

Overview of the analysis: The purpose of the analysis is to find out the best application when considering funding. By using ML we are able to find trends that can support nonprofits that have the best chance within their venture.

With knowledge in machine learning and neural networks we are then able to create a binary classifier that can predict whether applicants will be successful if funded by Alphabet Soup.

- **Results:**
- Data Preprocessing
 - I decided to remove AFFILIATIONS as I personally believe it to be a weak contribution.

```
[2] # Drop the non-beneficial ID columns, 'EIN' and 'NAME'.
     application_df = application_df.drop(columns = ['EIN', 'NAME', 'AFFILIATION'])
```

- What variable: The variables that I wanted to use would be INCOME_AMT and USE_CASE as I believe that both of these variables are important to consider. I am also keeping the code that has to do with both CLASSIFICATION and APPLICATION_TYPE.
 - When thinking about INCOME_AMT you want to see if the nonprofit is going to be heavily dependent on your money as well as seeing if they can provided services for themselves rather than constantly relying on more and more money

```
[12] income = application_df['INCOME_AMT'].value_counts()
      income
```

```
[13] # Choose a cutoff value and create a list of application types to be replaced
      income_to_replace = list(income[income<500].index)
      income_to_replace

      # Replace in dataframe
      for app in income_to_replace:
          application_df['INCOME_AMT'] = application_df['INCOME_AMT'].replace(app, "Millions")

      # Check to make sure binning was successful
      application_df['INCOME_AMT'].value_counts()
```

- When considering USE_CASE its important to think about what the money will be used for as it can be a big or small project that can possibly rely on more money or if the project solves an immediate problem left to be solved
- I found this to be interesting as the majority of projects are in need of Preservation funding followed by product dev. So it seems that there are those that are struggling to survive and others that are in need of improving their products and or services.

```
USE = application_df['USE_CASE'].value_counts()

case_use = USE[USE>1]
case_use
```

Preservation	28095
ProductDev	5671
CommunityServ	384
Heathcare	146
Other	3

Name: USE_CASE, dtype: int64

- What variable(s) should be removed from the input data because they are neither targets nor features?
 - I didn't find any reason to remove anything as OTHER being equal to 3 is way too small to impact anything substantially
- Compiling, Training, and Evaluating the Model
 - How many neurons, layers, and activation functions did you select for your neural network model, and why?
 - I added more nodes in both the first and second layers and found that i ended up making my model worse

- I have no idea why it got worse but i kept adding and subtracting and kept going back to the outcome that my model kept getting worse.
- Were you able to achieve the target model performance?
 - I was not, the model dropped from a 72 to a 63 indicating that need more experience with ML in order to create a better accuracy that is above .75

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.

Although I was unable to make the optimization better I was however able to see data that I found to be interesting to warrant a different type of analysis rather than ML. An example being that when I looked further into the amount of income that some of the nonprofits generated I was surprised to see that there were a substantial amount that had millions. I grouped those that had a range from 5M- 50+M. There were 564 instances of those and that could honestly lead a group to have success as they aren't heavily dependent on the money needed. Not to mention that there was a majority of nonprofits (24,383) that had little to no money.

Another variable that surprised me would be the used case for the money. The top two were Preservation (28,095) and ProductDev (5,671) and if we were to analyze the data without ML I would want to see what type of nonprofits were rich in money and to see if the project was a success. There are so many things to explore and felt limited by ML.