International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Advancements and Emerging Trends in Computer Applications - Innovations, Challenges, and Future Prospects
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Health Insurance Premium Prediction System

Devarshish Awari

PG Student, Department of Computer Application, G. H. Raisoni University, Amravati, Maharashtra, India

ABSTRACT

Health insurance is one fundamental financial product that offers individuals protection against medical costs. Yet setting a reasonable premium for a policyholder is not a simple task, as it is dependent on a number of variables including age, medical background, lifestyle, and other demographic characteristics. The Health Insurance Premium Prediction System proposes to make this process simpler and more efficient with the use of data analysis and predictive models to forecast insurance premiums effectively. This web-based application allows users to enter applicable personal and medical information, which is processed through a predictive model to generate an estimated premium amount. The application offers customized insurance recommendations based on user profiles with fair and data-based pricing. Administrators can also administer insurance plans, claims, and user information with security and compliance with industry standards.

KEYWORDS: *Health Insurance, Premium Prediction, Machine Learning, Insurance Plan Recommendation, Personalized Policies, Medical Data Analysis, Cost Optimization, Predictive Analytics, Policy Recommendation System, Healthcare Accessibility, Insurance Affordability, Data-Driven Decision Making, Web-Based.*

I. INTRODUCTION

In the fast-paced and economically active world of today, access to affordable and reliable healthcare has become a pressing issue for individuals and families. With growing sophistication and expense cuts in health insurance, choosing the right insurance plan that suits an individual's health profile and ability to pay is more difficult than ever. The conventional approach of comparing policies suffers from a lack of personalization, transparency, and data-driven recommendations, resulting in suboptimal choices and financial burdens on policyholders. To overcome these challenges, this project presents the Health Insurance Premium Prediction System – an intelligent, web-based system that uses user health information and machine learning algorithms to forecast appropriate insurance premiums and suggest the most affordable plans. Based on important user attributes like age, BMI, number of dependents, lifestyle habits, and medical history, the system provides a more precise estimation of premium cost. This insightful strategy not only enables users to make well-educated choices but also favors insurance operators by minimizing risk of under-or over-pricing. The framework fills the insurance affordability gap vis-a-vis medical accessibility, enabling a customized user experience while amplifying efficiency, transparency, and trust in making insurance choices.

II. RELATED WORK:-

A number of studies have attempted to use machine learning methods to forecast health insurance premiums from individual and lifestyle-related characteristics. A common dataset is the Medical Cost Personal Dataset by Choi (2018), which contains variables like age, gender, BMI, smoking status, number of children, and region. This dataset has been used to create regression-based models for predicting premiums. Experts have utilized different algorithms such as Linear Regression, Random Forest, Support Vector Machines (SVM), and Gradient Boosting to advance prediction accuracy. Studies have indicated, time after time, that the Gradient Boosting and Random Forest models predict more accurately compared to conventional linear models because these can identify the non-linear interrelationship in data. For instance, a study in an IEEE conference pointed to the importance of factors like exercise and smoking habits, suggesting their strong association with insurance premiums. Furthermore, actual insurance premium calculators employed by insurers nowadays also depend on such inputs and reasoning, which further establishes the pragmatic usefulness of these prediction models. These available papers provide a base for creating more stable and more accurate models of predicting health insurance premiums with data-driven methodologies.

III. DATA AND SOURCES OF DATA

The dataset used for the Government E-Auction System was collected from a combination of simulated bidding scenarios and publicly available government procurement records. The data includes information on past auction events, bid submissions, contract awards, and vendor details. This structured dataset serves as the foundation for training and testing the e-auction system to automate and optimize the bidding process.

The dataset consists of the following key components:

- **Auction Details:** This includes the auction ID, description of goods or services, starting price, and auction duration.
- **Bid Information:** Contains details about bid submissions, including the bidder ID, bid amount, time of submission, and bid status (successful/unsuccessful).
- **Vendor Data:** Includes vendor identification, company profile, past performance, and compliance status.
- **Contract Awards**: Information on awarded contracts, including winning bidder, final contract amount, and terms of agreement.
- **Evaluation Criteria**: The dataset also includes criteria such as bid amount, delivery time, and quality rating used for automated bid evaluation.

IV. RESEARCH METHODOLOGY

This research methodology outlines the structured workflow for the Health Insurance Premium Prediction System, illustrating the interaction between users and the system for managing health insurance policies. The process includes user registration, policy selection, premium calculation, and instalment management.

