An Analysis of AI Term Projects for Undergraduate Students

Sang-Hoon Oh*
*Mokwon University, Korea
E-mail : ohsanghoon16@gmail.com

1. Introduction

Artificial Intelligence (AI) is emerging as a key technology pioneering the future. For undergraduate students, universities are strengthening AI education in both general and major areas. In general education course at Mokwon university, undergraduate students study “Introduction to AI”, “Basic AI programming”, “AI experiments”, and “Applications of AI”. Since undergraduate students in general education course do not have concrete mathematics knowledge or programming skills to implement AI models or to process data, it is very difficult for them to do term-projects for real world application of AI. On the contrary, undergraduate students in major course have successfully completed rigorous education course such as programming skills and mathematics including derivatives and linear algebra, they can study theories and programming skills of AI. After successful completeness of AI theories and programming, it is essential for the undergraduate students in major course to do some term-projects for real-world applications. Students need to undertake AI term projects that solve real-world problems to solidify their knowledge, enabling them to apply AI effectively when they enter the workforce after graduation [1].

This study introduces the AI programming lecture as part of the major course and reports the shortcomings revealed in the term projects conducted over four years at Mowon university. This report indicates that which parts of the AI programming lecture should be strengthened to improve the execution of AI term projects.

2. AI programming lecture with term projects

At mokwon university, AI programming lecture consists of reviewing programming language, data preprocessing, AI programming from scratch, and Tensolrflow to implement deep neural networks. Firstly, the review of programming language focuses on Numpy, Pandas, and Matplotlib for processing data and data visualization. Secondly, data preprocessing is practiced using the California housing prices dataset[1]. This part covers data structure, histograms, correlation coefficient, scatter matrix, data celansing (handling missing values or outliers, removing duplicates), data transformation, and data scaling. Thirldy, AI programming from scratch is done with the MNIST dataset[2], which is a handwritten digit dataset with 60,000 traning images and 10,000 test images sized 28x28. In the third part, students program template matching, k-nearest neighbor(k-NN) algorithm, perceptron, mulitlayer perceptron, and CNN(convilutional neural network) from scratch [3][4]. Additionally, SVM(support vector mahine) and dimensionality reduction, including PCA(principal component analysis) and LDA(linear discriminant analysis), are conducted with MNIST dataset. Finally, Tensorflow is used to implement CNN and LSTM(long short-term memory) networks[5].

There is also a parallel process for conducting term-projects. Term-project groups are organized into 2 or 3 students. They select real-world application problems from public data repository such as Kaggle and Amazon datasets. Afterward, they report their activities for the term projects every week. The reports include the project’s goal, data structure, data preprocessing, AI model implementation, and learning results. At every step of the project, data visualization is an essential skill for easily understanding data snalysis and learning results. Fig. 1 shows the general workkflow of machine learning projects. Any unexpected results at every step must be followed by feedback to the appropriate previous step.

3. An Analysis of Term Projects

Let me explain the shortcomings revealed in the term projects conducted by ninety-six students over four years at Mowon university. There were thirty-five term projects conducted during last four years, from 2021 to 2024 school years. Twenty-one classification problems and fourteen regression problems were tackled, with only two projects failing. Eighteen projects were categorized as excellent, ten as ordinary, and five as unsatisfactory.

The shortcoming revealed in the term projects appears throughout the entire process, from data analysis to visualizing learning results, despite the comprehensive coverage of these topics in the AI programming lecture. In the first step of selecting a dataset to address with AI techniques, students struggle to explain the meaning of each data column in the selected dataset as well as the data structure. Additionally, data analysis process such as PCA, LDA, calculating correlation coefficient, scatter plots, and histogram are challenging for the students. The next step involves selecting an AI model and deciding on its structure appropriate to the dataset. Students must determine the parameters of the AI model, but they often show weaknesses in explaining how to decide parameter values through a reasonable process. After successfully training AI models, it is necessary to display the learning curve of performance index versus

learning epoch. It is also important to compare performances of various AI models. Furthermore, selecting an appropriate performance index for AI application problems is crucial for reporting the final learning results. These shortcomings should be addressed and strengthened during the AI programing lecture.

4. Discussion and Conclusion

In this presentation, we introduced the AI programming lecture, including term projects. To acquire the essential skills needed to solve real-world AI application problems, students went through many steps. However, they revealed shortcoming during conducting term projects. Finally, the shortcomings were analyzed.

5. References

[1] Aurelien Geron, Hands-On Mavchine Learning with Scikit-Learn, Keras & Tensorflow, O'Reilly Media, Inc., 2019.
[2] P. Grothet, "NIST special database 19 handprinted forms and characters database," Tech. Rep., National Institute of Standards and Technology, 1995.
[3] S. Raschka and V. Mirjalili, Python Machine Learning Third Edition, Packt, Birmingham-Mumbai, 2019.
[4] S. Goki, Deep Learning from Scratch, O'Reilly Japan, Inc. 2017.
[5] I.-S. Oh and J. Lee, Artificial Intelligence with Python, Hanbit Academy, Inc., 2021.

Figure 1. Workflow of machine learning