SANSKAR SIJARIYA

Project: Medical Insurance Cost Prediction
Industry: Healthcare Industry

Project Requirement Document: Medical Insurance cost Predictions

1) Project Overview

This project aims to predict medical insurance costs based on various factors such as age, BMI, smoking status, region, etc. By using machine learning techniques, we can estimate the medical insurance costs for individuals and gain insights into the factors that contribute to higher or lower costs.

2) Objectives

- Develop a predictive model capable of identifying Medical Insurance Cost.

3) Scope

The project will encompass the following aspects:

- Data Collection: Gather historical disease outbreak data, weather data, population density, and healthcare infrastructure information from reliable sources.
- Data Preprocessing: Clean and preprocess the collected data, handle missing values, and ensure data consistency.
- Feature Engineering: Identify relevant features and engineer new ones that can contribute to accurate predictions.
- Model Development: Utilize time series analysis techniques, predictive modeling algorithms, and feature-engineered data to build a Medical insurance cost identification.
- Model Evaluation: Evaluate the model's performance using appropriate metrics such as accuracy, precision, recall, and F1-score.

4) Data Source

The project will require data from the following sources:
-Historical Medical Insurance cost from reliable Healthcare industry

5) Methodology

- **Data Collection and Preprocessing** Gather historical data of Medical Insurance cost, Clean, preprocess, and integrate the collected data to create a comprehensive dataset.
- **Feature Engineering** Identify relevant features and engineer new ones that can contribute to accurate predictions.
- **Encoding the Categorical Features of Datasets**
- **Model Development** Choose appropriate machine learning algorithms, such as Linear Regression,Logistic Regresion machines
- **Model Training and Validation** Split the dataset into training and validation sets, train the model, and validate its performance using suitable evaluation metrics.