2024 HAWAII UNIVERSITY INTERNATIONAL CONFERENCES
SCIENCE, TECHNOLOGY & ENGINEERING, ARTS, MATHEMATICS & EDUCATION JUNE 6 - 8, 2024
PRINCE WAIKIKI RESORT, HONOLULU, HAWAII

NON-PROGRAMMING ARTIFICIAL INTELLIGENCE EDUCATIONAL PRACTICE USING FREE CLASSIFICATION SERVICES

MIYAMOTO, YUKINOBU
DIVISION OF INFORMATION SYSTEMS
OTEMON GAKUIN UNIVERSITY
IBARAKI, OSAKA,
JAPAN

Dr. Yukinobu MIYAMOTO
Otemon Gakuin University
Ibaraki, Osaka
JAPAN

Non-programming Artificial Intelligence Educational Practice using Free Classification Services

Synopsis:

In this paper, we present a practical example of a non-programming data analysis class for social science students. In the practice, three types of analysis methods were presented to the students, and they developed exercises in which they analyzed and discussed the data to the extent that they were able to do so. The results and effects are considered in a setting that allows students to recognize their own skill level and produce satisfactory results to the extent possible.

Non-programming Artificial Intelligence Educational Practice using Free Classification Services

Yukinobu MIYAMOTO
Otemon Gakuin University

Abstract

In recent years, Japan has been gradually building a platform for mathematics, data science (DS), and artificial intelligence (AI) education at universities. On the other hand, the reality is that children of the next generation are beginning to learn DS education ahead of their peers, as evidenced by the GIGA School Initiative and other initiatives to make information education compulsory in primary and secondary education. To bridge this generation gap, there is an urgent need to promote DS education at universities, but not necessarily only students who have mastered programming are being educated.

In this paper, we present a practical example of a non-programming data analysis class for social science students. Students in the so-called humanities and social sciences are not necessarily good at mathematics or information-related subjects, or have not yet learned them, so the prerequisite knowledge level cannot be set very high. In the practice, three types of analysis methods were presented to the students, and they developed exercises in which they analyzed and discussed the data to the extent that they were able to do so. The results and effects are considered in a setting that allows students to recognize their own skill level and produce satisfactory results to the extent possible.

Keywords: Non-Programming Education, Artificial Intelligence, Image Processing, Classification, Google Teachable Machine

1. Introduction

In recent years, the construction of platforms for mathematics, data science (hereinafter referred to as “DS”) and artificial intelligence (hereinafter referred to as “AI”) education at universities has been progressing gradually. On the other hand, the reality is that children of the next generation are beginning to learn DS education ahead of university students, as evidenced by the GIGA school concept [1] and other initiatives to make information education compulsory in primary and secondary education [2]. Although there is an urgent need for DS education to penetrate universities to bridge this generation gap, it is not necessarily the case that only students who have mastered programming are

eligible for education.

This paper presents a practical example of a non-programming data analysis class for social science students. Students in the so-called humanities and social sciences are not necessarily good at mathematics or information-related subjects, or they have not studied them yet, so the prerequisite knowledge level cannot be set very high. In the practice, three types of analysis methods were presented to the students, and they developed exercises in which they analyzed and discussed the data to the extent that they were able to do so. The results and effects are discussed in a setting that allows students to recognize their own skill level and produce satisfactory results to the extent possible.

2. AI/DS Education for Social Science Students

The literature [3][4] reports on practical examples of DS education at universities in the humanities and social sciences. According to reference [3], the image sought by companies is not only a data scientist who makes full use of advanced technology, but also a "citizen data scientist" in the business field who can perform analysis as naturally as using spreadsheet software. In our department, we have reorganized the Management Information Science course into the Data Science major from the 2023 academic year, aiming to develop human resources similar to these citizen data scientists.

Reference [4] describes the methods and effects of data science education in a two-layered structure of literacy and application levels at a so-called humanities and social science university that does not have science and engineering departments. Both levels target students who are not very good at mathematics, which is similar to the nature of our department. This approach includes not only students in the social sciences but also those in the humanities, which suggests that the general trend is for students to be less proficient in mathematics.

