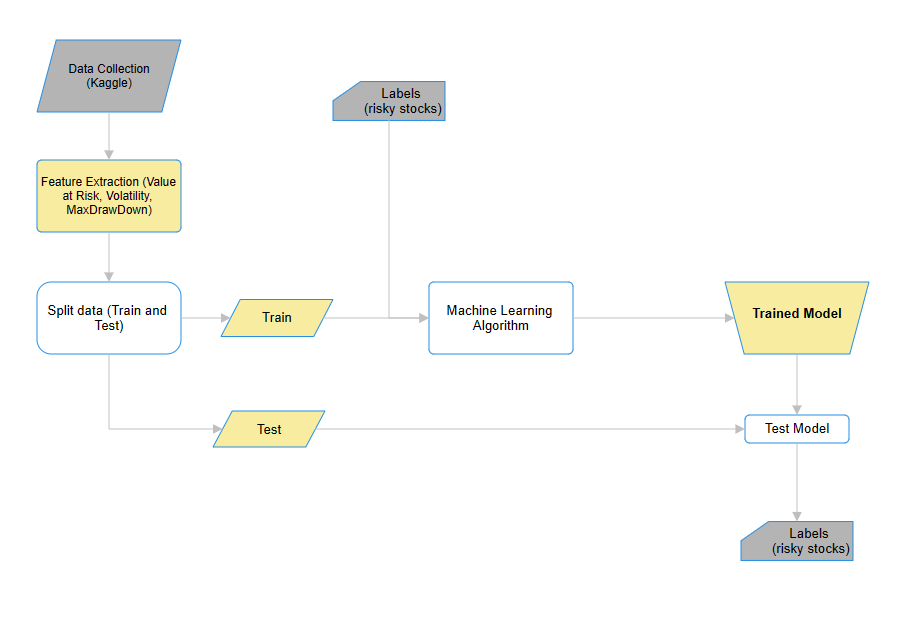
**-Formalize your problem. For example, is it essentially an optimization problem? Is it a searching problem? What is the objective? What will the agent/machine do to achieve the goal?**

The problem is a classification task aimed at identifying risky stocks within the S&P 500 fund. The agent’s objective is to learn to classify stocks based on their feature measurements, distinguishing between risky and safe options. This involves employing machine learning to analyze the data and perform accurate risk classification, aiding investors in making informed decisions about their portfolios.

**-Describe your framework(visualize with figure/chart if possible)**



**-Describe the components in your framework.(For example, Data, Model, Objective, States, Environment…)**

Data: Stock data S&P500 from Kaggle

Features:

* Value at Risk (VaR):
* **Volatility:** Volatility is the extent of price fluctuations in a financial asset over a specific period. It measures the degree of risk or uncertainty associated with the asset's price movements, with higher volatility indicating greater potential for rapid and unpredictable changes.
* Maximum Drawdown:

Objective: To classify stocks as either risky or safe for informed investment decisions.

Train Data: 90% of the preprocessed data

Train Test: 10% of the pre-processed data

Classification labels: 1 for risky stocks and 0 for safer stocks

Machine Learning Algorithm/ Machine Learning Model: Logistic Regression

Model Testing: To test the model, the accuracy of the predicted values from the test data will be calculated. This process involves comparing the predicted values generated by the model with the actual values in the test dataset to determine the proportion of correctly classified instances.

**-Describe the model. (For example, how does it processes the data, the input, output... refer to the textbook and see how the author describe every model.)**

**(Formal language)**

Model: Logistic Regression

-**Describe the implementation detail.(For example, programming tools, platforms, core algorithm/pseudo code, choice of hyper-parameters, and any possible pre-processing steps.)** add pseudo code, hyperparameters

**-Describe how would you like to evaluate the effectiveness of the proposed method.**

**-Discuss the challenge and limitations that you think might lead to unsatisfiable outcomes of your proposed result.**

**-Use some evidence to show your method could work.(For example, executable toy code, snapshots of execution results, or tables/charts/figures of the execution statistics.)**

-Less or equal to 6 pages, single column & 12 font size.

Both text and equations should be well written.

If you introduce mathematical notations, they need to be well defined.