

# **Capstone Project Proposal: Analysis of Cardiovascular Disease Risk Factors**

## **Overview**

The project aims to identify and analyze the key risk factors for cardiovascular disease (CVD), utilizing data analytics and machine learning. The focus will be on uncovering actionable insights that can predict CVD risk and identify high-risk patient groups, facilitating timely interventions.

## **Objectives**

1. **Key Risk Factors Identification:** Use feature importance and exploratory data analysis to highlight significant predictors of CVD.
2. **Predictive Model Development:** Construct an accurate model to assess the risk of CVD among individuals.
3. **Patient Segmentation:** Efficiently segment patients based on risk factors to identify those at highest risk.

## **Team Composition**

- **Project Manager:** Coordinates the project, ensuring timely progress and effective communication.
- **Data Analysts (2):** Conduct data cleaning and exploratory analysis, providing initial insights.
- **Data Scientists (2):** Develop the predictive model and perform clustering analysis, focusing on rapid iteration and validation.
- **Data Engineer:** Ensures data is accessible and structured for analysis, supporting rapid data operations.

## **Timeline and Methodology**

The project will be initiated immediately, with an aggressive but achievable timeline of 3.5 weeks, culminating in presenting findings and actionable strategies to stakeholders.

**Week 1:** Data preparation and exploratory analysis to quickly grasp key patterns and anomalies.

**Week 2:** Intensive focus on feature engineering and predictive modeling, with iterations based on preliminary results.

**Week 3:** Complete model validation, refine patient segmentation and draft insight reports.

**Week 3.5:** Finalize reports, prepare a comprehensive presentation, and rehearse findings.

## **Deliverables**

1. A concise report detailing exploratory data analysis findings and key insights.
2. An analysis of feature importance, emphasizing CVD risk factors.
3. A validated predictive model for assessing CVD risk, complete with documentation and user guidance.
4. A patient segmentation report, identifying and describing high-risk groups.
5. Targeted recommendations for interventions aimed at the identified high-risk groups.
6. A final presentation to stakeholders, summarizing findings, methodologies, and actionable recommendations.

## **Conclusion**

This project is designed to deliver critical insights into the prevention and management of CVD in a condensed timeframe. By leveraging a focused approach and the specialized skills of the analytics team, we aim to provide the client with valuable tools for identifying and intervening with high-risk individuals, thus making a significant impact on public health outcomes in a short period.