**Using stock exchange data, train a model that predicts the stock price level in the future. You may use R language but don’t put in source code in the answers. Your source code will be submitted separately**

* 1. Explore and visualize the dataset. Plot price level over time in training and testing datasets. Explain what steps you have taken in preprocessing and why.
  2. Develop two different models using Long short-term memory (LSTM), and ARIMA algorithms. Explain your methodology in determining model architectures in detail. For instance, answer which parameters did you use in the models, how did you determine those values and how did you optimize them.
  3. Visualize training and testing results by plotting actual vs predicted values as a function over time in test and train datasets for each model. Plot different models in different styles (dashes, lines, dots, etc.) or in different colors so that they can be distinguished.
  4. Using a table, compare average performances of the models using evaluation metrics of root mean square error (RMSE) and mean absolute error (MAE) for both testing and training datasets.
  5. Discuss the model performances and your key findings. Interpret the meaning of results in your opinion

1. **Using the transactions dataset, answer the following questions in the context of Market basket analysis. The dataset shows the transaction IDs and the list of products sold within the transactions, separated by a space. You may use R language but don’t put in source code in the answers. Your source code will be submitted separately.**
   1. Apply Apriori algorithm to the dataset and find all the frequent itemsets with the specified minimal support levels of (10%, 20%, 30%, 50%). Evaluate the results and discuss your findings.
   2. Apply Eclat algorithm to the dataset and find all the frequent itemsets with the specified minimal support levels of (15%, 25%, 35%, 45%). Evaluate the results and discuss your findings.
   3. Discuss the performance bottlenecks of Aprior/Eclat algorithms and explain how can they be improved? Describe potential issues of the algorithms which may occur on running large transaction datasets with millions of transactions and each transaction can contain tens or hundreds of items?
   4. Association rules with high confidence are usually considered more interesting. That said, association rules with 100% confidence level may not be considered as interesting. Why do you think that is the case? What might be the reason? Discuss why association rules with 97% confidence level can be regarded more important?
   5. [4 points]. Explain the concept of lift and how could you use the lift value in recommendation of books in e-commerce sites?
   6. [3 points]. Also, research the reasons why Netflix created the House of Cards TV series starring Kevin Spacey and Robin Wright. How did the concept of lift factor into Netflix’s decision to make the show?