## University of Toronto – Data Analytics & Visualization Bootcamp

# Project 4 – Machine Learning

**Correspondence #2 – June 5TH, 2023**

1. **Presented research on strengthening the topic**

* **Comment: Dataset is related to data science, therefore more beneficial to have turnover rates related to same industry as specific**
* **Obtain good key point from LinkedIn and Indeed (Madhur’s comment)**

1. **Recall not expected to get better**

* **Proper interpretation of resulting data will be required for:**

1. **Final narrative**
2. **Iterative process explanation and conclusion**
3. **Required for tomorrow:**

* **Accuracy score requested in rubric is >75%**

**This is doable as long as we can showcase the models**

* **Neuro-network may have been overfit which could be corrected by imputation and binning data**

1. **Create matrix to showcase the results across multiple models**
2. **All data has been loaded into SQL – complete**

**Items of note:**

* **Research data acquired – charts, important points, valuable story telling narrative**
* **Recall was identified as MOST important, more important than precision in additional reading**
* **Preprocessing**
* **Additional markdowns being generated**
* **Incorporation of Ronald’s code**
* **By Wednesday:**
* **Expectation to clean-up notebook and identify final set of data**
* **Identification of what recall value is required and how important it is going to be**
* **What feature is important**