**Machine Learning Workshop Project - Aravind, Katy, Kevin, Antony**

Mobile Bay, AL – Data provided by Aravind

**Project goal:** Determine the relationship between independent and dependent variables then compare machine learning and deep learning models to see which method is more accurate.

The first three datasets are independent variables - Atm temp, river discharge, wind, solar radiation

The other three are dependent variables - Temperature, Salinity, and DO

Train deep learning and machine learning model to build up the relationship b/w independent and dependent variables. Then apply the model to find the future of dependent variables using independent variables.

**Tidal elevation + Atm temp + River discharge + Wind + Solar Radiation = Temp + Salinity + DO**

1. Data consolidation and cleaning - timestep uniform, one main file for all data
   1. 12 years total - .csv format, some in netcdf that needs to be converted
      1. First start with 1 year (2020 - Cyclone Sally)
      2. First station (Bon Secour)
      3. 10 months for training and 2 months for testing
      4. 24 hour/daily average for variable
   2. Increase to more years and multiple stations - time & computing dependent

Make all columns numeric, no empty columns, combine all variables into one file

Do not focus on the statistical side of things

1. Make a repository, with two branches for master and edits - share access
   1. Original data files in repos, then separate folder for edited/cleaned versions
2. Develop model
   1. Focus on relationship between two variables to start with
      1. We are going to ask Dr.’s what they think - multiple independent variables and one dependent, which variables would be the best to start
      2. Don’t focus on stats part, just machine learning parts
   2. Deep learning model - Antony
   3. Machine learning model - Kevin
3. Presentation - combined effort - 10 - 15 mins
   1. Introduction to each person + research focus
   2. Data background - where, how collected, missing gaps, etc.
   3. Create model comparison for each variable

Corre