**Lesson 03 Discussion**

Ensemble models can provide significant reduction in variance of predictions. These algorithms can also be used as powerful data mining algorithms. Describe a situation where this data mining capability can provide valuable information on the relationship between variables in a data set. How will the limitations of these method limit their application in your scenario?

One of the biggest limitations of the Ensemble Model is the interpretability which depends on the complexity of the models used. One example of this could be a random forest model, where each individual decision tree provide a clear path to the target variable (predicted variable) and by itself a decision tree could be a great tool to also obtain information on which features are the most relevant to the mode, but when multiple decisions trees are combined to bust the performance on a prediction model it becomes really difficult to use it to identify directly clear meaningful relationships between the variables in comparison with the target feature. But even in this situation ensemble models can help to identify important/relevant features by providing an overall summary of importance of each prediction using RSS (regression) or Gini index (classification). The limitation of this is the ability to clearly pin point the reason each of the features were given more importance than others.