.mscsics@gmail.com), [leonard.alejandro@adamson.edu.ph](mailto:leonard.alejandro@adamson.edu.ph)

ABSTRACT

Examinations play an important aspect in teaching and assessment process. Questions are utilized to evaluate the knowledge and competence of students. Educators in higher education use major examinations to assess learning and skills acquired by students. Whereas, questions provided in an examination must be in different cognitive levels as categorized on the revised Bloom’s Taxonomy. This art of questioning in tertiary level major examination is one of the challenges faced by educators especially those who are not graduates of education programs. This study aims to provide faculty members with an authoring tool that can help them with their tasks in creating examination questions according to the different cognitive levels. This study proposed the application of Natural Language Processing (NLP) in classifying examination questions into their appropriate category based on this Taxonomy. It applied NLP techniques such as word tokenization and stop-word removal to identify important keywords and verbs, which may help in the identification of the category of a question in an examination. Naïve Bayes classifier was also utilized as the text classifier algorithm to create an expert application built in Laravel framework and Python programming language. Results show that using the NLP and Naïve Classifier is a viable approach in classifying questions according to revised Bloom’s Taxonomy. The conducted evaluation of the trained classifier has an accuracy and precision of 97% correspondingly. Applying the trained classifier in the authoring tool, helped the faculty members to identify and classify examinations correctly based on the cognitive levels of Bloom’s Taxonomy.

kEYWORDS:

Natural Language Processing, Naïve Bayes Classifier, Revised Bloom’s Taxonomy, Examination, Laravel Framework, Python

# Introduction

Web-based platforms and technologies provide ease of accessibility in delivering services in different educational institutions. Educational institutions made use of these kinds of technology to provide quality education. Strategic techniques are being developed and employed by educators to utilize computerized systems, online technologies, and mobile applications integrated with expert modules and intelligent systems. Such systems are capable of delivering real-life applications of the teaching-learning process in a physical environment (IGI Global, 2021).

The teaching and learning process is bounded by different activities that can be used to provide and do their tasks to educate their students. One of these activities is the conduct of learning assessments done through the administration of the major examination. Thus, examinations act as a medium for educators to confirm student knowledge and understanding and assess the extent to which students can adapt learning theory and applications (Abduljabbar & Omar, 2015). This particular assessment process is a convention in educational institutions. According to Osman and Yayha (2016), educational institutions should have the extent to develop valuable questions to drive the students' thinking process. Levels of questions in any type of examination are significant as students will tend to think more creatively and divergently when given questions based on higher order thinking skills (Carnegie Mellon University, 2022).

According to Abduljabbar & Omar (2015), Bloom’s Taxonomy developed by Benjamin Bloom and a group of specialist, revised by his students through the leadership of Lorin Anderson and David Krathwol, is one of the hierarchies in educational institutions, which is widely accepted and used as an important framework for teachers to assist them in crafting questions to test different cognitive levels and to ensure student’s cognitive mastery. The revised Bloom’s Taxonomy comprises six cognitive levels arranged in hierarchical form. These levels are (1) Remembering (recalling information); (2) Understanding (understanding the main idea of material heard, viewed, or read); (3) Applying (applying abstract ideas in a concrete situation to solve a problem); (4) Analyzing (breaking down a concept or idea into parts); (5) Evaluating (making informed judgments about the value of ideas or materials) and (6) Creating (integrating together parts of knowledge to form a whole) (Colorado College Website, 2022).

The said framework is being used in conduct of examination as an assessment in Marinduque State College (MSC). Like other educational institutions, it also encounters different problems such as lack of integrated and centralized platform in creating examinations, slow and unorganized management in monitoring, and submission and evaluation of examinations. Additionally, faculty members are having difficulty in classifying examination questions in proper cognitive level especially those who were not education graduates. With the said problems continuously experienced, means of providing the quality examination to students may be highly compromised. Teaching and learning process will not be as effective which may lead to the poor performance and low level of students’ adaptation in learning theories and applications.

Thus, having examination authoring tool with the integration of machine learning and application of Natural Language Processing (NLP), a text mining technique would provide benefits in escalating the learning performance of students and improving the means of educators’ teaching and assessment strategies (studoco.com, 2022). As cited in IBM.com (2020), NLP can enable computers to process human language in the form of text or voice data and to understand its full meaning. It can help perform text classification of examination questions using different processes like word tokenization, normalization and stop-words removal (javatpoint.com, 2021). NLP combined with machine learning will liberate authoring system to have capabilities that can manipulate examination questions driven by human natural language to be automatically classified based on the cognitive levels of the revised Bloom’s taxonomy (Kumara, et.al, 2019). Therefore, an authoring tool can help faculty members of the institute to deliver high quality standards in creating good examinations. Faculty members can have a web-based platform in creating and submitting their examinations. With the help of NLP and machine learning, questions in created examination would be properly classified according to the correct cognitive level of Bloom’s Taxonomy they belong, thus high cognitive levels of learning’s can be provided. Program Heads can be provided a monitoring feature with regards to the submission examinations. With the help of the authoring tool’s expert module, program head would have an effective way of checking the examinations since the system would provide already the statistical analysis of the submitted examination.

# METHODOLOGY

## Requirement Analysis and Design

Preliminary investigation was done thru fact-finding techniques such as interview to the faculty members and staff of Marinduque State College – Institute of Information Systems and Technology. Open ended questions were raised to the Faculty members and the Associate Dean of the Institute to collect pertinent information regarding with their examination system. Their faculty and staff is composed of thirty-four (34) faculty members, two (2) program heads, one (1) associate dean and one (1) technical and support staff. Questions involved were about the existing processes of the organization and underlying problems that they encounter in the creation of examinations. Observation with a group of faculty members was conducted to clearly understand how the examination creation and evaluation process works. Gathered data was used on the conceptualization of the design of the project. These were also used to determine the possible solutions that could solve the problems being encountered. the gathered data and the result of the requirements analysis was used to developed different models. The Data Flow Diagram of the Existing System was illustrated to enable the proponents to properly view the flow of data in every process. Use Case Diagram was also illustrated to demonstrate how the different users will interact with the proposed system. This diagram was able to define and organize the functional requirements of the system andmodel the basic flow of events in a use case.

## Data Processing & Classifier Training

The gathered data was subjected to different data processing phases in order to have best results when injected to the machine learning algorithm. These phases were presented in Figure 1 below.

Question Collection

Question Cleaning

Verification of Question Classification

Question Preprocessing

Tokenization

Stop Words Removal

Keyword Extraction

TFIDF

Classification

Naïve Bayes

Classifier Evaluation

Question Text File

(Data Set)

Figure 1: Phases of Data Preprocessing and Classifier Training

*Question Collection.* Questions were taken from the examination papers of the Institute of Information Systems and Technology. The questions were consolidated from the year 2016 to the year 2020 with a total of 4, 450. These questions were taken from random general education and major subjects under the BS Information Technology and BS Information Systems. See Table 1 for the breakdown of data.

**Table 1: Distribution of Questions per Year**

|  |  |
| --- | --- |
| **Year** | **No. of Questions** |
| 2016 | 399 |
| 2017 | 634 |
| 2018 | 874 |
| 2019 | 1182 |
| 2020 | 1361 |
| **Total** | **4450** |

*Question Cleaning and Manual Classification.* This step was performed to ensure that all questions conforms with the different levels of the revised Bloom’s Taxonomy. Specific verbs and keywords used in the questions were identified to determine whether they satisfy with the cognitive classification they belong. After, removing inconsistencies on the questions, less than 1000 records have been identified and labeled correctly. Type of tests helped to easily classify some questions. According to studoco.com (2022), there are types of tests that can be automatically classified according to cognitive levels of Bloom’s Taxonomy presented in Table 2.