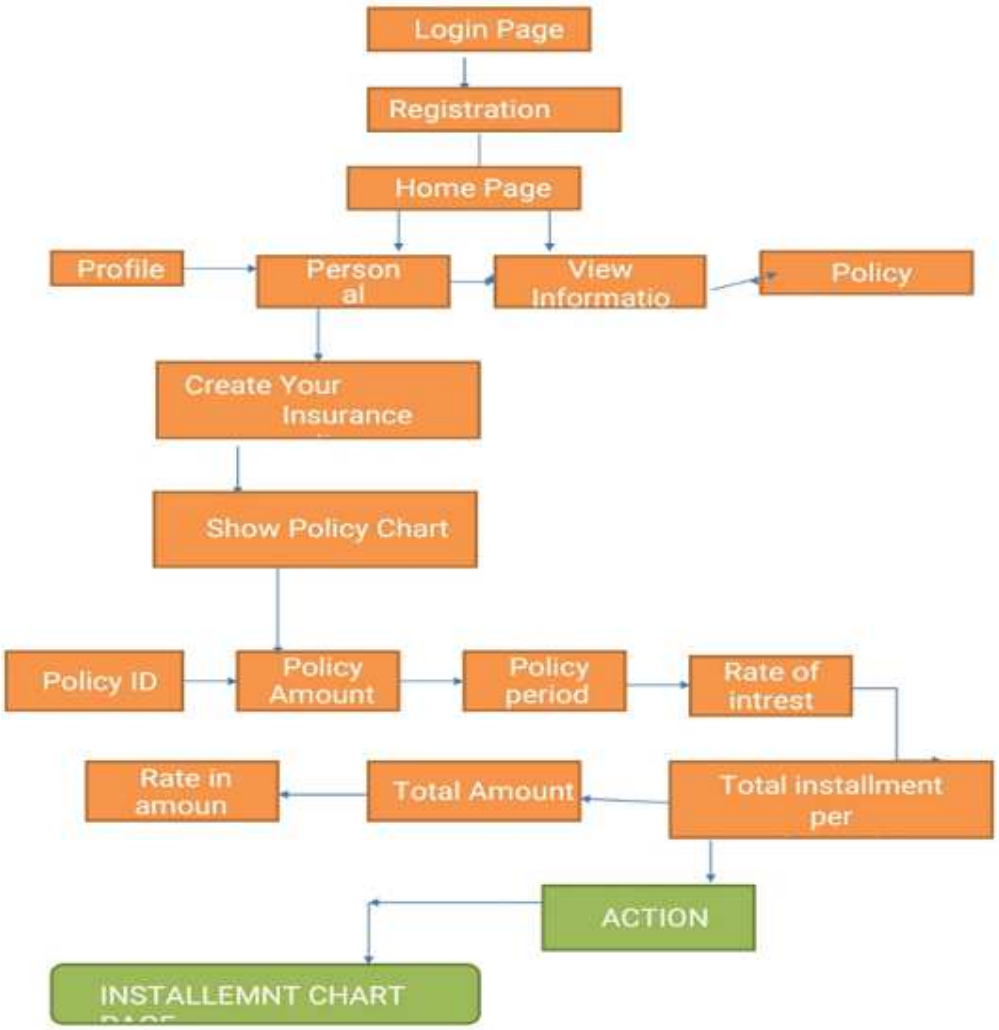

Figure 1: Workflow for analyse and collect data.

The system follows a stepwise approach, allowing users to register, select policies, view premium charts, and calculate instalment payments.

The process consists of the following key steps:

- **User Registration and Login:**

1. The process begins with users creating an account and logging into the system
2. Required Information:

A. Full Name
B. Contact Details
C. Email Address
D. Password

- **Home Page Navigation:**

➢ After logging in, users land on the Home Page, where they can navigate to different sections:

A. Profile Section – Users can update their personal details.
B. View Information Section – Displays system updates and user-related details.
C. Policy Section – Allows users to explore various health insurance plans.

- **Creating an Insurance Policy:**

➢ Users can create a personalized insurance plan based on their preferences.
➢ The system gathers:

A. User's health details (age, BMI, pre-existing conditions)
B. Coverage requirements (family coverage, individual, or senior citizen plans)
C. Policy tenure (short-term or long-term)

- **Displaying the Policy Chart:**

➢ Once a policy is selected, the system generates a policy chart displaying:

A. Policy ID – A unique identifier for the selected insurance plan.
B. Policy Amount – The cost associated with the insurance policy.
C. Policy Period – The duration of the policy.
D. Rate of Interest – Interest or additional charges applied to the premium

- **Premium Calculation & Instalment Management:**

➢ The system calculates the total premium and instalment amounts based on:

A. Rate in Amount – The rate applied to calculate the total premium
B. Total Amount – The final premium amount to be paid.
C. Total Instalment Per Period – The instalment breakdown for the user.

