**We are Team WOLFS**

**Team Members include: Krishanu Das**

**Rohit Gedam**

So, basically our line of action was to modify the features to extract more info about the dataset. After some trial and error we were successful in doing so.

And we were also able to find the significance of the output variable y.

* This is the link to our modified dataset:

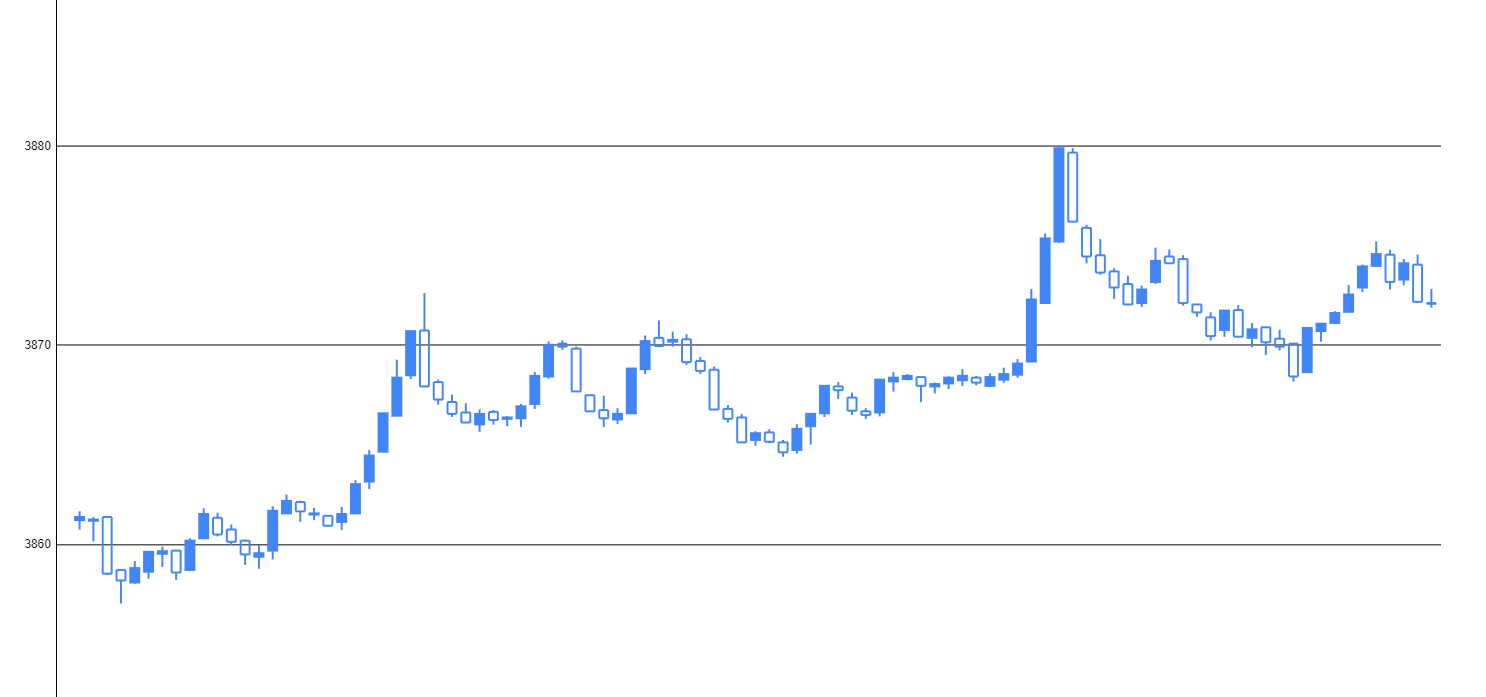
[Copy of test\_data\_Analytics GC.xlsx - Google Sheets](https://docs.google.com/spreadsheets/d/10A2IBA_9rtbuOhhBXC_iKlZBGroZ0yO6/edit#gid=388645770)

* The significance of the target variable was, when its

0 : it shows bearish trend

1 : it shows bullish trend

[Copy of Extract.xlsx - Google Sheets](https://docs.google.com/spreadsheets/d/1x_erQ_1zsdox_vxBhPr0ctY5_h4SvzRk/edit#gid=1270391576)



We research on various models and applied them in the code but what we found that Logistic Regression was best fitting among Random Forest, Logistic Regression & Neural network

We guess thar maybe the other 2 method were overfitting the model because of their complexness.

References:

* [machine learning - Techniques for increase random forest classifier accuracy - Data Science Stack Exchange](https://datascience.stackexchange.com/questions/67493/techniques-for-increase-random-forest-classifier-accuracy)
* [Session 3 - neural networks.ipynb - Colaboratory (google.com)](https://colab.research.google.com/github/yannsia/tutorials/blob/master/session3/neural_networks.ipynb#scrollTo=fQ4m9VfZ9VSD)
* [funcode/BreastCancer.ipynb at master · lmoroney/funcode · GitHub](https://github.com/lmoroney/funcode/blob/master/BreastCancer.ipynb)
* [Building Neural Networks With PyTorch in Google Colab - DZone AI](https://dzone.com/articles/building-neural-networks-with-pytorch-in-google-co)

Our code:

[Analytics GC.ipynb - Colaboratory (google.com)](https://colab.research.google.com/drive/1M7tDh4f6nD41T0shGbfgsHrYDoaFsRPZ#scrollTo=BiwMGsrU9-Hi)