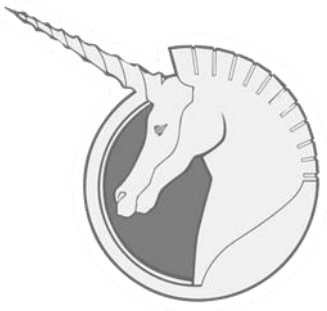


PROJECT UNICORN

NCU machine learning course final project

Group 12

劉裕方,楊宇翔, 顏紹同, 陳禾旻

1. MOTIVATION

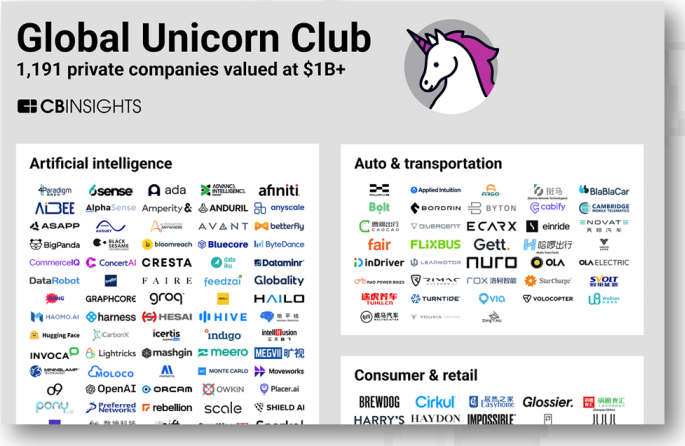
Due to our interest in economics, at first, we chose to build a model in order to predict stock based on the analysis of ppt's discussion. After some research, we found the idea of predicting stock prospects is great. However, gaining the feature of discussion would be difficult for us. Fortunately, we discovered an interesting data set on Kaggle which is about the start-ups. Also, it's related to our original idea about predicting the prospect of a stock/company, so we came up with a new plan based on this data set.

2. GOAL

- Provide an interactive and fun way to know about start-ups.
- Give some possible investment targets.
- Find out the key traits of a company to be successful.
- Passing the class.

3. RELATED WORK/ MARKET SURVEY

"Unicorn", Aileen Lee called a successful company, refers to a privately held startup company with a value of over $1 billion. Our work is to know the potential of this company whether it could become a "Unicorn" company or not, in the other word, the probability of becoming a "Unicorn" company will be our target value. The feature of explanatory value used for fitting the model will be company's funding date, industries, financial, people, technology.



In order to gain these features ,we consider to use the data from CrunchBase, a web database about start-ups. The original source of the data set we use is from here.

crunchbase

4. USERS

By giving the users characteristics of unicorn company, the people who have potential to use Project Unicorn would be Investors that are seeking a practical way to know if a start-up will become a unicorn. People that are seeking the chance to enter a start-up. Start-up founders who want to evaluate their own or other's company by objective index(s).

5. EXPLANATION OF PRODUCT FEATURES

5.1 Backend: NT-D

Our model will be able to tell the probability of a company becoming “Unicorn”. With algorithm capable to Predict the Future Development of Start-up Companies.

5.2 Frontend: La+ Program

Since we won’t want users to see directly into our code, we will have a UI for users to Interact with NT-D.UI will show the user the possibility if a start-up will become a unicorn. Suggest the users what features they should focus on. By default, showing what are in common between unicorns gives the users idea to deploy their resource.

6. THE SOLUTION ARCHITECTURE

6.1 Python 3.10.8

powerful programming language, able to use the package we need.

6.2 PyTorch

PyTorch has tensor computation (like NumPy) with strong GPU acceleration to shorten our training time. DNN we used will be built on a tape-based autograd system. NumPy, SciPy, and Cython are all able to extend PyTorch. PyTorch also has multiprocessing which is useful for data loading and Hogwild training.

6.3 PyTorch Lightning

PyTorch Lightning can make the PyTorch model easier to build, connect, optimize.

6.4 Rust

Rust is memory-efficient; it can power performance-critical services, run on embedded devices, and easily integrate with other languages. It is also memory-safety and thread-safety. Publish to the network would have more reliability by using Rust.

6.5 Tauri

Tauri is a framework for building tiny, blazingly fast binaries for all major desktop platforms. Developers can integrate any front-end framework that compiles to HTML, JS and CSS for building their user interface. The backend of the application is a rust-sourced binary with an API that the front-end can interact with.