MPE Annotations

Andrés Camilo Viloria García
aviloriagarcia00@mylangara.ca
Data Analytics
Langara College
Vancouver, BC

Diana Sayuri Ortiz Montes
dortizmontes00@mylangara.ca
Data Analytics
Langara College
Vancouver, BC

This document states how 2 phases of project were used two different ways to compute MPE metric.
The method used to calculate the MPE in Phase 3 and 4 differs from the approach described in the paper titled "Short-Term Stock Price Forecasting using Exogenous Variables and Machine Learning Algorithms." The variation in calculation methods is illustrated below:

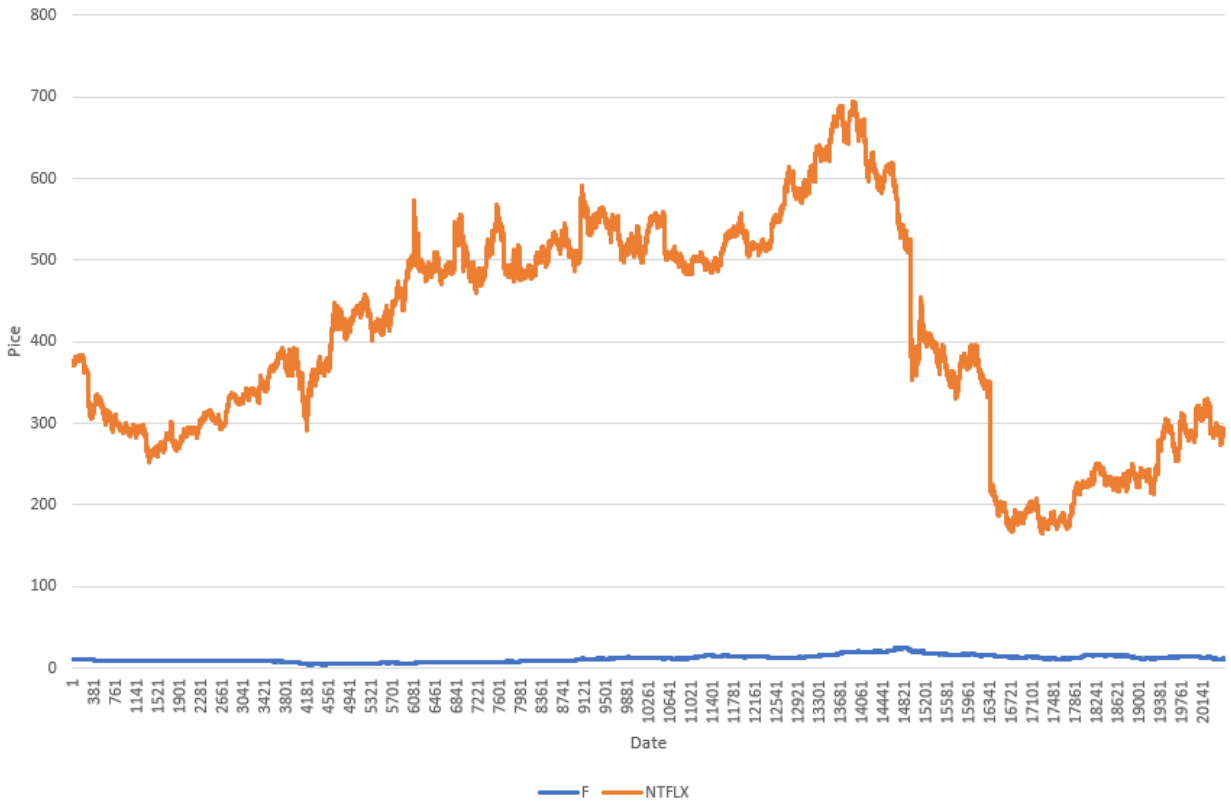
Article ($) - Dolar unit	**Code (%) - percentage unit**
$$MPE = \frac{1}{n}\sum_{i=1}^{n} max(\hat{Y}_i - Y_i, 0)$$	$$MPE = \frac{1}{n}\sum_{i=1}^{n} max(\hat{Y}_i - Y_i, 0)/Y_i$$
This was the method used to compute the MPE in whole phase 3 and 4.	

`def mean_positive_error(y_true, y_pred):`
 	
`  y_true, y_pred = np.array(y_true), np.array(y_pred)`
 	
`  error = np.mean(np.maximum((y_pred - y_true),0))`
 	
`return error`	

In the article, it computed ratios or percentages to compare the positive prediction errors with the actual values. Then the average is calculated of these ratios or percentages. In the current code, simply subtract the predicted value from the actual value in dollars per share and then find the average of these differences.

The decision to decide which one is more appropriate depends on the purpose of the metric. The percentage version standardizes the measurement, making it more meaningful on comparing different stock's price level. For example, consider two stocks: Ford (F) with a current price of $20 and NETFLIX (NFFLX) with a current price of $450. If the AI model overestimates Ford by $5 and Netflix by $100, a simple numerical difference metric would suggest that the model's overestimation is higher for Netflix. However, in terms of percentage difference, the model's overestimation for Ford would be 25% (($5/$20) * 100) while for Netflix it would be 22.2% (($200/$450) * 100). This allows for a fairer comparison between the two stocks.

Please for your reference of the previous example check the following graph with the ranges of prices of all stocks evaluated from July to December 2022.



In order to compare the differences in results between two calculations in a real situation, Revise the file MPE.xlsx that demonstrates the variance in calculations for a specific stock and its corresponding results. With access to the file, I will be able to visualize the calculations and provide a meaningful comparison of the outcomes.

It was decided to continue computing the MPE as previously done in the code (dollar unit).