

Research paper 3

Yan Shi



October 24, 2020

Covid-19 has become the most severe breakout in 2020. Everyone is struggling financially this year. Firms, organizations, and even countries took a big hit from this worldwide pandemic. As the most vulnerable population, college students have suffered a real financial crisis. The impact is enormous; students' financial aid is sliding while student loans' interest is increasing. Income reduction is another concern, along with the limitation of the job market. Reports showed eighty percent of students in the U.K. worried about their financial situation as a result of Covid-19. (TopUniversities, 2020) As a college student, I understand their situation and share the same feelings. Facing an unexpected financial crisis is painful, but it is life. We cannot predict an emergency in life, but we can prepare ready for it.

To minimize the risk of an unpredictable financial crisis, students should take a simple finance course such as personal finance. More specifically, they should learn how to save and invest. However, even most high schools require students to take a course in personal finance, a 2016 survey still indicated that 43 percent of college students don't track their spending while 58 percent said they aren't saving money each month. (Rathmanner, 2016) It is surprising how schools focus on students' financial futures, yet students barely listen.

Americans are addicted to spending is not a secret as consumer spending consistently accounts for about 70% of the U.S. economy. (Louis, 2020) However, asking students to save money directly would induce a reverse effect that they may not understand and resist. Instead, we provide students real investing experience that allows them to realize the importance and benefit of saving. As is said by Julius Caesar, "Experience is the teacher of all things." We understand students are fearful of losing their money in investing in the stock market. This is why we choose credit card cashback bonus as a source for financing. Thus, students don't have to worry about losing real money and still can learn how to manage personal finance. Another reason for selecting a credit card is that students are more willing to spend. As we expect, more spending, more investing. Addressing college students' financial issues while corresponding their favor, we developed a plan that offers students real investing experience while spending on a consumer product. Credit card is the most common financial tool for students, and we can implement an option that allows students to use cash bonuses directly to invest in the stock market or funds. If we can identify students interested in using a cash bonus as an investment option, we can better target them and provide the services. As a result, students will be offered opportunities in real-life investing and consumer product spending habits. Banks could also benefit from attracting more students and expand their markets.

Understanding the problem, we intend to design supervised learning that we will try to divide the respondents into several categories. They have different choices about whether to invest, and if they choose to invest, they choose other channels and have a different risk appetite. If we act as the manager of an investment company or university student financial plan, we can customize the recommendations plans for different students. Although the return on cashback investment is relatively small, we think the manager's goal has been achieved to inspire and motivate young people's investment consciousness. The expected next step is that young people voluntarily invest in some financial products. This would help them maintain good consumption habits and prevent potential economic risks.

Based on our findings of college students, specifically within the context of finance and investing, I decided on one central question that served as the base for the project.

The questions are as follows:

1. Would college students be more likely to invest money if it was money given to them in the form of cashback that could be used as investment credit (instead of their own money)? Would spending money that is not technically coming from their pockets, and therefore carries no risk, encourage millennials to invest? If so, could this potentially lead to a healthy habit of constant, small investments that would eventually serve as a pathway to more significant saving and investment habits?

The data is collected through questionnaire that distributed in school, and each question will be served as an input feature in the machine learning model. The questionnaire receives 100 responses from Clark University. Before data analysis, we have to clean and preprocess the data as the response is raw data and some answers have a different format or missing values. After data cleaning, we can choose different machine learning models and apply the selected model to predict students interested in the credit card investing option. Our model's input is the respondents' features in multiple dimensions, from objective age, majors, monthly income, etc. to subjective consumption bills, investment channels, etc. The output of our model is the categories of respondents.

Of course, the dataset's size also limited model choices; our idea model would be SVM. The model is widely used in classification problems because its simplicity and flexibility; even the dataset is limited, SVM still supports balanced predictive performance. (Derek A. Pisner, 2020) We will also apply data to different models and compare their performance, such as decision tree or KNN, to see which model works best. Meanwhile, we will include feature engineering, enabling us to determine features that are correlated with the output and pick the most critical element. The primary python library will be scikit-learn, which include most of the machine learning model.

The dataset will be split into train and test dataset using train\_test\_split, which allows us to determine train and test dataset sizes. Cross-validation will also be introduced to test our model and avoid overfitting. To provide a better view of our prediction outcome, we defined the confusion matrix for accuracy. The general idea is to count the number of instances of class A classified as class B.

We would visualize the result in both the data cleaning stage and the final output test stage. In the data cleaning stage, the correlation matrix would help us understand the relationship between features and labels. In the test stage, a confusion matrix will be presented. It is a specific table layout that allows visualization of the performance of the model. We can evaluate our model from both precision and recall perspective.

Utilizing established machine learning models, we can identify students interested in investing cash bonus options and provide a reference guide for financial service to better help students who are willing to experience investing while spending. Meanwhile, students will also benefit from offering an alternative option to invest. They could gain dynamic financial knowledge and learn how to manage personal finance from real investing experience.

Throughout the project, we would face a couple of challenges:

1. Dataset limitation is the most crucial problem. The initial responses were only around 100. The data size even shrinks to 70 after data cleaning. Lacking qualified data will affect model's performance and may induce an overfitting problem. Our model selection is also constrained as our dataset is small. Since deep learning requires tons of data, we won't apply it here as the model performance would be impeded with a small dataset.

2. As the survey is not mandatory, a full response to every single question is very rare. We have lots of missing values in the data set; some numeric values could affect model performance. Here, the problems we have come up with logic methodologies to fulfill the missing value and minimize the impacts.

3. Data format is another challenge. We have various data types in our response, such as int, str, and object. Convert data to a proper data type is our priority, and we have to be very cautious. Meanwhile, some of our questions give options to choose from; we have to use the get\_dummy method to convert a categorical variable into indicator variables. We also have to create numeric value for output prediction.

We believe our project will guide banks to target and help students financially. Even though we don't provide direct financial aids to students, those real investing experiences will develop students' financial skills and help them learn how to manage personal finance. We hope students could achieve in academics while managing their wealth.

# References

Aurelien, G. (2017). *Hands-On Machine Learning with Scikit-Learn and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems.*

Derek A. Pisner, D. M. (2020). Support vector machine. *Methods and Applications to Brain Disorders*, 101-121.

Louis, F. R. (2020). Shares of Gross Domestic Product: Personal Consumption Expenditures.

Rathmanner, D. (2016, May 16). *2016 College Students and Personal Finance Study*. Retrieved from lendedu: https://lendedu.com/blog/college-students-and-personal-finance-study

*TopUniversities*. (2020, May 19). Retrieved from 80% of UK Students Struggling Financially Due to COVID-19: https://www.topuniversities.com/student-info/student-finance/80-uk-students-struggling-financially-due-covid-19