Our university is planning to expand the program to students in the social sciences, especially business administration students, and mathematics and programming courses at the remedial education level are also provided in the first year of the program. On the other hand, these mathematics and information-related courses are not compulsory, so it is necessary to assume that the background and prerequisite knowledge of mathematics and programming is not uniform when implementing the initiatives described below.

3. Non-programming AI/DS Education Methods

In this chapter, we propose a non-programming exercise instruction method as a new AI/DS education method for social science students. The details are described below.

The first point of view is the analysis method for data analysis. Since the author mainly

developed analysis methods that lean toward AI rather than toward mathematics, the methods can be summarized into the following three categories.

- Classification
- Clustering
- Regression analysis (linear simple regression analysis)

First, classification is simply the process of assigning cases to known categories. For example, it corresponds to the process of forming groups of similar images such as dogs and cats from several types of animal images. Clustering, like classification, is the process of forming groups of similar examples, but the process of forming groups is very different from that of classification. Clustering belongs to unsupervised learning, in which no examples belonging to a known category are given in advance as training examples, and similar examples are gathered together in a situation where no prerequisite knowledge is available, and the resulting groups (clusters) are named later. Finally, regression analysis is a method that can mathematically perform a variety of analyses, but for the sake of simplicity, only correlation analysis based on linear simple regression analysis was used in this study. All of the analysis methods are limited to the extent that they can be taught simply during the class period, and it is assumed that students who wish to learn more about them should take other courses that precede this one.

Next, we focused on analytical tools as the second perspective. Although all of the above analysis methods are mathematically formulated, it is desirable to lower the mathematical hurdle for social science students who do not have comparatively much time to study mathematics. Therefore, they tend to prefer analytical tools in a form that does not show mathematical formulas in a clear manner. In addition, students tend to choose tools with which they are familiar rather than tools with which they have no experience. Based on these criteria, the following analysis tools were prepared and introduced in the class along with a simple usage method.

- Google Teachable Machine
- Spreadsheet tool (Microsoft Excel)
- Python programming language

The first is Google Teachable Machine (GTM) [5]. GTM is a non-programming Web service that uses machine learning. Once the image or audio samples to be classified are prepared, the operation can be completed with a full mouse operation on the browser,

making it an intuitive and very easy-to-use tool even for novice users (Figure 1). Currently, GTM is equipped with a standard function to classify images, sounds, and poses, and is scheduled to be expanded in the future. The training model can also be output as Python or JavaScript program code, although the operation is completed in a non-programming manner.

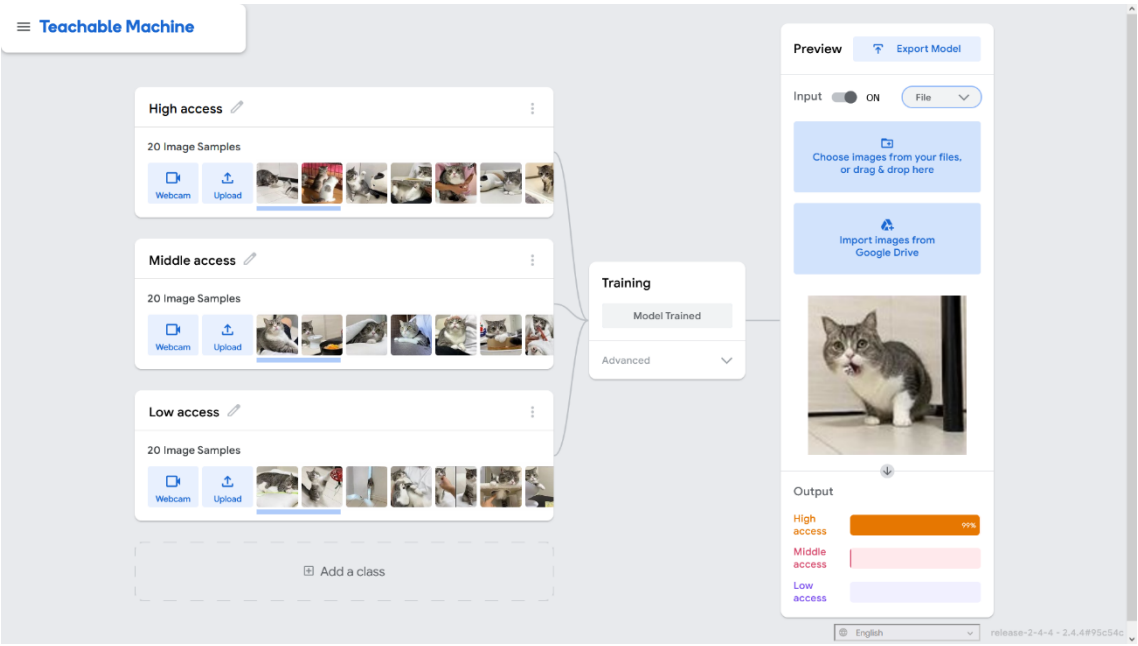
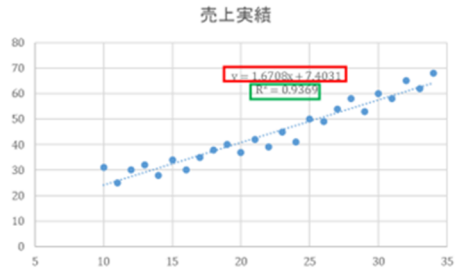


