Group 5: Miracle Awonuga, Stiven Lavrenov, Abby Roberts, Charles Tippett

Module 5 Week A Summary

**Key Points Discussed:**

* The group debated between different research papers and datasets, ultimately deciding on a dataset from **UC Irvine** related to **heart disease prediction**.
* The dataset is well-documented and aligns with the project requirements, including the need for machine learning models.
* The team will use **two machine learning techniques**: **Neural Networks (MLP)** and **K-Nearest Neighbors (KNN)**, alongside exploratory data analysis (EDA).
* Discussion on **saving the project environment** (e.g., exporting a list of installed Python libraries) for reproducibility, as required by the instructor.

**Decisions Made:**

* **Stiven** will implement the **Neural Network model** since he has prior experience with it.
* **Charles** will handle the **KNN model**.
* **Exploratory Data Analysis (EDA) and preprocessing** will be done by **Miracle**, covering data cleaning, standardization, and dummy variable creation.
* **Writing tasks** will be shared, with everyone drafting their sections and finalizing the paper collaboratively.
* The team agreed to **meet on Monday at 7 PM** to review progress before the **Wednesday deadline**.

**Challenges Identified and How They Will Be Addressed:**

* **Dataset availability:** Some papers lacked accessible datasets, but the selected UC Irvine dataset met requirements.
* **Understanding Neural Network implementation:** The team ensured that sufficient details were provided in the source paper for accurate reproduction.
* **Reproducibility requirement:** The instructor requires an **environment export** (listing installed packages). The team discussed multiple methods (e.g., pip freeze, conda list) to generate this file.