**Table 2: Type of Test per Cognitive Levels of Bloom’s Taxonomy**

|  |  |
| --- | --- |
| **Cognitive Level** | **Type of Tests** |
| Remembering | Identification, Fill in the Blanks, Enumeration |
| Understanding | Matching Type, True or False |
| Applying | Application |
| Analyzing | Analogy |
| Evaluating | Critiquing |
| Creating | Illustration, Composition |

The types of test in Table 2, were the commonly found on the gathered examination. Multiple choice and essay types of test as commonly used in examination are not present in the table above because these types of tests may fall in any cognitive level of the taxonomy depending on the keywords used and the questions’ construction. Some of the following keywords utilized in the examination questions were summarized below in Table 3. Using the keywords in the table below, multiple choice and essay type questions were easily categorized to what level they should be long.

**Table 3: Commonly used Keywords per Cognitive Level**

|  |  |
| --- | --- |
| **Cognitive Level** | **Keywords Used** |
| Remembering | Cite, Definition, Define, Term, Refers, Identify, Describe, Label, Name, Recall, Select, State, Call, Known, Who, Correct, Different |
| Understanding | Associate, Classify, Compare, Infer, Explain, Generalize, Outline, State, Summarize, Say About, Observe on, Restate, Meaning, Main Idea |
| Applying | Develop, Alter, Apply, Change, Construct, Demonstrate, Execute, Implement, Modify, Present, Solve, Debug, Examples, Utilize, Find, Choose, Run, Result, Output, Codes |
| Analyzing | Differentiate, Examine, Investigate, Motive, Dissect, Relationship, Simplify, Discover, Analyze, Inference, Inspect, Categorize, Conclusion, Related, Scrutinize, Study, Resolve, Distinguish, Motive, Divide, Compare, List, |
| Evaluating | Assess, Discuss, Recommend, Convince Defend, Prioritize, Check, Classify, Justify, Conclude, Criticize, Critique, Defend, Importance, Opinion, Disprove, Agree, Support, Convince, Evaluate, Figure Out, Why, Criteria, Assess, Prove, Gauge, Decide, Influence, Judge, Value, Rate, Conclude, Measures, |
| Creating | Formulate, Imagine, Generate, Invent, Happen, Devise, Brainstorm, Solution  Forecast, Result, Design, Propose, Suggest, Changes, Make, Portray, Predict, Improve, Create, Combine |

After cleaning the questions, the dataset was trimmed down to 200 records of questions per cognitive levels of the revised Bloom’s taxonomy. The proponent subjected the data for secondary validation. Subject matter experts were tasked to evaluate the identified cognitive levels of the exam questions. SME’s did the evaluation manually by reading and examining the questions one by one. Based on the result, 94% (or 1,128 questions) were correctly classified and 6% (or 72 questions) were incorrectly classified. The incorrectly classified questions were revised, others were replaced in order to equalize the number of questions per cognitive level, with 100 questions each. This data was randomly divided into training data (80%) and testing data (20%).

*Question Preprocessing.* After splitting*,* the dataset in excel format was converted to a text file (than can be called as a corpus), so that the NLTK package of python can easily read each question. At first, the corpus underwent the process of word tokenization (ExpertSystem.com, 2016). Tokenization is a procedure in which the questions as a long string of texts are broken into smaller pieces or tokens. After that, normalization was also involved in removing punctuation marks and converting the corpora to lower case. Stopwords removal was also conducted. This step is the process of removing the words that commonly occur across all the questions in the corpus (Algarni, 2016).

­

*Keyword Extraction.*  This process was conducted in order to determine how relevant the keywords are in question. This was performed using the. TF-IDF (term frequency-inverse document frequency). TF-IDF works by increasing proportionally to the number of times a word appears in a question, but is offset by the number of question that contain the word. For example, if the word “definition” appears many times in questions under a single cognitive level, while not appearing many times in others, it probably means that it’s very relevant (CapitalOne.com, 2021). TfidfVectorizer was used as the algorithm to transform text into a meaningful representation of numbers which is used to fit machine algorithm for prediction. This transformed the preprocessed tokens in the vectors so that it can used as input to the machine learning algorithm to start the training of the Naïve Bayes Classifier.

*Classification.* Naïve Bayes classifier is a probabilistic machine learning model that’s used for classification task. Since the classification of exam questions according to the levels of Bloom’s taxonomy is considered as text classification, the Multinomial Naïve Bayes was used. This classifier is based on the Bayes Theorem which predicts the tag of a text or token. When applied to an individual question it calculates the probability of each token or keyword for a given sample question and then gives the token with the highest probability as output (Abduljabbar & Omar, 2015). The vectorized corpus of the preprocessed dataset of questions, was exposed to the MultinomilaNB classifier to start training the said algorithm. With that, the classifier would be familiarized on how the questions are constructed. Keywords and patterns would be recognized in all cognitive levels of the Bloom’s taxonomy. As a result, the classifier would be able to predict or determine the classification of new questions that would be applied to it.

*Classifier Evaluation.* To measure the performance of the classifier, the proponent used confusion matrix as an evaluator. This matrix is composed of information on the actual (column) and predicted (row) by the Naïve Bayes Classifier. This would determine if the trained classifier can correctly classify new questions that would be given to. Table 4 represents the metrics that can be used to balance the classifier estimates. It is composed of the Accuracy, Precision, Recall and F1-score.

**Table 4: Evaluation of the Naïve Bayes Classifier**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Accuracy** | **Precision** | **Recall** | **F1-Score** |
| **Training Set** | 0.979058 | 0.979186 | 0.979198 | 0.979085 |
| **Testing Set** | 0.949791 | 0.951295 | 0.950050 | 0.949228 |

The value of the accuracy of the dataset, describes the number of correct predictions over all predictions. Precision on the other hand, measures the number of positive predictions that are correct while recall measures the number of positive cases the classifier correctly predicted, over all the positive cases in the data. F1-Score represents the harmonic mean of the precision and recall. Thus, the classifier has 96% chance of correctly classifying new questions according to the levels of Bloom’s taxonomy, given that newly created questions has similar sentence constructions with that of the datasets used in training the classifier algorithm.

## Development and Testing

After the data processing, actual development started by creating the web-based system. Sublime Text editor was used as the main application in coding the program using the various programming language such as PHP: Hypertext Pre-processor 7, and JavaScript. Laravel framework was also used in making the web interface. It served as a great tool in providing features like inheritance and displaying data that added flexibility to the web development process. Jupyter Notebook was also utilized in implementing the Natural Language Processing pre-processing and the Naïve Bayes classifier. Python programming language was used and its capabilities served its purpose in simulating machine learning process on the prepared dataset. All data processing that occurred in the system was managed using XAMPPserver, a database management system employed by MySQL structured query language to store, retrieve and process information.

After the development phase, testing was conducted. The system underwent cross browser compatibility test to verify if it works across different browsers as it is expected, ensuring that the system would not have behavioural errors. The system was also subjected for stress testing to verify the stability and reliability of the system. This test determined the system’s robustness and error handling under extreme heavy load conditions. Additionally, assessing the acceptability of the system was the last part of the testing phase. Acceptability was ensured by conducting user acceptance testing. User acceptance testing (UAT) was conducted to the end-users; the target users of the system verified the systems workability meeting the provided system requirements and user expectations. Aside from UAT, software quality evaluation based on ISO/IEC 25010 quality model was conducted. This software evaluation assessed software criteria such as functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability and portability. A devised questionnaire containing the above criteria was prepared and results was measured using the Likert Scale (Table 3.5) rating basis. The conduct of this software quality evaluation will be administered by IT experts and professionals.

