Presentation

ABSTRACT AND OUTLINE REVIEW
Stock price movement forecasting using machine learning in trading

Sarvsav Sharma (2020SC04239)
Supervisor: Vinaya Sathyanarayana
DSECLZG628T DISSERTATION

BITS Pilani

Pilani Campus

Aim

Stock price movement forecasting using machine learning in trading that will help traders to invest wisely and with increased accuracy.
The dataset collection will be realtime from:

https://pypi.org/project/googlefinance/
https://tradingeconomics.com/nifty:ind

Introduction

The trading is almost happening in every domain, now we can trade in stocks (zerodha), crypto coins(vested.com), commodities (zerodha), players (sixer.com), gamings, etc. And, for the gain in market, a trader should have knowledge of domain as well as external factors of the market. With this model, by working on realtime data and obtaining high accuracy, an investor can choose wisely to invest in the market about the good selling and buying time of the stock.

Problem Statement

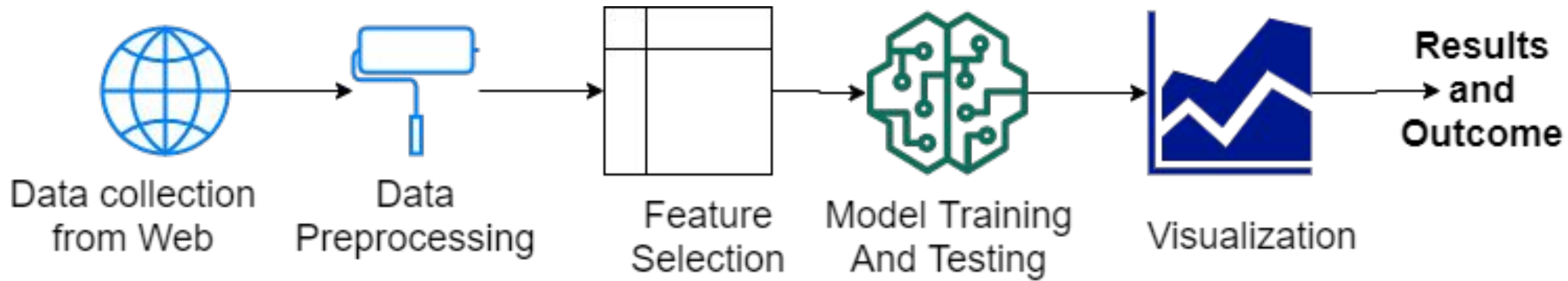Failure to obtain adequate stock market returns.

Existing models are built on older and limited datasets, like 1 month dataset with 1 minute interval or 1 year database with 1 day interval.

How much money is safe to invest?

Detailed explanation of topic

Work process will be uploaded in github continuously, after data collection

Data collection from Web → Data Preprocessing → Feature Selection → Model Training And Testing → Visualization → Results and Outcome

Stock price influencing factors

Problem Statement

Failure to obtain adequate stock market returns.

Existing models are built on older and limited datasets, like 1 month dataset with 1 minute interval or 1 year database with 1 day interval.
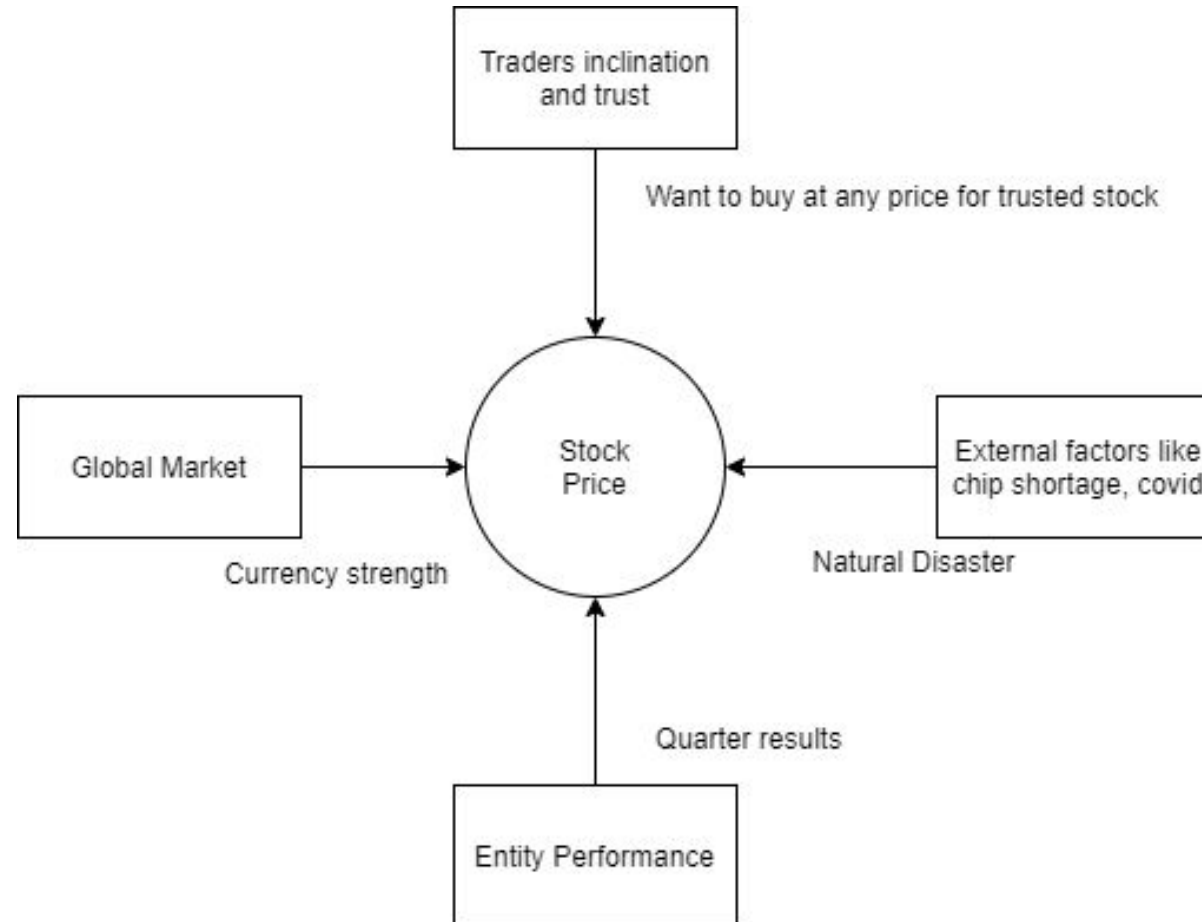
How much money is safe to invest?

