

Stock Price Prediction With Relative Strength Index(RSI)

Objective :

Development of a predictive model for stock price prediction so used rsi crossover strastergy

To predict direction of market

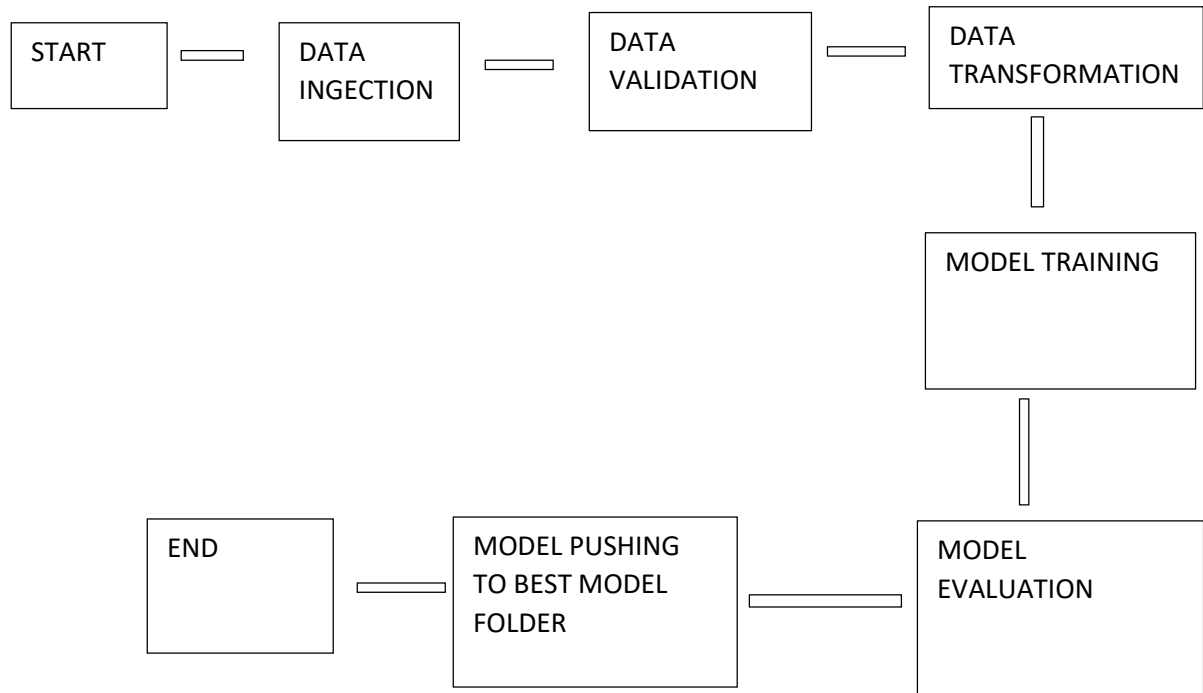
Benefits:

Avoid wrong entry in market

Help to do systematic trading

Avoid overtrading and emotion

Architecture



Data Description

Data contains open close, high ,low ,datetime values and data taken from stock broker huge amount of data needed for this problem then only it will work because we need to transform data then we need to filter our condition matching data and sort them.

Data ingestion

Data injected from Casandra database before that we need to connect with Casandra database and fetch data from table and insert into csv and store inside artefact folder

Data validation

Here we are validating our data set is having numerical columns and number of columns present in dataset and validated data store in validation artefact

Data transformation

Transform validated data first we will pass through robust scaler then impute missing values then we having only open,close,low,high,datetime values from this we are generate RSI(14), RSI(9), EMA(20) ,EMA(5) values and we will drop datetime column and autoscaling and transformed data we will split traintestdata and store numpy numpy array and pre-processor object also store inside datatransformation

Model training

We are using xgboost classifier for prediction and training transformed data on xgboost classifier and trained model saving artifact directory

Model evaluation

Here evaluating trained model and if this model have good accuracy and compare with previous model accuracy if this model have more accuracy than previous model then insert this model to best model folder

model pushing

After model evaluation we compare accuracy of current model with previous model and then if trained model has high accuracy than previous model then we push current model to best model folder and remove old model from that folder.

Q/A:

Q1) what is the source of data?

I will fetch data from fyersr api then I will dump into Cassandra database

Q2) what was the type of data ?

The data was the combination of numerical and datetime

Q3) what format have you used for datetime?

The format was " %d-%m-%Y % H: %M "

Q4) How logs are managed

In log folder log file will created whenever we started execution of Pipeline then log file will create and record all logs

Q5) what technique used for data pre-processing?

Date time column dropped and out layers removed and missing value imputed
And data transformed to rsi14,rsi15,ema5,ema20 values

Q6) How training was done or what models were used?

For training model we used xgboost regressor and transformed nnpay
Will take for training after training evaluate model and if accuracy is
Then push current model to saved model folder

Q 8) How Prediction was done?

For prediction user input will take and convert as pandas dataframe
and then convert numpy array then call best model from saved_model
directory and numpy array passed for prediction and output displayed in
frontend