**Table 5: Likert Scale Rating Basis**

|  |  |
| --- | --- |
| **Weighted Mean** | **Adjectival Interpretation** |
| 5 | Very Strongly Agree |
| 4 | Strongly Agree |
| 3 | Agree |
| 2 | Disagree |
| 1 | Strongly Disagree |

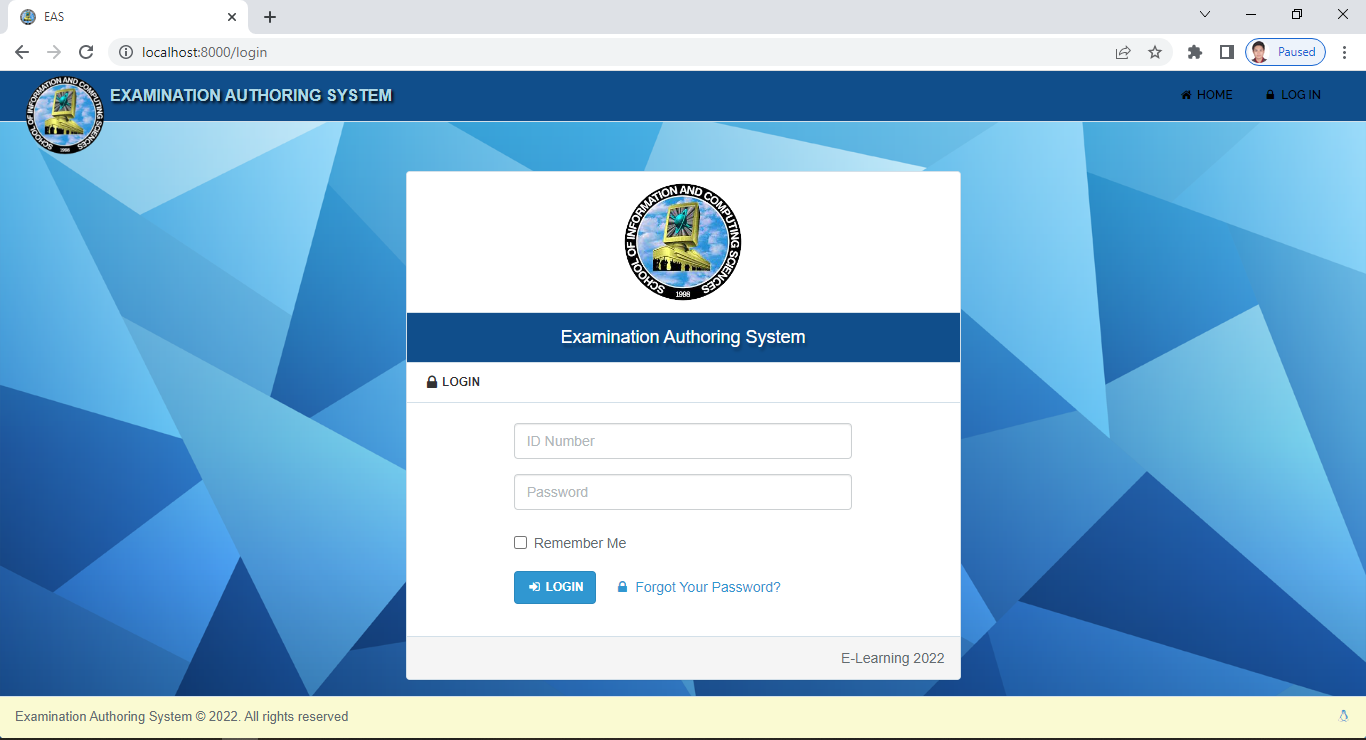
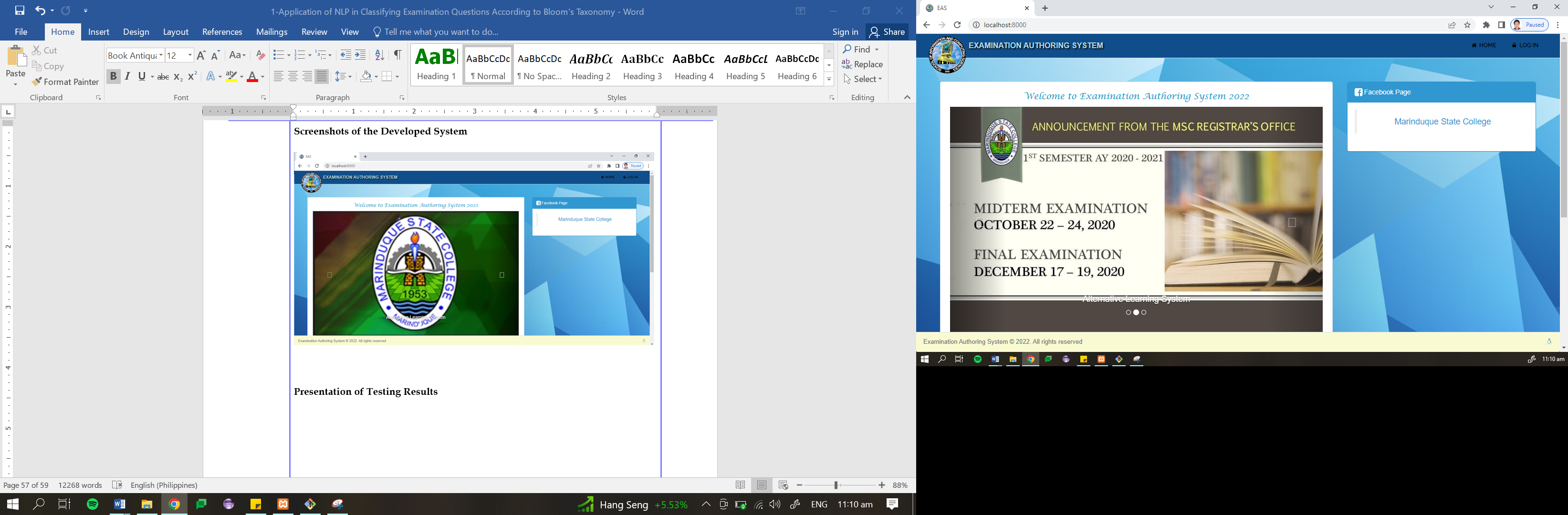
Table 5 shows the rating basis that will be used in the software quality evaluation based on the ISO/IEC 25010 quality model.

