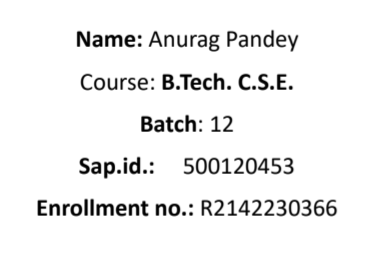


AIML ASSIGNMENT



**Assignments 1**

**Objectives:**

This project aligns with the Sustainable Development Goal of *Good Health and Well-being*, which emphasizes promoting healthier lives and well-being at all ages. The primary aim here is to use machine learning to uncover patterns and insights in health-related data that can lead to actionable improvements in public health. By predicting health outcomes and identifying key factors influencing well-being, the project seeks to inform better decision-making for policymakers and healthcare providers, ultimately contributing to healthier communities.

**Methodology:**

1. **Data Collection:**
   * The dataset, chosen specifically for its relevance to health and well-being, contains various indicators that influence health outcomes, such as lifestyle factors, access to healthcare, and demographic information. This dataset provides a foundation for exploring health trends and identifying risk factors.
2. **Data Preparation:**
   * Before analysis, the dataset underwent thorough cleaning and preprocessing. Steps included handling missing values, standardizing or normalizing continuous data, and converting categorical variables into a numerical format suitable for machine learning. To address any class imbalances (e.g., if some health outcomes are underrepresented), methods like SMOTE or ADASYN were applied, ensuring that the models can perform consistently across all categories.
3. **Model Selection and Training:**
   * A variety of machine learning models were applied to the data, including options like Logistic Regression, Decision Trees, Random Forest, and Support Vector Machines. Each model was trained using K-Fold Cross Validation, a method that divides the data into multiple parts, allowing each model to be validated across different data subsets for robust performance.
4. **Model Evaluation and Comparison:**
   * To assess and compare the models’ effectiveness, various performance metrics were used, such as accuracy, precision, recall, and F1-score. For a more comprehensive evaluation, the AUC score and confusion matrix were analyzed to understand each model's strengths and limitations. These comparisons helped identify the model best suited to reliably predict health-related outcomes.
5. **Result Interpretation and Insights:**
   * The final step involved interpreting the model outputs to uncover meaningful insights. By analysing patterns and significant factors, the project highlights key areas that may benefit from targeted interventions. These insights could serve as valuable guidance for healthcare strategies aimed at improving overall health and well-being in line with the SDG.



