**MODULE 21 Challenge**

1. **Overview of the analysis: Explain the purpose of this analysis.**

The purpose of the analysis is to analyze the performance of nonprofit foundation Alphabet Soup in selecting the applicants for funding based on the predictions done by using machine learning i.e., Artificial Neural Network(ANN). The dataset contains more than 34,000 organizations in various industry and affiliation.

1. **Results**: Using bulleted lists and images to support your answers, address the following questions:

* Data Preprocessing
  + **What variable(s) are the target(s) for your model?**

The variable used in the ANN model are:

* APPLICATION\_TYPE
* AFFILIATION
* CLASSIFICATION
* USE\_CASE
* ORGANIZATION
* STATUS
* INCOME\_AMT
* SPECIAL\_CONSIDERATIONS
* ASK\_AMT
* IS\_SUCCESSFUL
  + **What variable(s) are the features for your model?**

The variable ‘IS\_SUCESSFUL’ is the feature of the model which was used in preprocessing the data as training and testing datasets.

* + **What variable(s) should be removed from the input data because they are neither targets nor features?**
    - EIN and NAME
* **Compiling, Training, and Evaluating the Model**
  + **How many neurons, layers, and activation functions did you select for your neural network model, and why?**

I did iterate various neurons and layers combination. Adding high number of neurons doesn’t improve the model performance and the output performance was almost same. So, I decided to keep the model with 3 layers, first layer of 16 neurons, second layer with 8 and third layer with 4 neurons, respectively.

* + **Were you able to achieve the target model performance?**

Yes, the model performance (after optimization) had the accuracy of 75% and Loss was 49%.

* + **What steps did you take in your attempts to increase model performance?**

I tried the combination of various random state (training dataset), layers, neurons and changing the epochs no with each iteration.

1. **Summary**: **Summarize the overall results of the deep learning model. Include a recommendation for how a different model could solve this classification problem, and then explain your recommendation.**

ANN is a strong tool to analyze any non-parametric or non-linear data. The great benefit is that it can handle long datasets and work well with missing values in dataset. If the model achieves the optimum accuracy (minimum of 75%), it can give predictions that are very close to the actual values.

One of the limitations with ANN is that a person needs to be very good in handling ANN when making any predictions and forecast with neural network.

Neural network is useful in handling complex and non-linear dataset which means data with high variance, skewed or large number of outliers. One way to handle the non-linear data is smoothening or transformation by applying filters, autoregression etc. (to reduce the variance in data and to attain normal distribution). Later, using non-supervised learning (like regression analysis), predictions/forecasts can be made.