

Project Proposal

Capstone: The Art of Approximation

Ojas Chaturvedi, Zaheen Jamil, Saianshul Vishnubhaktula, Ritwik Jayaraman

November 21, 2023

Contents

Idea 2: AI-Based Precision Medicine Platform	1
Proposal	2

Idea 2: AI-Based Precision Medicine Platform

Purpose: Build an AI-powered platform that can analyze genetic data, medical records, and research to provide personalized treatment recommendations.

Member 1: Bioinformatics Scientist

- Develop algorithms for processing and analyzing large genomic datasets.
- Work on integrating genetic data with patient electronic health records (EHR).
- Lead the research on disease markers and drug response prediction.

Member 2: Machine Learning Engineer

- Build predictive models for disease risk and treatment outcomes.
- Implement natural language processing to extract insights from medical literature.
- Ensure the interpretability and fairness of AI models.

Member 3: Data Security Expert

- Design a secure infrastructure to protect sensitive health data.
- Ensure compliance with HIPAA and other health data protection regulations.
- Implement robust access controls and audit trails.

Member 4: Clinical Specialist and User Experience Designer

- Provide clinical expertise to guide the platform's development.
- Design the user interface for clinicians and patients, focusing on usability.
- Facilitate clinical trials and collect user feedback for continuous improvement.

Proposal

1. **Language:** Python, a simple and popular language for machine learning and data science due to its extensive libraries and frameworks.
2. **Objective:** Build an AI-powered platform that can analyze genetic data, medical records, and research to provide personalized treatment recommendations.
3. **Implementation:**
 - (a) **Overview of steps:**
 - i. **Data Collection:** Collect genetic data, medical records, and research data from various sources and databases.
 - ii. **Data Processing:** Process the data to extract relevant features.
 - A. **Homomorphic Encryption** will be utilized to protect sensitive health data instead of having to conform to HIPAA and other health data protection regulations.
 - iii. **Model Training:** Train machine learning models to predict disease risk and treatment outcomes.
 - iv. **Model Deployment:** Deploy the models on a secure platform to be used by clinicians and patients.
 - (b) **Libraries:** Matplotlib, Pandas, NumPy, Scikit-learn, TensorFlow, Keras, PyTorch, Flask, Django
 - (c) **Manual Work:**
 - i. Building own model?
 - ii. Collection of dataset of diseases and percent chance of symptoms