Figure 1: Example of Google Teachable Machine use (cat and dog image classification)

Next, we used a spreadsheet tool since most students have prior experience with Microsoft Excel. Although it is a known tool, we used an analysis tool with additional functions (add-ins) that students have little experience using (Figure 2).



回帰統計

重相関 R	0.967931
重決定 R2	0.93689
補正 R2	0.934146
標準誤差	3.260096
観測数	25

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	1	3628.911	3628.911	341.4408	2.69E-15
残差	23	244.4492	10.62823		
合計	24	3873.36			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	7.403077	2.093347	3.536479	0.001764	3.07266	11.73349	3.07266	11.73349
X 値 1	1.670769	0.090419	18.47812	2.69E-15	1.483724	1.857815	1.483724	1.857815

Figure 2: Example of spreadsheet tool use
(correlation between temperature and beer sales)

Finally, the Python programming language is an option for students who wish to engage in serious data analysis. Since the purpose of this project is non-programming data analysis, there is no particular need for programming, but it was introduced as an advanced technology for students who are interested in this topic and wish to work on it more advanced (Figure 3).

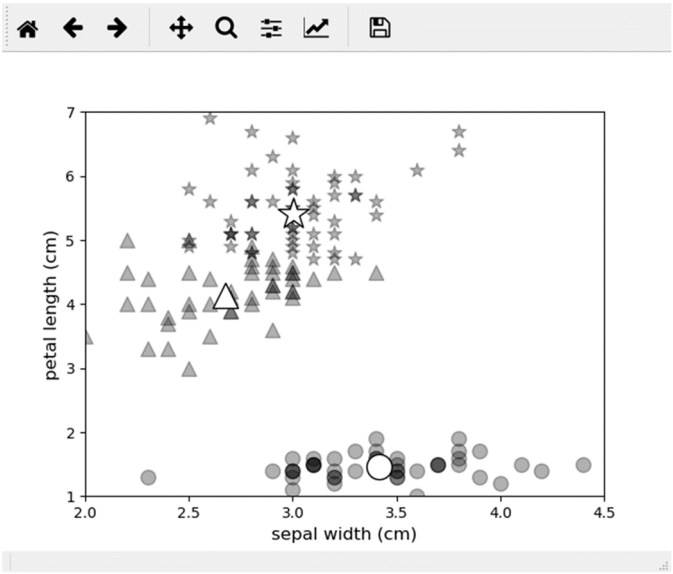


Figure 3: Example of Python use (clustering of three types of objects)

4. Practical Examples and Discussion

In this chapter, we describe examples of actual implementation and their effects based on the policies described in the previous chapters. First, Table 1 summarizes the setting of the subjects and the status of the classes as examples of practical application. The period of practice was from 2021 to 2023, the semester was the second half of the school year, and the duration per week was one 90-minute class. The course was offered at two universities, the University A as our main university and the University B belonging as part-time lecturer, and the number of students enrolled in each university were 14 and 30, respectively.