# RESULTS AND DISCUSSIONS

## Technical Description

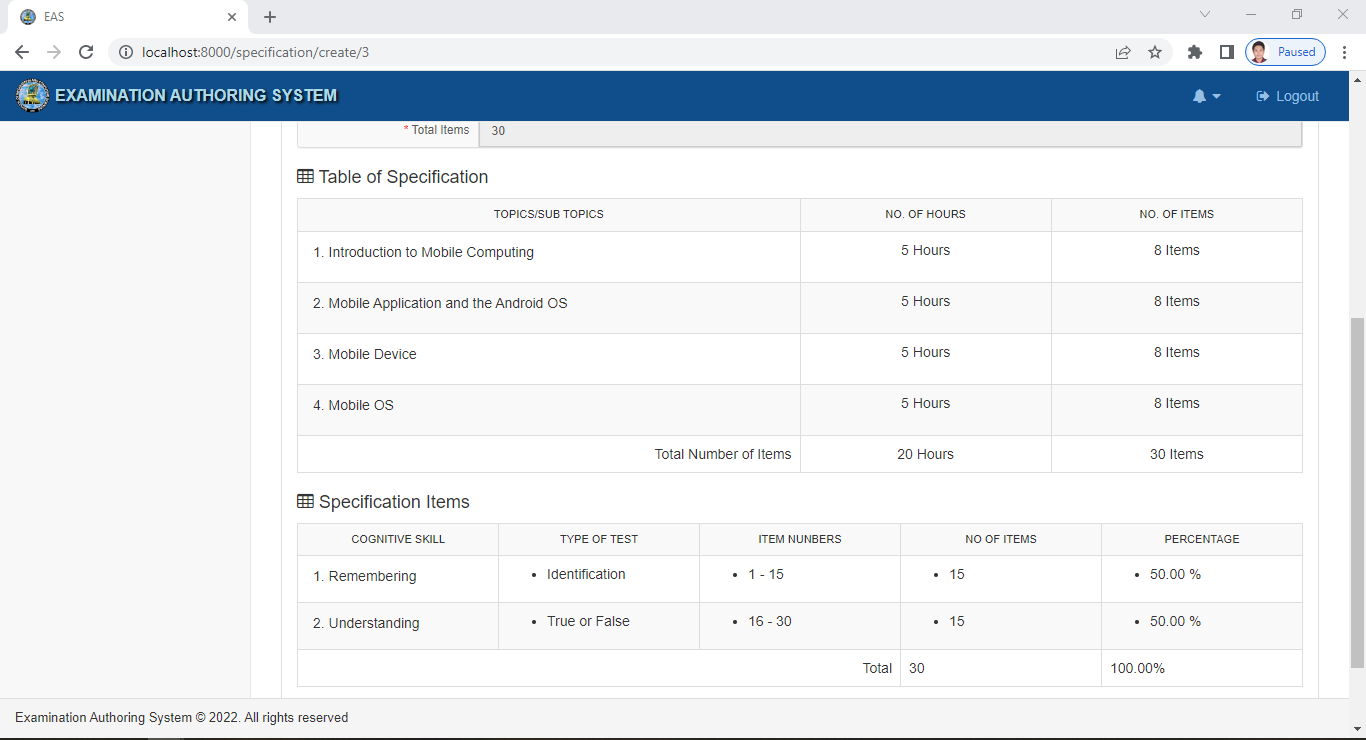
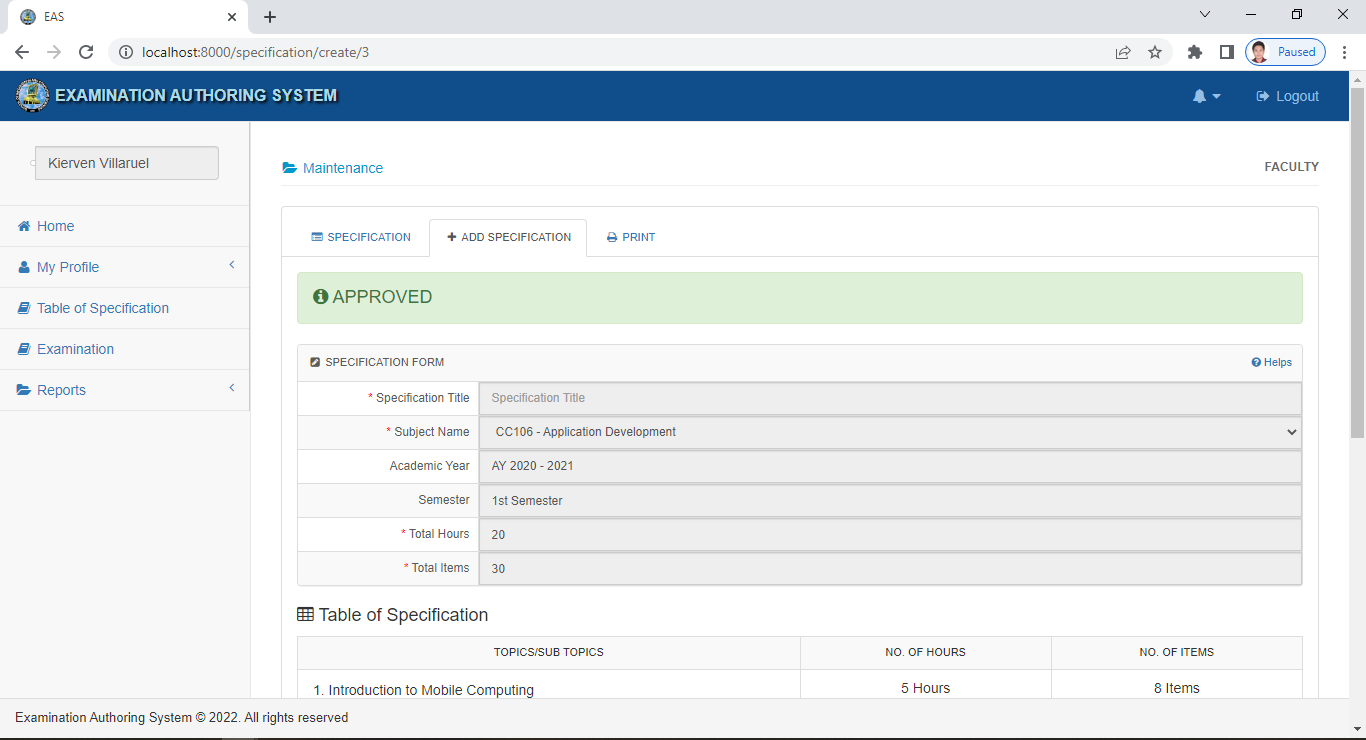
The Examination Authoring System as the product of the Application of Machine Learning and Natural Language Processing in Classifying Examination Questions according to Revised Bloom’s Taxonomy presents an interactive system that can accommodate faculty members of the MSC – Institute of Information Systems and Technology, in creating Table of Specification and Major Examination in a centralized web-based platform. The following are the functional specifications of the developed system based on the projects’ objectives:

