# Report

Project Phase 1 - Data Collection Module (Group Task)

1. System Introduction
   1. How to fetch information

The application that can retrieves stock information is mainly combined with the following two parts, Real-time Data and History Data. Real-time data is the data that the application fetches each time the website update, and as time goes by, the latest data will be stored and become history data. Also, in order to initialize application, after the application started, it will also get history data from website API or database. Real-time data will get after updating. It will run continuously as a background process as it required. We use python and some useful libraries to implement our application. Searching ‘https://finance.yahoo.com’, and get specific html tags to get data. Use BeautifulSoup4 to get HTML element and fetch data after them. The application gets the stock history and latest stock price at the same time. Building an index of name of the stock, and get all information including date, open price, high and low price of the day, close price and adjusted close price.

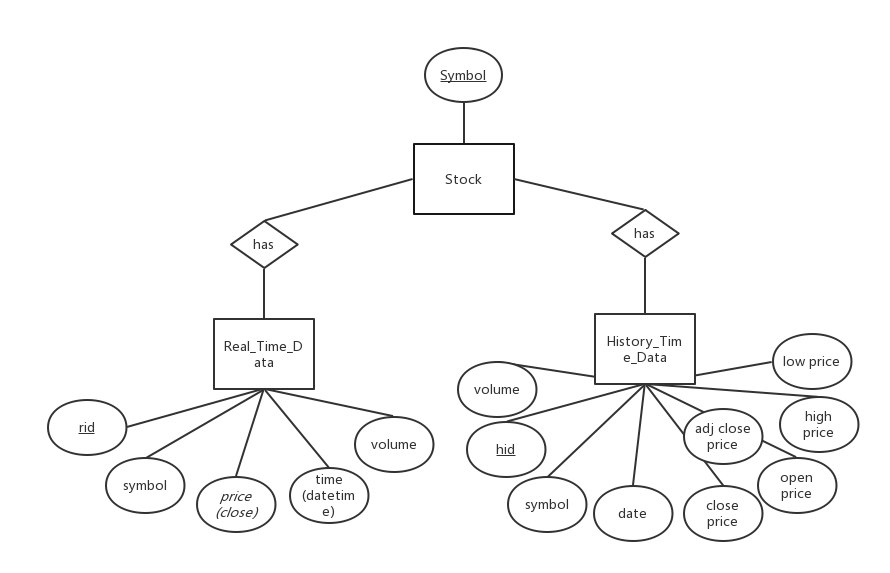


Figure 1: E – R Diagram

* 1. Storage

We will store all information not only in MySQL database but in JSON format. In MySQL, the structure of data listed in the graph below.

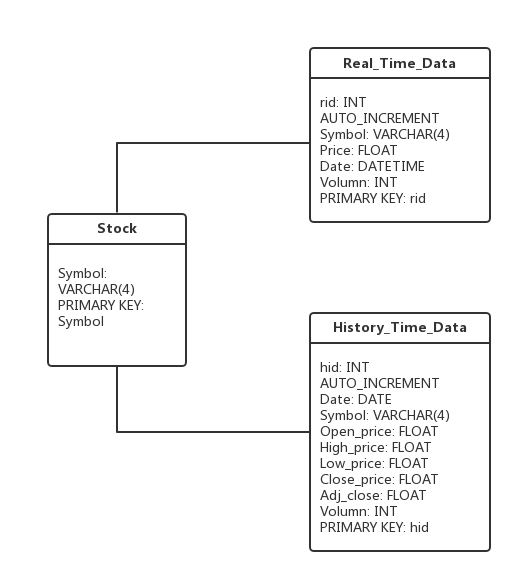


Figure 2: UML Diagram