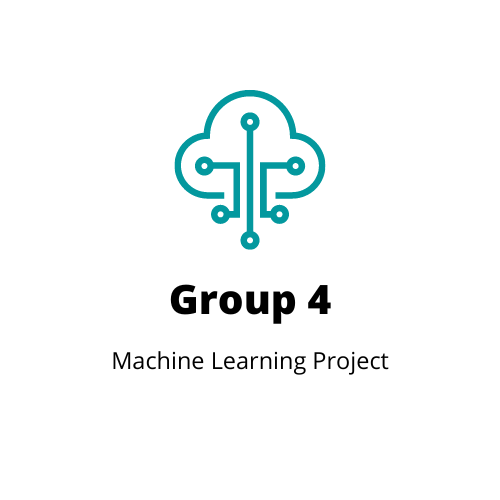
**Project Name: Risk Prediction Models For Diabetes Using Machine Learning**

**Project Logo:**

****

**Team Members:**

|  |  |  |
| --- | --- | --- |
| **Name** | **Email** | **Phone Number** |
| Isabella Taylor | isabellajadet@gmail.com | (928)-863-8788 |
| Leah Nash | leah\_nash@icloud.com | (918)-928-5443 |
| Sandra Braun | seweiland@gmail.com | (405)514-6170 |
| Valerie Grannemann-Barber | slipknotsangel2@hotmail.com | (918)978-0953 |

**Motivation & Summary Slide**

Leveraging machine learning technology holds promise to reduce health care costs through the ability to predict disease state and potentially identify at-risk individuals earlier in the disease process to target education and intervention activities.

**DATA sets to be Used:**

* <https://www.kaggle.com/datasets/alexteboul/diabetes-health-indicators-dataset>
  + Derived from: https://www.kaggle.com/datasets/cdc/behavioral-risk-factor-surveillance-system
* https://www.cdc.gov/brfss/annual\_data/2015/pdf/codebook15\_llcp.pdf

**Tools/Modules to be Used:**

* Python Pandas
* Tableau
* PySpark
* Tableau

**Breakdown of Tasks** (Ownership by Group Member)**:**

* All: Will work on model research for implementation and optimization. Attempt initialization over weekend and use slack channel to communicate which model we’re currently investigating to avoid duplicative effort
* Isabella: Tableau Dashboard
* Leah: GitHub Documentation and Repository Upkeep/Merge, etc.
* Sandra: Data Model Optimization & Presentation
* Valerie: Data Model Optimization & Presentation

**Schedule for Completion of Tasks:**

|  |  |  |
| --- | --- | --- |
| **Date** | **Task** | **Notes** |
| 04/06/23 | Project Work/ Project proposal |  |
| 04/06/23 | Collecting the data |  |
| 04/10/23 | Work on Data |  |
| 04/10/23 | Code development |  |
| 04/11/23 | Code development |  |
| 04/13/23 | Create Visualizations |  |
| 04/16/23 | Mock Presentation |  |
| 04/17/23 | PROJECT PRESENTATION |  |