### Web-based Authoring System



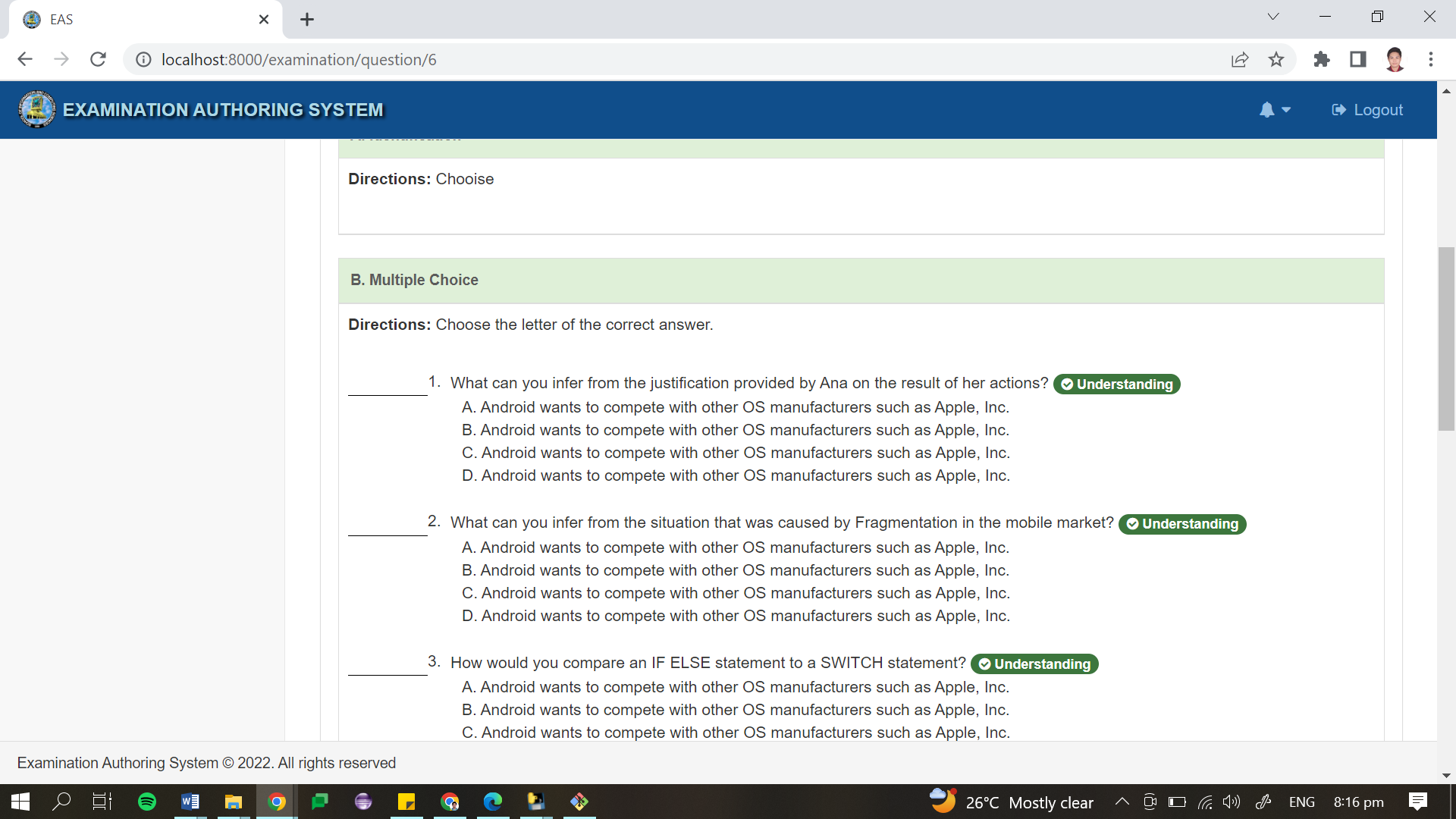
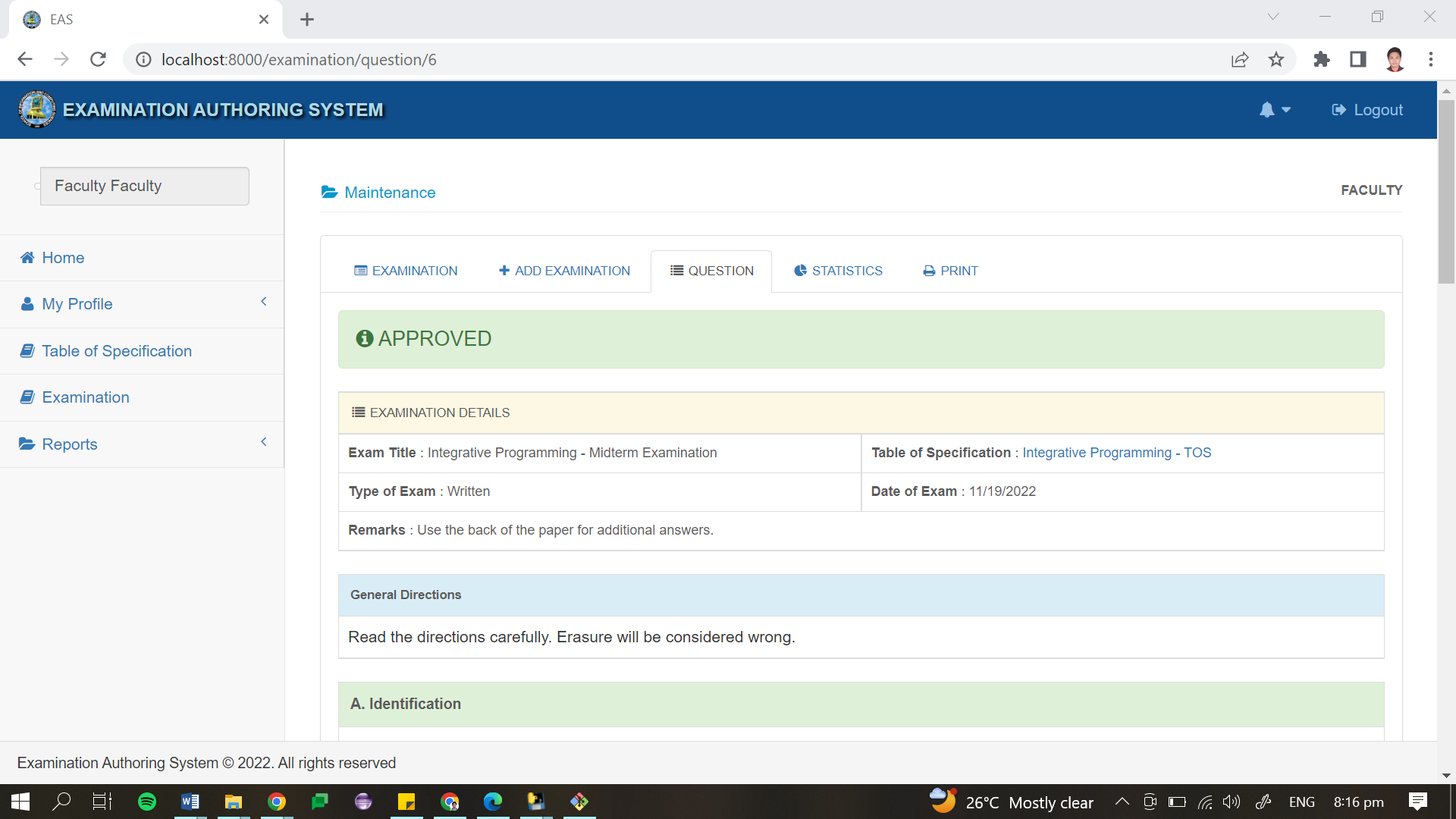
**Figure 2: Landing pages of the Web-based Authoring System**

The Web-based system is capable of authoring two important documents which are the Table of Specification and the Major Examination. It can help create the said documents by having a standardized web interface. As an authoring tool, the system follows the prescribed design and layout of the output document. Faculty members who are the primary user of the system can experience automated process of creating their TOS and examination. The system provides faculty members to conveniently add topics that are included in the TOS. It can automatically compute the number of items per particular topic based on the number of hours entered in the TOS. As for the type of test that will be included in the examination, faculty members can add type of tests that is applicable to the cognitive skills of Bloom’s Taxonomy.



