IST 718 Project Proposal:
Do Schools Really Consider Students' Success When Allocating Expenditures?

Group "SASI & Co":
Samuel Bull, Aminata Sanogo,
Sabina Seredneva, Ian Ustanik

Part 1

1. Area: The financial situation of public schools plays a crucial role in children's development and academic achievements. This project will look into a schools' financial state and analyze the impact of schools' expenditure allocation on the performance of its students.

2. The Problem Definition: There is a lot of critique about public schools not being transparent and efficient with their spendings. The main debates are going about how schools' financial decisions affect children's education. However, no specific financial model has been created to make sure that public schools in the United States use their allocated funds effectively.

3. Importance of the Problem: The driving force behind schools' expenditures should be student performance. This is not the case in many school districts in the United States. As a result, students are not receiving the best quality education. If school districts learn to allocate their spending more strategically, the quality of student learning will increase. Low student performance can be explained by many other factors including disparities between communities economically. However, there are ways to magnify effective practices. The mismanagementment of funds greatly affects the way kids are taught and the environments they're taught in (i.e. suburbia vs inner city vs rural). It is important to improve the education standards in local school districts because poor quality education can lead to higher crime, higher poverty rates, and more people believing misinformation, in addition to many other societal problems that are not beneficial to the country as a whole.

4. Project's Objective: As a group, we plan on looking at data that pertains to both the finances of school districts around the United States, as well as the testing and overall performance of students within those districts. From there, we hope to be able to find a possible correlation between the allocation of funds within different schools and the success of students overall due to that allocation.

5. Expected results: Hopefully, when all is said and done, we will have a set of classifications as to the school districts who have a positive correlation between the splitting up of their funds with student performances as well as those who fail. From there, we will be able to come up with a model that school districts will be able to follow so that all students get the resources they need to succeed.

6. Statement of Relevance: We are morally obligated to improve the lives of children as they are the future. It is necessary to start the process by implementing changes that will improve their quality of education.

Part 2

1. Goal title: Identifying the correlation between fund management of high performing schools and their students achievements.

2. Tasks:

a) Identify and collect data on school performance (such as results of annual standardized tests).
b) Clean and prepare the dataset for future analysis.
c) Determine successful performing schools based on state and national averages (benchmarks).
d) Analyze main financial indicators of schools' budgets through visualizations (such as histograms and maps).
e) Apply several Machine Learning algorithms (such as logistic regression, k-means clustering, and association rule mining) on the dataset to find patterns that lead to higher performance of students on tests.
f) Provide recommended allocation breakdown of spendings.

3. Expected results: First we expect to identify school districts that successfully allocate their finances. We expect to find patterns in the school districts that allocate their finances successfully. Then we plan to share the effective methods that successful school districts use to underperforming school districts in hopes that their administration considers applying those methods.

4. Expected problems: There are several potential areas of this project where issues may arise. Firstly, there are thousands of school districts over the course of many years that we will be analyzing. This will make the classification of successful performing school districts to be quite complex. To remedy this issue, our group will look to take advantage of associated rule mining and clustering methods to create helpful profiles/benchmarks. There is also a potential for inconsistent data amongst school districts. This type of error is almost a guarantee for how large the dataset is that we will be dealing with missing values will be filled in with the average for that feature. Finally, some of the models that we produce may overfit. To prevent overfitting, we can use tools like cross-validation, regularization, ensemble methods, in addition to proper feature engineering and early stopping of our models if necessary.

Part 3

Datasets: Public Elementary-Secondary Education Finance Data (U.S. Census Bureau)

Education finance data include revenues, expenditures, debt, and assets [cash and security holdings] of elementary and secondary public-school systems. Statistics cover school systems in all states and include the District of Columbia. Data is available in viewable tables and downloadable files. The survey covers all public-school systems that provide elementary or secondary education. The data include revenue by source (local property tax, monies from other school systems, private tuition and transportation payments, school lunch charges, direct state aid, and federal aid passed through the state government), expenditure by function and object (instruction, support service functions, salaries, and capital outlay), indebtedness, and cash and investments.

Tools: Spark ML, Hadoop, Python, Scikit-Learn, R

Models: Linear Regression Analysis, Logistic Regression Analysis, Association Rule Mining, K-means Clustering, Support Vector Machines, Random Forests, K-nearest neighbors

Criteria: MSE, Support, Confidence, Lift, Cross Validation, Classification Accuracy (Test Error), Confusion Matrix

Bibliography:

U.S. Census Bureau (1992-2018). *Public Elementary-Secondary Education Finance Data.* Retrieved from https://www.census.gov/programs-surveys/school-finances/data/tables.html