Kickstarter Projects Prediction

Question/need:

Nowadays, we need a solid foundation to determine success of projects in any field through data analysis. This foundation consists of specific steps to perform the necessary actions and conduct a well-structured and comprehensive data analysis. The main idea of this project is to provide the best possible predictive model for Kickstarter projects that shows whether they will succeed or fail before it is released, depending on different factors.

Data Description:

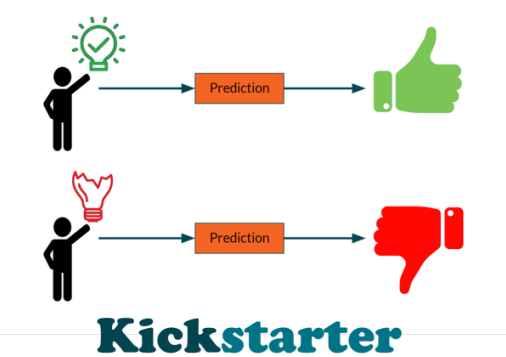
An available dataset was chosen from Kaggle [1] in order to perform this project, which sourced projects data from the Kickstarter Platform. It contains one table with various features, which contain integer values denoting a goal and pledged amount for each project, as well as textual values such as name, category, currency and state. Moreover, the dataset size consists of 324k rows.
However, the target of predicting for the modelling will be the state variable since it will contain two states of either success or failure of the project. Therefore, it will help to know the expected success of projects before they are implemented.

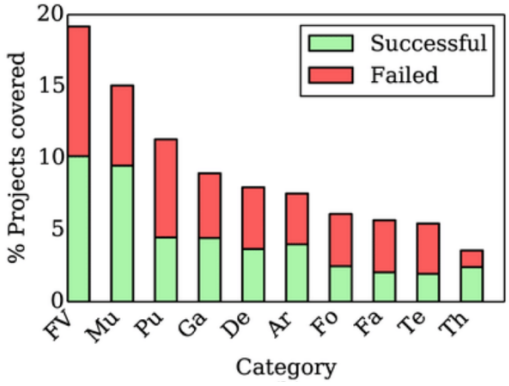
Tools:

Mainly, the python programing language will be used to perform this project. In addition, various libraries will be used as mentioned below. Nonetheless, the project result will be presented using Tableau software for interactive visualizations.

- NumPy and Pandas for data manipulation.
- Scikit-learn for modelling.
- Matplotlib and Seaborn for plotting.

MVP Goal:





References:

[1] Kaggle. Kickstarter Projects Dataset. Retrieved from https://www.kaggle.com/kemical/kickstarter-projects/activity