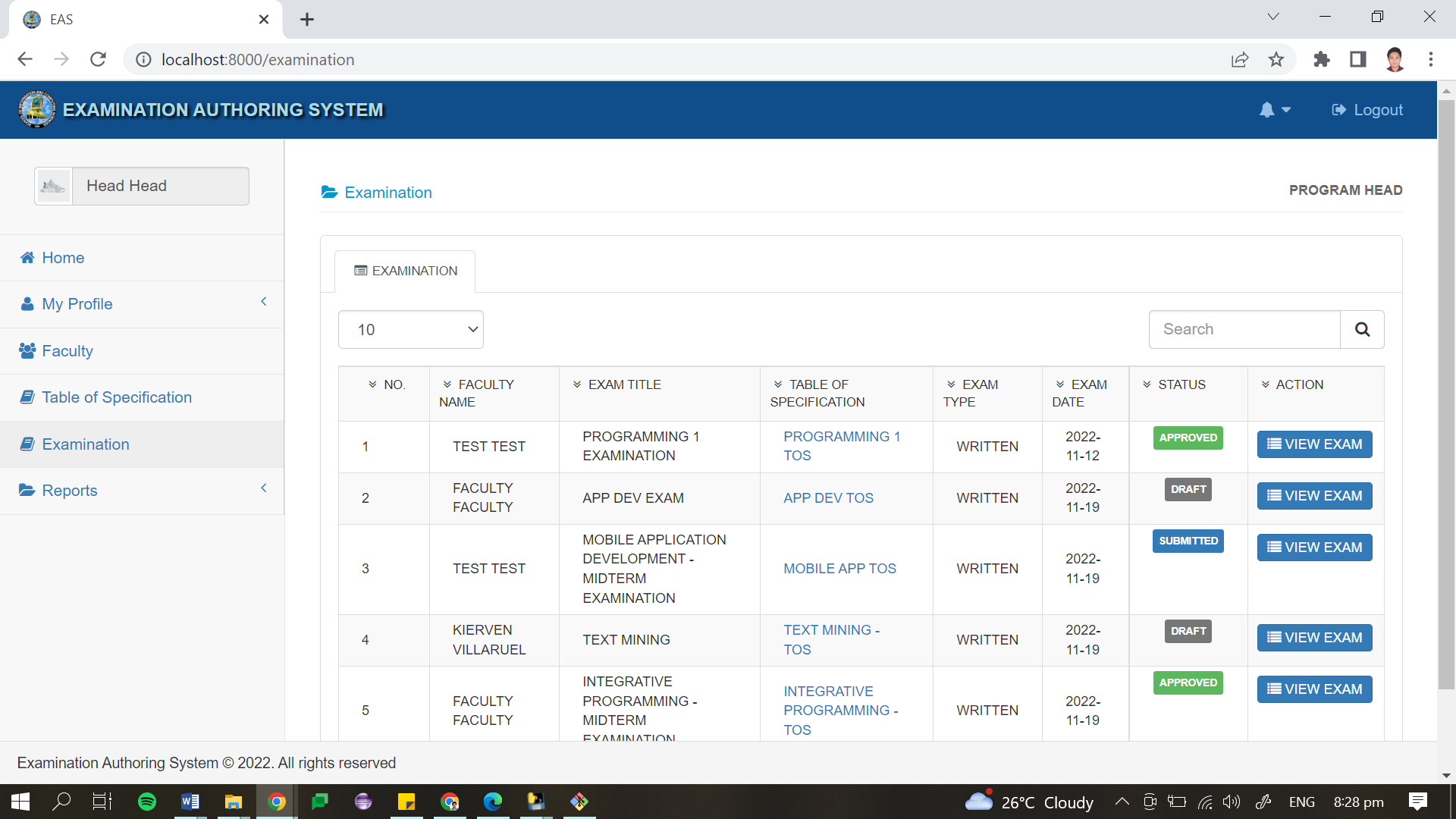
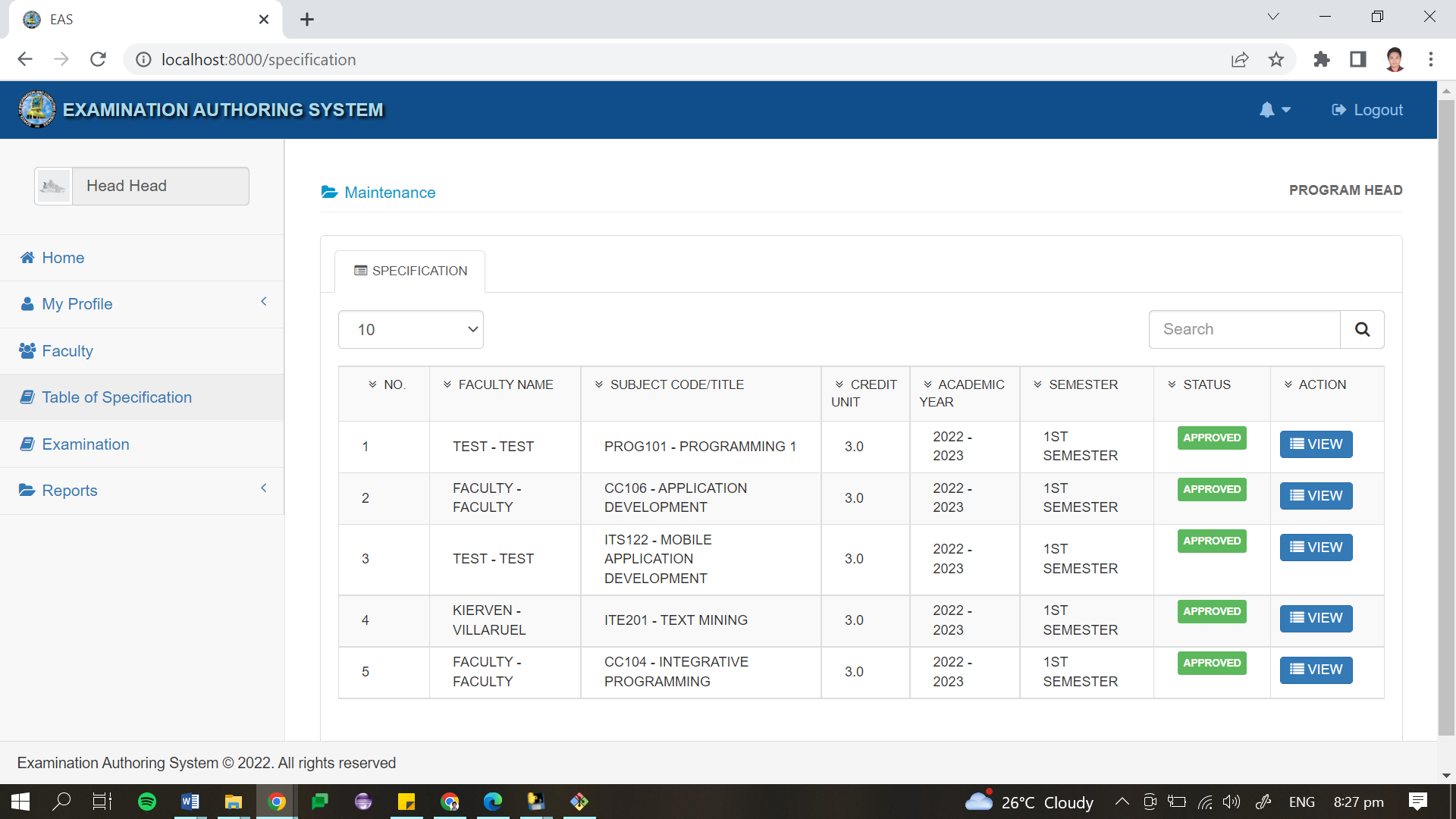
**Figure 3: Creation of Table of Specification Interfaces**

Automatic computation of the number of items per type of tests and its percentage equivalent can be also found in the system. Once the created TOS is approved, faculty members may start creating their examination. In creating examination, only the type of tests present in the created TOS shall be added in the exam. Several types of test are already classified according to the proper cognitive level. Questions under the type of Multiple Choice and Essay are the once who would be automatically classified to what cognitive level they belong. The system provides a facility to edit or reconstruct the created questions. Proper formatting of type of tests are also provided in the system. The system also provides a facility to print the created TOS and examination.



**Figure 4: Creation of Examination Interfaces**

### Management of TOS and Examination



**Figure 5: Monitoring Interface of TOS and Examination**

The created TOS and examination of faculty members can be submitted to their respective Program Heads. Program Heads as secondary users can access the system for the monitoring of TOS and examination submissions. The system provides a facility to approve and disapprove for revision the submitted TOS and examination. A list indicating names of the faculty members who have submitted their TOS and examination is provided in the system. Different status on creating TOS and examination is also provided such as Draft, Submitted and Approved. This functionality can help program heads determine faculty members who are not yet creating, and or submitting their TOS and examination.

