<u>AI under India AI Mission</u>

Personalized Treatment Recommender System

A Full Stack Dev AI Project

Bachelors in Technology

by

Vir Garg (21124119)

Under the Mentorship of

Dr. Nisha Chaurasia

Department of Information Technology

Dr B.R Ambedkar National Institute of Technology

Jalandhar-144008, Punjab (INDIA)

June 24, 2024

Problem Statement

Medical practitioners use survival models to explore and understand the relationships between a patient's covariates (e.g., clinical and genetic features) and the effectiveness of various treatment options. Standard survival models like the linear Cox proportional hazards model require extensive feature engineering or prior medical knowledge to model treatment interaction at an individual level. While nonlinear survival methods, such as neural networks and survival forests, can inherently model these high-level interaction terms, they have yet to be demonstrated as effective treatment recommender systems.

Personalized medicine (PM) aims to tailor disease prevention, diagnosis, and treatment to individuals based on their genes, lifestyle, and environment. This domain is currently dominated by supervised learning, which allows physicians to select from limited sets of diagnoses or estimate patient risk based on symptoms and genetic information. Patients and interested organizations may potentially play an important role in the realization of PM, which is closely related to better disease assessment and more effective treatments based on individual health data paired with predictive analytics.

Background

PM relies on building statistical models able to predict personalized treatments tailored to the characteristics of specific patients. Developing such statistical models calls for large amounts of patient data. However, in practice, patients often prefer not to divulge certain kinds of data due to privacy concerns.

Medical researchers use survival models to evaluate the significance of prognostic variables in outcomes such as death or cancer recurrence and subsequently inform patients of their treatment options. A standard survival model is the Cox Proportional Hazards model (CPH), a semi-parametric model that calculates the effects of observed covariates on the risk of an event occurring (e.g., death). The model assumes that a patient's log-risk of failure is a linear combination of the patient's covariates, referred to as the linear proportional hazards condition. However, for providing personalized treatment recommendations, this assumption may be too simplistic. A richer family of survival models is needed to better fit survival data with nonlinear log-risk functions.

Objectives

1. **Develop a Robust AI Model:**
   - Utilize DeepSurv and Random Survival Forest (RSF) to create a reliable model that predicts personalized treatment recommendations based on patient data.
2. **Ensure High Predictive Accuracy:**
   - Measure and optimize the model's predictive accuracy using the concordance-index (C-index), ensuring it effectively ranks patient survival times.
3. **Enhance Data Privacy and Security:**
   - Implement stringent data privacy and security measures to protect patient information, addressing concerns about data sharing and confidentiality.
4. **Integrate with Existing Systems:**
   - Design the software to seamlessly integrate with existing hospital information systems and electronic health records (EHRs) for smooth implementation.
5. **Provide User-Friendly Interface:**
   - Develop an intuitive interface for medical practitioners to easily input patient data and receive actionable treatment recommendations.

Architecture of DeepSurv

To model nonlinear survival data, researchers have applied three main types of neural networks to the problem of survival analysis: classification methods, time-encoded methods, and risk predicting methods. This third type is a feed-forward neural network (NN) that estimates an individual's risk of failure. Faraggi-Simon's network is an example of a nonlinear extension of the Cox proportional hazards model.

In this section, we explain the main methodology for providing personalized treatment recommendations using DeepSurv, an open-source Python module that applies recent deep learning techniques to a nonlinear CPH network. DeepSurv is defined as a prognostic model, and we demonstrate how to use the network's predicted log-risk function to provide personalized treatment recommendations.