Table 1: Lecture status for each university

University	Year	Semester	Hour / Week	Class style	Grade	#Students
Univ. A	2021-2023	Fall	90 min.	Practice	2nd	14
Univ. B	2021	Fall	90 min.	Lecture + Practice	2nd	30

Table 2 shows the lesson plan of the subject. The content of the 15 lessons in the semester was largely divided into two parts, with the first half and the second half of the course to be addressed in each of the seven lessons. In addition, a presentation session was set up at the end of each of the first and second halves. The first half is referred to as Exercise 1 and the second half as Exercise 2.

Table 2: Lesson schedule of each university

#Lesson	Contents	#Lesson	Contents
01st	Lecture Guidance	09th	Data Analysis in Python
02nd	AI Application Experience	10th	Clustering
03rd	Classification and Clustering	11th	Regression Analysis
04th	Classification by GTM	12th	Exercise 2 (1)
05th	Exercise 1 (1)	13th	Exercise 2 (2)
06th	Exercise 1 (2)	14th	Presentation on Exercise 2
07th	Presentation on Exercise 1	15th	Review
08th	Fundamentals of Data Analysis	Exam	None

In terms of the exercises, all participants were required to use GTM for exercise 1, and any combination of the analysis methods and tools described in Chapter 3 was allowed for exercise 2. In the final week of each exercise, students were given time to make a presentation. Table 3 shows the number of participants who submitted the exercise 2, in which they could freely choose the analysis method and analysis tools.

Table 3: Status of exercises (Unit: persons)

	GTM	Spreadsheet	Python
Classification	5	-	0
Clustering	-	-	0
Regression	-	38 (1)*	1

Note: 1 person indicated by symbol * is included in the number and is the same as the student who worked on the regression analysis in Python.

Table 3 shows that regression analysis using a spreadsheet tool was the analysis method used by an overwhelming majority of the students (approximately 86%). The reason for this is assumed to be that the students found it easiest to use the analysis tools that they were familiar with before the course. In addition, some students commented that the example of “correlation between temperature and beer sales” presented as an example of

regression analysis in the lecture was easy to understand. A certain number of students (about 11%) were not satisfied with the results of the analysis of Exercise 1 and continued to work on the classification using GTM. On the other hand, a small number of students used programming techniques in their analyses, confirming the existence of a segment of students who are willing to try new approaches.

In both universities, the target year of this course is the second year or above, and Python is set as the preceding programming exercise class. However, since both courses were elective, not all students who took the courses had mastered Python programming techniques, and in fact, an overwhelmingly small number of students had already learned Python in the preceding courses. Since there was not enough time to explain all of the Python syntax in the course, the course was limited to explaining how to use the analysis tools. However, since the purpose of this course is "non-programming data analysis," the fact that most students were able to analyze data using existing tools without prerequisite knowledge of programming languages can be considered to have a certain effect.

Analysis themes included "judging the best color of avocado to eat," "how to distinguish anime characters," and "analysis of individual parts in group songs" in classification. In regression analysis, there was a "correlation between the economy and the number of pets kept," "correlation between the weather and race results in horse racing," "linkage between the ratings of animation TV programs and stock prices," "correlation between the number of police drama TV programs and the number of crimes," and others. None of these were included in the examples given by the instructors but were based on the original ideas of individual students. It can be interpreted that students' spontaneous learning activities were encouraged by providing them with analytical tools for phenomena that they themselves have been concerned about on a daily basis.

A lecture questionnaire was administered at the last class session at each university. The following is a summary of the questions and options, and the results of the questionnaire are shown in Figure 4.

Question: How do you think you have gained knowledge, skills, and abilities by taking this class?