### Question Classification

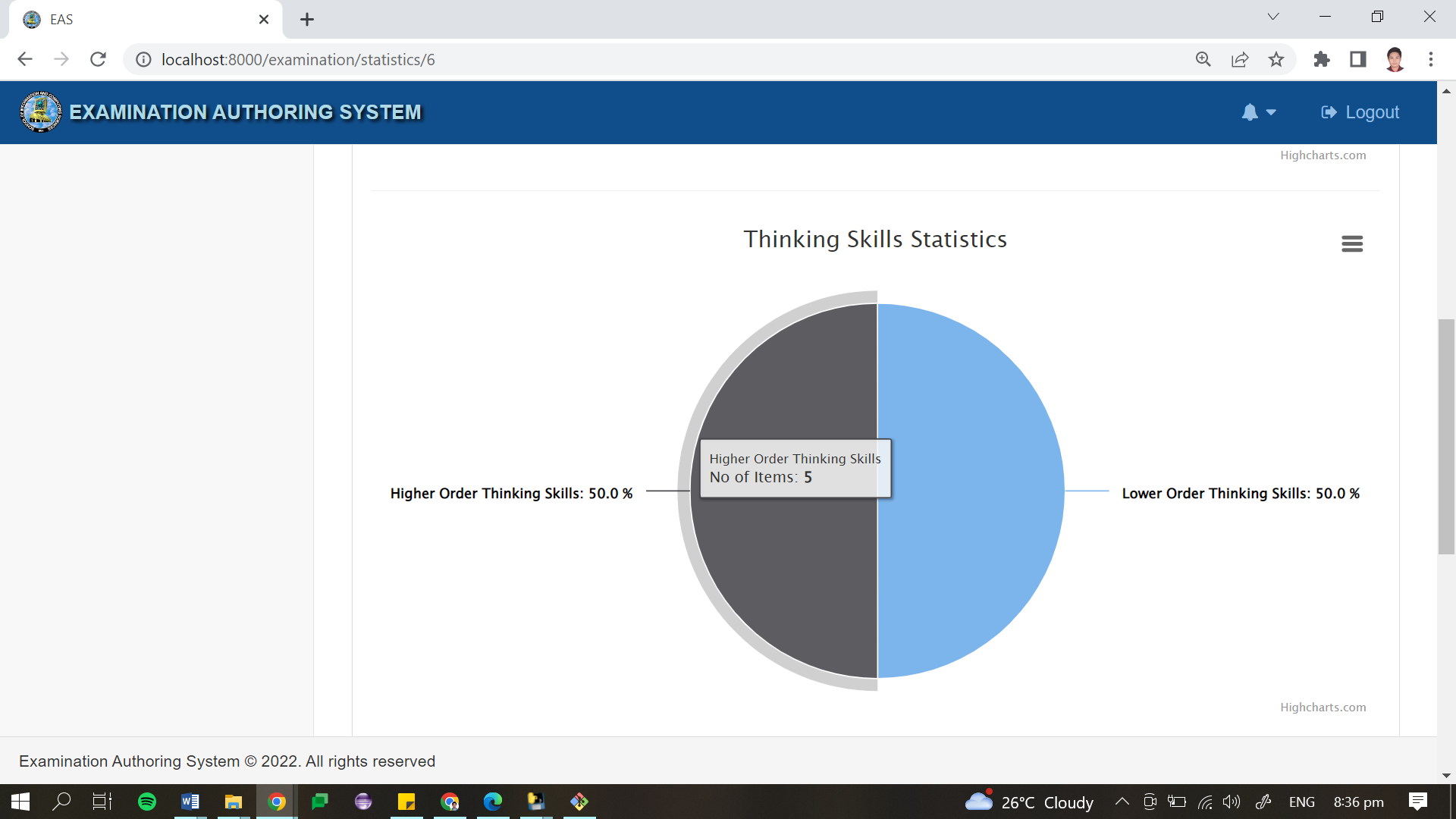
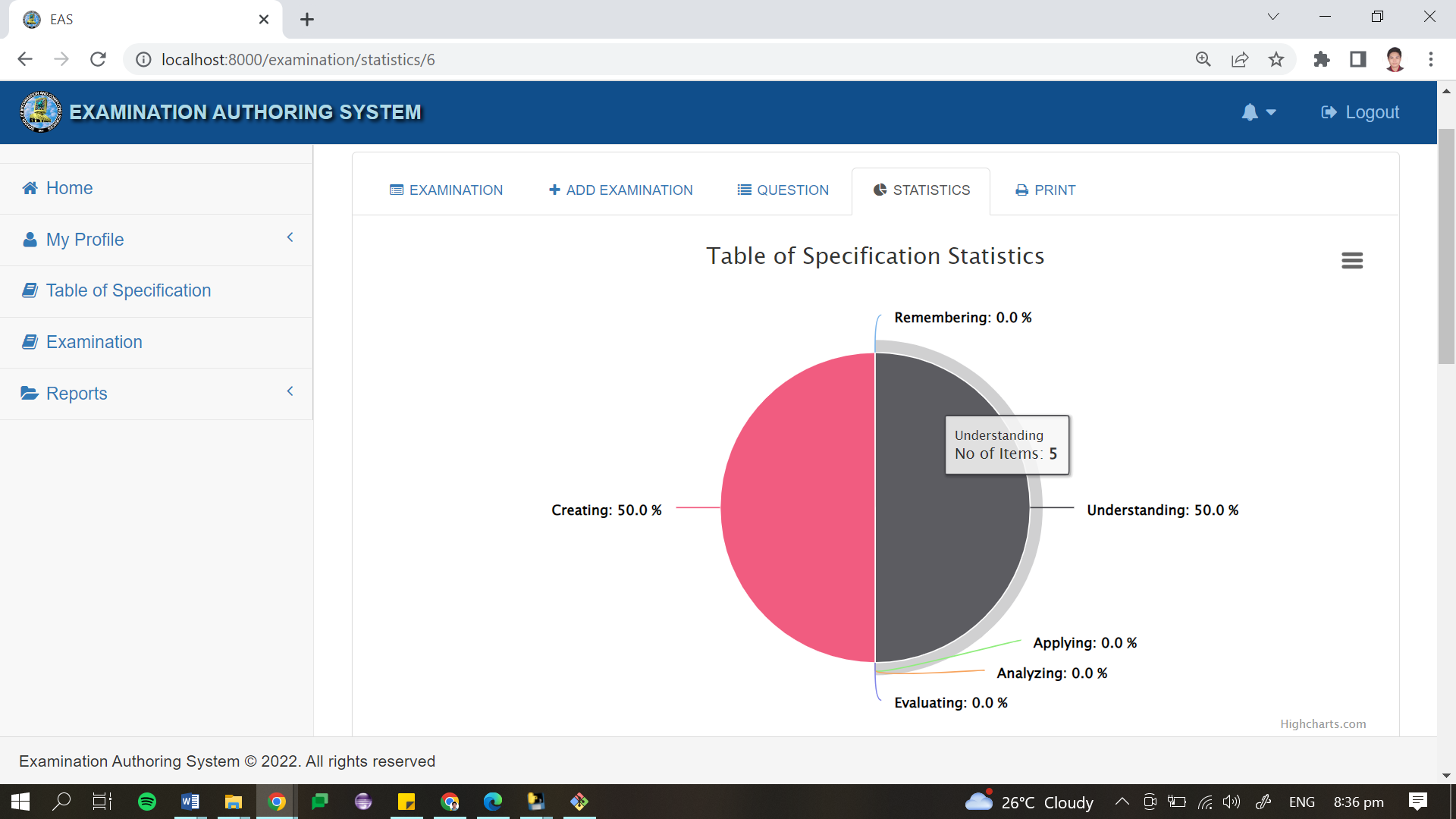
The question classification is the main feature of the system. It was made possible by utilizing machine learning techniques and natural language processing. This function is employed using Python programming language. Script that holds its program is being called inside the web-based system created with Laravel framework. The python script is powered by a naïve bayes classifier caller MultinomialNB(). It is a powerful algorithm that is for text data analysis and classification. Since it implements the Bayes Theorem, it uses the probability to determine a possible prediction or classification of newly created questions. This classifier was trained using question dataset that underwent text preprocessing using NLP. The training set of the question dataset has an accuracy of 97%. When tested, it provides correct classification of the questions based on the cognitive level of Bloom’s Taxonomy. Question Classifier is visible on the system when faculty members create questions for their exam. It will automatically provide a status of their classification beside the created questions.