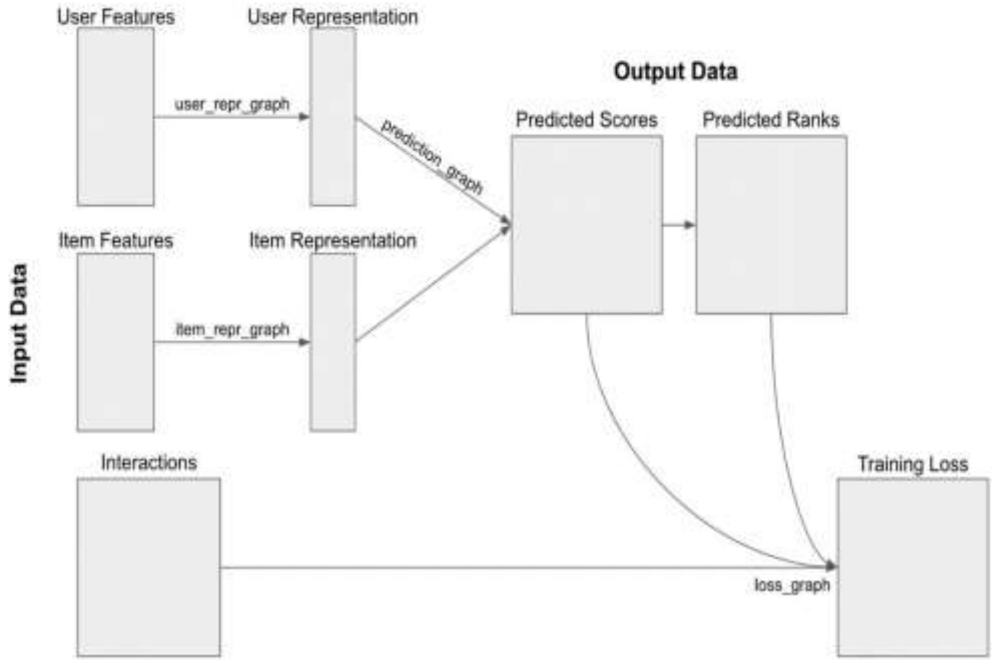

Fig:1 Recommendation system using Deep learning

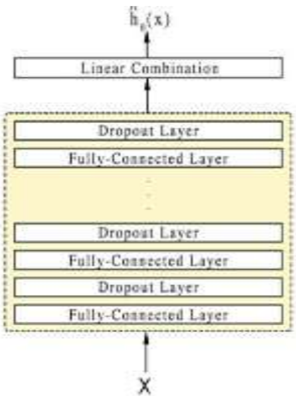

Fig:2 Diagram of the DeepSurv

Experimental Design

Dataset

Datasets such as those from the Indian Council of Cancer Research (ICCR) and local hospitals will be used for experimentation and evaluation.

Evaluation Measures

Measures such as accuracy of prediction of hazards will be computed by comparing the proposed treatment and hazards with the ones prescribed by doctors. The concordance-index (C-index) will be measured to evaluate the model's predictive accuracy on the survival data. The C-index is the most common metric used in survival analysis and reflects a measure of how well a model predicts the ordering of patients' death times.

Methodology

Step 1: Data Collection and Dataset Preparation

- **Survival data** will be collected from the ICCR dataset and local hospitals.
- **Activities**: Secure necessary permissions and agreements, aggregate and preprocess the collected data to ensure quality and consistency.

Step 2: Developing a DeepSurv-Based Treatment Recommendation System

- A recommended treatment for each patient in the test set will be built using DeepSurv and the Random Survival Forest (RSF).
- Without preselected treatment-interaction terms, the CPH model will compute a constant recommender function and recommend the same treatment option for all patients, effectively comparing the survival rates between control and experimental groups.
- DeepSurv and RSF can predict an individual's hazard per treatment by computing relevant interaction terms.

Step 3: Training and Experimentation on Datasets

- The DeepSurv model will be trained on large-scale datasets.
- **Activities**: Implement the DeepSurv architecture, train it on the survival data, optimize model parameters, and compare performance against standard survival models.

Step 4: Deployment and Analysis in Real-Life Scenario

- The trained and tested recommender system will be developed in a real-life scenario where historical medical records of cancer patients will be collected from local hospitals.
- **Activities**: Integrate the system with hospital databases for real-time data collection, monitor performance, gather feedback from medical practitioners and patients, and refine the system.

