Project Title

Cirrhosis Prognosis Pro: AI-Driven Drug Prediction

Done By: Nayera Ammar

Department: AI & Data Science

Acknowledgment

We express our deepest appreciation to Dr. Moshira Ibrahim Ghaleb for her invaluable guidance and support throughout the project. We also extend our gratitude to [University Name] for providing the resources necessary to complete this research successfully.

Table of Contents

1. Abstract
2. Introduction

2.1 Project Name
2.2 Problem Statement
2.3 Project Idea and Objective

3. Project Overview

3.1 Data Engineering Section
3.1.1 Tools and Technologies
3.1.2 Data Pipeline
3.1.3 Data Warehousing
3.1.4 Data Visualization
3.2 Machine Learning Section
3.2.1 Tools and Technologies
3.2.2 Workflow
3.2.3 Model Selection and Training
3.2.4 Model Evaluation and Deployment

4. Implementation

4.1 Data Engineering Implementation
4.2 Machine Learning Implementation

5. Results

5.1 Data Insights and Visualization
5.2 Model Performance and Evaluation
5.3 User Interface (UI) for Clinician Interaction

6. Conclusion
7. Future Work
8. Team Members
9. References

1. Abstract

This project develops an AI-driven drug recommendation system for cirrhosis patients. By leveraging advanced machine learning models, the system predicts the most effective drug for treating cirrhosis based on patient-specific clinical data. It integrates data engineering for efficient data warehousing and processing with AI-powered predictive modeling to support clinicians in making informed, data-driven decisions about liver disease treatment.

2. Introduction

2.1 Project Name:

Cirrhosis Prognosis Pro: AI-Driven Drug Recommendation for Liver Disease

2.2 Problem Statement:

Cirrhosis is a complex liver disease where predicting the optimal drug treatment is challenging due to variations in patient responses. Current methodologies often lack precise, data-driven approaches to identify the most effective drug for individual patients. An accurate AI system is needed to help healthcare providers recommend personalized drug treatments for cirrhosis.

2.3 Project Idea and Objective:

This project aims to build a system using advanced data engineering techniques for storing and processing patient data, combined with machine learning algorithms to predict the most suitable drug for cirrhosis patients based on their clinical profiles. The goal is to provide real-time insights to clinicians, enabling more effective and personalized treatment decisions.

3. Project Overview

3.1 Data Engineering Section

3.1.1 Tools and Technologies

- Tools: Python, SQL, Power BI
- Libraries: Pandas, NumPy, Matplotlib, Seaborn

3.1.2 Data Pipeline

- Patient data ingestion
- Cleaning and preprocessing steps
- Storage in a structured database

3.1.3 Data Warehousing

- Centralized storage for efficient data access

3.1.4 Data Visualization

- Interactive dashboard development using Power BI

3.2 Machine Learning Section

3.2.1 Tools and Technologies

- Tools: Scikit-learn, XGBoost, TensorFlow
- Libraries: Pandas, NumPy
- Techniques: Gradient Boosting, Support Vector Machines, Neural Networks

3.2.2 Workflow

- Data Collection
- Feature Engineering
- Model Training
- Model Evaluation

3.2.3 Model Selection and Training

- Comparison of multiple models
- Selection based on performance metrics

3.2.4 Model Evaluation and Deployment

- Rigorous testing
- Deployment in real-time environments

4. Implementation

4.1 Data Engineering Implementation

- Detailed steps for data ingestion and preprocessing

4.2 Machine Learning Implementation

- Step-by-step guide for model training and evaluation

5. Results

5.1 Data Insights and Visualization

Power BI dashboards provide interactive insights into patient demographics, drug effectiveness, and treatment outcomes for cirrhosis patients.

- **Patient Demographics:** Age, gender, and medical history breakdown
- **Drug Effectiveness:** Comparative performance of different drugs

- **Treatment Success Rate:** Success rates segmented by disease severity and patient groups

5.2 Model Performance and Evaluation

Machine learning models, especially XGBoost and Random Forest, were evaluated based on metrics like accuracy, F1-score, and recall. The XGBoost model achieved the highest accuracy, demonstrating its ability to predict optimal drug treatments effectively.

- **Evaluation Metrics:**
  - Accuracy: 71% & 85.26

5.3 User Interface (UI) for Clinician Interaction

The UI was designed to enable seamless interaction between clinicians and the AI system.

- **Data Input Form:** Allows clinicians to enter patient details (e.g., age, gender, clinical stage, lab values).
- **Prediction Output:** Provides drug recommendations with confidence scores.
- **Visualization Links:** Directs clinicians to Power BI visualizations for further analysis.

6. Conclusion

Summarizes the project findings and their implications for clinical decision-making.

7. Future Work

1. Enhance models with larger and more diverse datasets.
2. Integrate genomic data for improved personalization.
3. Validate model predictions in real-world clinical settings.
4. Expand the drug library for broader treatment recommendations.
5. Develop real-time patient monitoring systems.

8. Work Done

Nayera Ammar Imam:

- Led machine learning efforts, including data cleaning, feature engineering, and model development.
- Focused on data warehousing and SQL queries, ensuring robust data pipeline and storage.
- Developed the UI and interactive Power BI dashboards for clinician insights.

9. References

- **Dataset:** Obtained from Kaggle and Google Colab.
- **Research Literature:** Sourced from Google Scholar.