(1) Very much so
(2) Somewhat agree
(3) Cannot say either way
(4) Not really
(5) Not at all

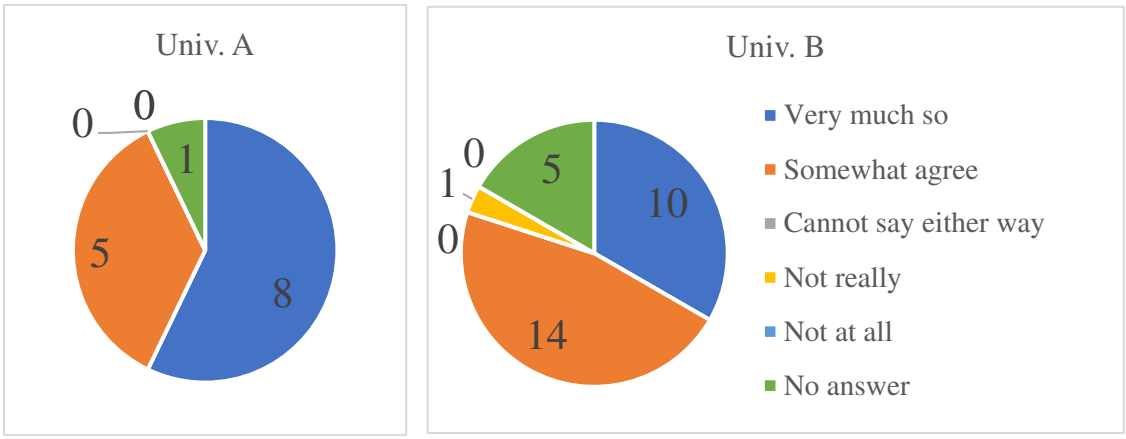


Figure 4: Survey results tabulation (unit: person)

Figure 4 shows that 93% (100% of the valid responses) of the students in University A and 80% (96% of the valid responses) in University B chose options (1) and (2), which were favorable responses to the questions. The grade point averages for University A and University B were 5.00 and 4.32, respectively, which are very high. The results can be interpreted as satisfactory results in terms of mastery of this subject, which does not belong to the so-called social science subjects of the major, although the faculties of each university are the Faculty of Business Administration and the Faculty of Contemporary Social Studies.

5. Conclusions

In this paper, we described a practical example of non-programming AI/DS education for social science students. The response from the students was generally favorable, and they were able to compile their own analysis results and finally produce results that could be presented in the course. From the above results, it can be evaluated that the non-programming AI/DS education program based on the proposed method was effective to some extent.

As for future prospects, education such as that provided in this project may be extended not only to some limited students but also to a wide range of students in other faculties. Currently, only students who have selected the author's courses and students who belong to seminars are eligible for this program, but since the content of this program consists of simple elemental techniques that do not require much prerequisite knowledge, it is relatively easy to expand the program to all students, and there is sufficient room to consider reorganizing the common basic undergraduate courses. In the future, it will be essential to develop the results not only for students who wish to take the course, but also across faculties at the literacy level[6] from a bird's-eye viewpoint.

References

[1] Realization of the GIGA School Concept, Ministry of Education, Culture, Sports, Science and Technology (MEXT) Official Web Site, https://www.mext.go.jp/a_menu/other/index_00001.htm, Last Accessed on 2022.7.22.

[2] Special page on Information Science in High Schools, Ministry of Education, Culture, Sports, Science and Technology (MEXT) Official Web site, https://www.mext.go.jp/a_menu/shotou/zyouhou/detail/1416746.htm, Last Accessed on 2022.7.22.

[3] Tsuji S.: Practical cases of data science education in universities of humanities and social sciences, University Education and Information, No. 172, pp. 13-18, 2020.

[4] Masukawa J., Tsuji S., and Tamura K.: Data Science Education in Universities of Humanities and Social Sciences, Information Processing, Vol. 63, No. 2, pp. d61-d75, 2022.

[5] Google Teachable Machine, Google LLC, https://teachablemachine.withgoogle.com/, Last Accessed July 22, 2022.

[6] Mathematical and Data Science Education Enhancement Center Consortium Model Curriculum (Literacy Level): http://www.mi.u-tokyo.ac.jp/consortium/model_literacy.html, Last Accessed on 2022.7.22.