# Health Monitoring System - Project Report

## 1. Introduction

The Health Monitoring System is a Big Data project designed to process, analyze, and visualize patient health data. It leverages Apache Spark, Hadoop, Kafka, and NoSQL databases to handle large-scale healthcare data efficiently. The project enables doctors and patients to track key health parameters and provides real-time alerts for critical conditions.

## 2. Objectives

* ✔ Generate large-scale patient health data
* ✔ Process data using Apache Spark & Hadoop
* ✔ Visualize health trends with an interactive dashboard
* ✔ Stream real-time data using Kafka
* ✔ Store patient feedback in a NoSQL database
* ✔ Use machine learning for risk classification

## 3. Technology Stack

🔹 Big Data Processing: Apache Spark, Hadoop (HDFS)

🔹 Real-time Streaming: Apache Kafka

🔹 Database: MongoDB (NoSQL)

🔹 Dashboard: Streamlit (Python)

🔹 Data Visualization: Matplotlib, Seaborn

🔹 Machine Learning: Decision Trees, Random Forest

## 4. Implementation

### 4.1. Data Generation

A Python script generates 10,000 patient records, each with attributes like demographics (Patient ID, Age, Gender) and health metrics (Blood Pressure, Sugar Level, Cholesterol, Haemoglobin). Data is stored in CSV format.

### 4.2. Data Processing with Spark

The patient data is loaded into Apache Spark, where MapReduce and Spark transformations are applied to compute statistics, filter, and classify data based on health risk levels.

### 4.3. Real-time Streaming with Kafka

Kafka Producer sends processed data to a Kafka topic, and Kafka Consumer (doctor's system) subscribes to receive real-time patient health updates. Patients can view real-time results and submit feedback.

### 4.4. NoSQL Storage for Feedback

Patient feedback is stored in MongoDB as `{Key: Patient ID, Value: Feedback}`. Data sharding techniques ensure scalability.

### 4.5. Dashboard Development

The dashboard provides filters (Age, Gender), visualizations (Histograms, Scatter plots, Trend analysis), and a machine learning model to predict high-risk patients.

## 5. Results & Analysis

* ✅ Identified high-risk patients based on BP, Sugar, and Cholesterol levels
* ✅ Real-time Kafka alerts notify doctors of critical cases
* ✅ Interactive visualizations for monitoring patient trends
* ✅ Collected patient feedback for system improvement

## 6. Future Enhancements

* 🚀 Mobile app integration for remote patient monitoring
* 🚀 Deep learning models for more accurate predictions
* 🚀 Automated doctor recommendations based on AI insights
* 🚀 Cloud deployment for scalability

## 7. Conclusion

The Health Monitoring System successfully implements Big Data techniques to process and analyze healthcare data. By combining Spark, Kafka, and NoSQL, it enables real-time monitoring and decision-making. Future advancements will focus on enhanced machine learning models and mobile accessibility.

\*\*Developed by:\*\* [Bogala suneel]

\*\*Date:\*\* [01/03/2025]

\*\*Email:\*\* [suneelreddy556700@gmail.com]