Data Collection

Data source:

1. Using API from google or yahoo finance to collect realtime data for better accuracy

What data needs to be collected:

1. Date,Open,High,Low,Close,Adj Close,Volume (Similar to below, image is from
    https://www.marketwatch.com/investing/index/nifty50/download-data?countrycode=in)

DATE	OPEN	HIGH	LOW	CLOSE
12/21/2022	18,435.15	18,473.35	18,162.75	18,199.10
12/20/2022	18,340.30	18,404.90	18,202.65	18,385.30
12/19/2022	18,288.10	18,431.65	18,244.55	18,420.45
12/16/2022	18,319.10	18,440.95	18,255.15	18,269.00
12/15/2022	18,614.40	18,652.90	18,387.70	18,414.90
12/14/2022	18,671.25	18,696.10	18,632.90	18,660.30
12/13/2022	18,524.40	18,617.25	18,490.20	18,608.00
12/12/2022	18,402.15	18,521.55	18,345.70	18,497.15
12/09/2022	18,662.40	18,664.70	18,410.10	18,496.60
12/08/2022	18,570.85	18,625.00	18,536.95	18,609.35
12/07/2022	18,638.85	18,668.30	18,528.40	18,560.50
12/06/2022	18,600.65	18,654.90	18,577.90	18,642.75

DAILY   WEEKLY   MONTHLY

⬇ DOWNLOAD DATA (.CSV)
Limited to a maximum of one year.

Data Analysis

1. Data Cleaning
   a. Removal of duplicate and null values
   b. Dropping unnecessary columns
   c. Data type conversion of fields
   d. Creation of addition features
   e. Reduce data dimensions

2. Data Analysis
   a. Discovering data trend like open, close, high, low
   b. time wise price variations
   c. Visualizing data
   d. Data corelation
   e. Bar chart comparison

ML Models

1. SVM POLY, LINEAR based model

Business Process Flow

Monitor opening and close index of a stock

Monitor lowest stock value

Monitor highest stock value

Buy or Sell the stock

Client: sixer.com

Uniqueness of the project: Model with higher accuracy to predict the price of stock every minute

Benefit to the organization: With this functionality, the traders will be able to make profitable and defensive investments in intraday trading.

Scope of work:

Using a Python script, data is scraped as well as processed, and fitted to machine learning (ML) models using different algorithms to predict the price of the stock, and deploy it as a web application.

Resources needed for the project, including people, hardware, software, etc.

Tools: Jupyter Notebook, Anaconda, Streamlit, flask

Potential Risks

Potential challenges & risks in doing the project :

    1. Limitation on web scraping of data for number of records

    2. Availability of latest dataset

    3. Choosing the right features for best accuracy

    4. Choosing the best visualization

    5. ML model's implementation via Streamlit

Plan of work

#	Task	Expected date of completion	Names of Deliverables
1	Web scraping of data from Nifty	20 Dec, 22	Data collection
2	Pre processing of data	10 Jan, 23	Valid data
3	Visualising and insights of data by plotting	02 Feb, 23	Data visualization
4	Obtain best model using ML algorithms	20 Feb, 23	ML model
5	Deployment using streamlit and flask	21 Feb, 23	Web Application

Thank You