Why Selected the specified algorithm?

For predicting the type of shipment we use an Artificial Neural Network using TensorFlow/Keras' deep neural networks. This algorithm has been chosen as there are a large number of input parameters in the dataset which can affect the shipment mode.

For predicting the vendor the shipment came from we have used Decision Tree using sklearn's decision regression algorithm. This algorithm has been chosen as only a limited set of of input fields contributed in predicting the vendor

About algorithms used in model building:

We will be using TensorFlow/Keras neural network to make our prediction. An Artificial Neural Network(ANN) is the component of a computing system designed in a way similar to how the human brain analyzes and makes a decision. Artificial neural networks work like a human brain. The human brain has billions of neurons and each neuron is made up of a cell body that is responsible for computing information by carrying forward information towards hidden neurons and provide final Output. ANN initially in the training phase learns to identify patterns based on inputs given to the input layer.

For vendor prediction we first used linear regression, Linear regression analysis is used to predict the value of a variable based on the value of another variable. For better accuracy we used DecisionTreeRegressor from skLearn. Decision Tree is a decision-making tool that uses a flowchart-like tree structure or is a model of decisions and all of their possible results, including outcomes, input costs, and utility.

Decision-tree algorithm falls under the category of supervised learning algorithms. It works for both continuous as well as categorical output variables.

Process followed for model building:

Shipment Type Predictor:

The dataset contains a lot of non numeric data points which the model cannot use. We first perform integer encoding on the non numeric columns and then one - hot encoding can be applied to it. This is where the integer encoded variable is removed and a new binary variable is added for each unique integer value. The column which is to be classified has just been integer encoded. All the values were scaled down to -1 and 1 range for the neural networks. The training and test data have a 70 : 30 split. Then the model is trained

using an Artificial Neural Network with two hidden layers. The model was then tested and the accuracy was recorded.

Vendor Predictor:

The dataset contains a lot of non numeric data points which the model cannot use. We first perform integer encoding on the non numeric columns. Only the relevant columns were used for training and testing data. The training and test data have a 67 : 33 split. Then the model is trained and accuracy is recorded based on the output on the testing data.

Technologies used for model building:

- Tensorflow/Keras - Keras is used for creating deep learning models
- Scikit-learn - LinearRegression, DecisionTreeRegressor

Accuracy of algorithms with analysis:

Shipment type predictor: 88.13%

There are 4 types of shipment available in the dataset. By air, ocean, air charter, truck. Depending upon details of the order we can predict which shipment to be used

Vendor prediction: 86.42%

For analyzing vendors we analyzed various terms such as what type of product the vendor sells, Country of shipment, mode of shipment, product group and its pack price. Depending upon most suitable input we can predict that this vendor can be a good choice or not.

Hyperparameters used in models:

Shipment Type Predictor:

- Train-test split ratio: 70 : 30
- Learning rate in optimization algorithms: gradient descent
- Choice of optimization algorithm: Adam Optimizer
- Choice of activation function in a neural network layer: softmax
- Number of iterations (epochs): 100

Vendor Predictor:

- Train-test split ratio: 67 : 33

Links :

https://colab.research.google.com/drive/1ZmS5FaWQiqICrCKytuV60JBCuF473iro?usp=sharing

https://colab.research.google.com/drive/16_85RJqhA6aIbnCuJ_ctwEjJI-Z6VLCw?usp=sharing

https://colab.research.google.com/drive/1c4JDZ94V6od1EpVd0nZdH-jSFGlKQnLP?usp=sharing

Repository:

https://github.com/SpecularAura/8_ZeroTwo_3