Importance at National Level

The development of a personalized treatment recommender system under the IndiaAI Mission is of great national importance. Personalized medicine is a transformative approach that can lead to more effective treatments, improved patient outcomes, and optimized resource allocation within the healthcare system.

This initiative aligns with India's vision to become a leader in AI-driven healthcare innovations and addresses critical challenges in medical research, such as the need for robust statistical models to predict personalized treatments while respecting patient privacy concerns. Successful implementation could revolutionize clinical practices, enhance the precision of medical interventions, and contribute significantly to the nation's public health goals.

India’s Healthcare Challenges and Statistics

1. **Healthcare System Ranking:**
    - India ranks 112th out of 195 countries in the Global Healthcare Access and Quality (HAQ) Index. This low ranking highlights significant gaps in the quality and accessibility of healthcare services.
2. **Mortality Rates Due to Inefficient Treatments:**
    - It is estimated that nearly 5.2 million people die each year in India due to medical errors and adverse effects of treatments. This includes deaths caused by incorrect diagnoses, wrong treatments, and ineffective management of chronic diseases.
3. **Cancer Statistics:**
    - India reports over 1 million new cancer cases annually, with a significant proportion of patients receiving late diagnoses or inappropriate treatment plans. Personalized treatment approaches could potentially reduce mortality rates by ensuring timely and accurate treatment.
4. **Chronic Disease Burden:**
    - The burden of chronic diseases like diabetes, cardiovascular diseases, and hypertension is rising in India. With over 77 million diabetics and increasing numbers of heart disease patients, personalized medicine can play a crucial role in managing these conditions more effectively.
5. **Hospital Resource Utilization:**
    - Hospitals often face resource constraints, with many operating at full capacity. Efficient treatment recommendations can help optimize the use of these resources, reducing the strain on healthcare facilities and improving patient throughput.

Deliverables with Timeline

Month 1-3: Data Collection and Preparation

- **Deliverable**: Collected survival data from ICCR and local hospitals.
- **Activities**: Secure permissions, aggregate and preprocess data.

Month 4-6: Model Development and Training

- **Deliverable**: Developed and trained DeepSurv model and RSF.
- **Activities**: Implement DeepSurv, train on survival data, optimize model parameters.

Month 7-9: Experimentation and Evaluation

- **Deliverable**: Conducted extensive experimentation and evaluation.
- **Activities**: Evaluate using C-index, compare predicted hazards with doctor-prescribed treatments.

Month 10-12: Deployment and Real-Life Analysis

- **Deliverable**: Deployed recommender system in real-life scenarios.
- **Activities**: Integrate with hospital databases, monitor performance, gather feedback, refine system.

Usage as a Product: Software

The personalized treatment recommender system will be a software product designed to integrate seamlessly with existing hospital information systems and electronic health records (EHRs). The software will provide medical practitioners with an intuitive interface to input patient data, receive personalized treatment recommendations, and track patient outcomes. This approach ensures flexibility, scalability, and ease of updates, which are crucial for adapting to evolving medical standards and practices.

In conclusion, the proposed personalized treatment recommender system represents a critical step forward in the application of AI to healthcare. It promises to enhance the precision and efficacy of medical treatments, thereby improving patient outcomes and contributing to the overall health and well-being of the nation.

Impact

1. **Improved Patient Outcomes:** By providing personalized treatment recommendations, this system can significantly enhance patient outcomes, leading to better recovery rates and improved quality of life.
2. **Optimized Resource Allocation:** Hospitals and clinics can utilize resources more efficiently by tailoring treatments to individual needs, reducing unnecessary treatments and associated costs.
3. **Advanced Medical Research:** The system will contribute to medical research by identifying effective treatment patterns and potential new therapies through data-driven insights.
4. **Enhanced Doctor-Patient Relationship:** With precise and personalized treatment plans, doctors can build stronger trust with patients, improving overall patient satisfaction.
5. **National Healthcare Advancement:** Aligning with the IndiaAI Mission, Personalised Recommender System will position India as a leader in AI-driven healthcare, fostering innovation and attracting global attention.