### Exam Question Validation

This validation functions refers only with the process of checking the creation of exam based on what is provided in the created TOS. The system shall verify whether a particular question with its classification is parallel on the item numbers in the TOS. The TOS shall serve as the guidelines in providing the questions as entries in the examination.

### Reports Generation and Visualization

Examination reports and statistics are presented in the system. It consists of graphical representation depicting the percentage distribution of the number of questions classified according to the levels of Bloom’s Taxonomy. Exam reports also includes the categorization of the examination questions for the Higher Order Thinking Skills and Lower Order Thinking Skills.



**Figure 6: Examination Statistics Visualization**

## Test Results

Testing and evaluation procedures were conducted after the development of the final prototype. The developed examination authoring system underwent cross browser compatibility test. It was accessed using Google Chrome, Microsoft Edge, Internet Explorer and Mozilla Firefox. From the different browsers used, Google Chrome was the most compatible as the system did not experience any software malfunction in the web interface and system functionalities. Meanwhile, when tested with other browsers, like Internet Explorer, some pop-up menus and JavaScript prompts experienced irregularities due to outdated browser version and blocked web technology extensions. Stress testing was also conducted to verify the system’s robustness and error handling. This testing procedure determined possible error occurrence if invalid and out of range inputs were entered into the system. This helped the developer to plan accordingly the response that the system would display if such errors happen unexpectedly.

User Acceptance Testing were also conducted. The system was presented to the Associate Dean, Program Heads and faculty members of the Institute of Information Systems and Technology. The system was made available via local area network using XAMPPServer remote configuration. UAT forms were given to program heads and faculty members to gather their acceptance, views, comments and recommendation when it comes to the functionalities of the system. Based from the result, most of the retrieved forms, gathered positive outcome as most of the faculty members strongly agreed that the system has executed all the functions that were tested. Hence, the acceptance of the system’s features and function is strongly agreeable. While some faculty members provided comments on some functionalities, the developer took these comments for the improvement of the functions of the whole system.

After UAT, the system was also subject to quality evaluation. The developer used an evaluation form which sample is presented in Appendix I. Results of the conducted evaluation based on the ISO/IEC 25010 is detailed in Table 4.15.

**Table 6: Software Quality Evaluation Result of the Developed System**

|  |  |  |
| --- | --- | --- |
| **System Attribute** | **Mean** | **Adjectival Interpretation** |
| Functional Suitability | 4.7 | Very Strongly Agree |
| Performance Efficiency | 4.8 | Very Strongly Agree |
| Compatibility | 4.7 | Very Strongly Agree |
| Usability | 4.7 | Very Strongly Agree |
| Reliability | 4.6 | Very Strongly Agree |
| Security | 4.6 | Very Strongly Agree |
| Maintainability | 4.8 | Very Strongly Agree |
| Portability | 4.8 | Very Strongly Agree |
| **Grand Mean** | **4.7** | **Very Strongly Agree** |

This table displays the overall result of the software evaluation. The testers evaluated the system with a mean of 4.7 for Functional Suitability, Compatibility, and Usability. The mean of 4.8 was given to Performance Efficiency, Maintainability and Portability. On the other hand, Reliability and Security got the lowest mean of 4.6. These mean that the evaluators are in agreement with the overall quality of the system. The resulting grand mean of 4.7 is interpreted as “Very Strongly Agree”, which means that the system suits to what are being expected by the respondents.

# CoNCLUSION AND RECOMMENDATIONS

## Conclusion

After conducting series of tests, the system has proven its capability to improve the existing system used by the MSC – Institute of Information Systems and Technology. All objectives of the project have been met including the creation of table of specification and examination in the system wherein these two related documents are placed in one centralized platform. The authoring system has also the capability to monitor submission and checking of TOS and examination. The best thing about this system, is that it can classify questions on created examination based on the cognitive level of the revised Bloom’s Taxonomy which is a significant feature that can eliminate problems experienced by faculty members in making good examinations. And lastly, it is a great help for faculty members to see the statistics of their examinations through graphical representation that will enable them to have better outlook in proving better assessment tools for their students.

Therefore, it is concluded that the developed system would be a great help to the faculty and staff of MSC – IIST, in upgrading their ways on how to do things and in solving encountered problems.

## Recommendations

Having the summary of findings and conclusion as basis, the following recommendations were drawn;

* Upon using the system, make sure to have the required minimum hardware and software requirements;
* The exam authoring system would function better if uploaded to a dedicated server.
* The system must be prepared to be connected with the other modules of a bigger system to serve better purpose.
* Faculty members should have familiar with the Art of Questioning.
* It also recommended to try improving the NLP process with deeper and high level approach.
* Other classification algorithm like Support Vector Machine and Voting Algorithm may be also used to compare the accuracy in classifying examination questions according to the revised Bloom’s Taxonomy.
* Plenty and diverse collection of data must be used in training the classifier.
* Lastly, it is recommended to test the system into a bigger number of possible users to measure the full potential of the question classification.

# REFERENCES

IGI Global: Publisher of Timely Knowledge. (2021, October). What is Higher Education Institution?. Retrieved from <https://www.igi-global.com/dictionary/inciting-grassroots-change/13097>

Abduljabbar, D.A. & Omar, Nazlia. (2015). Exam questions classification based on Bloom’s taxonomy cognitive level using classifiers combination. *Journal of Theoretical and Applied Information Technology 78*. 447 – 455

Osman, A. & Yahya, A. (2016). Classifications of Exam Questions Using Linguistically-Motivated Features: A Case Study Based on Bloom's Taxonomy.

Carnegie Mellon Univeristy. (2022). Why should assessments, learning objectives, and instructional strategies be aligned?. Retrived from <https://www.cmu.edu/teaching/assessment/basics/alignment.html>

Colorado College. (2022). Bloom’s Revised Taxonomy: Using Bloom’s Revised Taxonomy in Assessment. Retrieved from <https://www.coloradocollege.edu/other/assessment/how-to-assess-learning/learning-outcomes/blooms-revised-taxonomy.html>

Studoco. (2022). Use Bloom’s Taxonomy to Align Assessments. Retrived from <https://www.studocu.com/en-us/document/university-of-the-people/online-education-strategies/blooms-taxonomy-handout-for-critical-thinking-skills/>

IBM. (2020, July). Natural Language Processing. Retrieved from <https://www.ibm.com/cloud/learn/natural-language-processing>

Javatpoint. (2021). NLP Tutorial. Retrieved from <https://www.javatpoint.com/nlp>

Kumara, B.T.G.S., Brahmana, A., & Paik, I. (2019, June). Bloom's Taxonomy and Rules Based Question Analysis Approach for Measuring the Quality of Examination Papers. International Journal of Knowledge Engineering, Vol. 5, No. 1 doi: 10.18178/ijke.2019.5.1.111

ExpertSystem.com. (2016, September). Natural Language Processing Applications. Retrieved from <https://expertsystem.com/natural-language-processing-applications/>

Algarni, Abdulmohsen. (2016). Data Mining in Education. International Journal of Advanced Computer Science and Applications. 7. 10.14569/IJACSA.2016.070659.

CapitalOne.com. (2021, October). Understanding TF-IDF for Machine Learning. Retreived from <https://www.capitalone.com/tech/machine-learning/understanding-tf-idf/>