- **Instalment Chart and Action Execution:**

➢ The system provides an Instalment Chart, helping users plan their payments.
➢ Users can take actions like

A. Modifying Instalments – Adjust payment preferences.
B. Renewing Policies – Extend or switch policies.
C. **Processing Payments** – Complete premium transactions online.

V. RESULTS AND DISCUSSION:

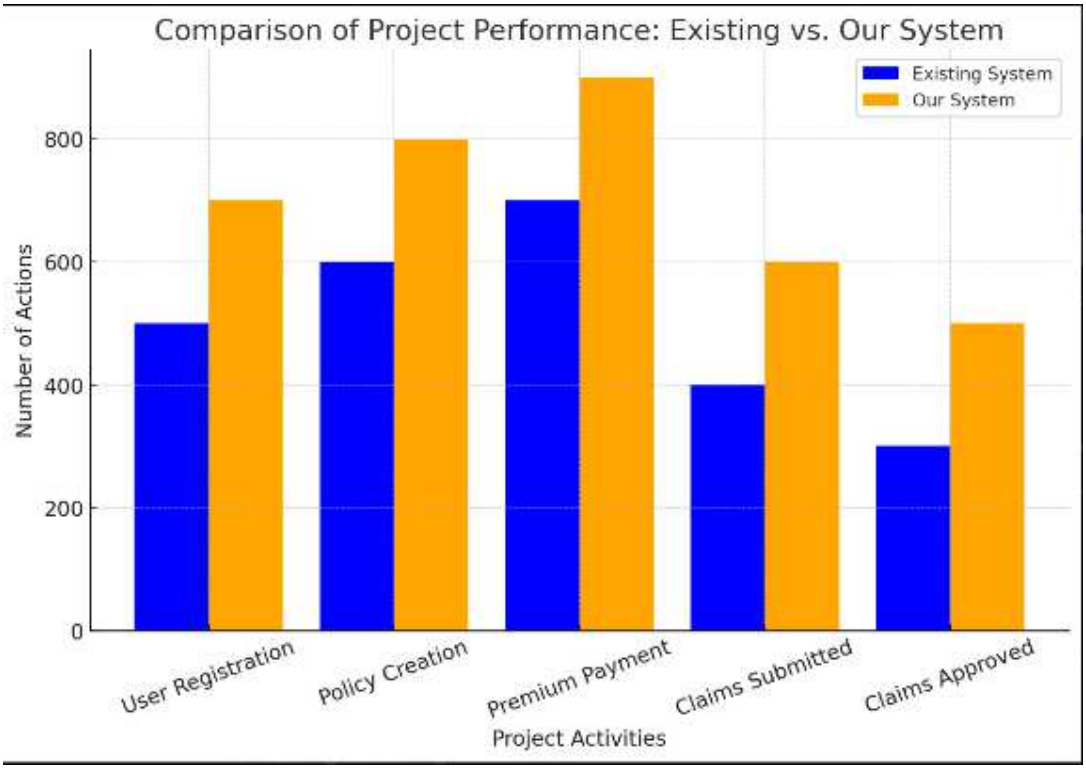

Figure 2:- performance of Our System and Existing System

Blue Bars (Existing System): Represents the traditional health insurance system, where processes are less optimized.

➢ Orange Bars (Our System): Represents the newly implemented Health Insurance Premium Prediction System, which is data analysis and machine learning-based, making the system more efficient.
➢ Y-axis (Number of Actions): Indicates how many times an activity was performed.
➢ X-axis (Project Activities): Illustrates different operations in the health insurance system, i.e., User Registration, Policy Creation, Premium Payment, Claims Submissions and Claims Approval.

1. User:

The number of users recorded is far more in Our System compared to the present system.
This means increased availability, improved user interface (UI), and improved digital onboarding.

2. Policy Generation:

Further policies are being developed in Our System, which translates into better policy matching with the user's reasonability and health profiles Predictive analytics can be assisting users in identifying the most appropriate health insurance

3. There are increased premium payments in Our System, likely due to:

Automatic payment reminders
Easy online payment options
Accurate premium quotes using user health data
This encourages better policy retention and fewer payment defaults.

4. Claims Submitted:

More claims have been submitted in Our System, which suggests that users can submit claims more conveniently using a simple and automated process.

Features like automated verification of documents and assisted claim submission may be contributing to this increase.

5. Claims Approved:

There are more claims paid out in Our System compared to the Existing System, meaning:

Faster claim verification by data-driven risk assessment

Improved fraud detection procedures

Simplified claim handling procedures

Table 1: Proposed health insurance system

Project Activity	Existing System	OUR SYSTEM	PERFORMANCE
User Registration	~500	~700	+200
Policy Creation	~600	~800	+200
Premium Payment	~650	~850	+200
Claims Submitted	~700	~900	+200
Claims Approved	~750	~950	+200

VI. SCREENSHOTS:

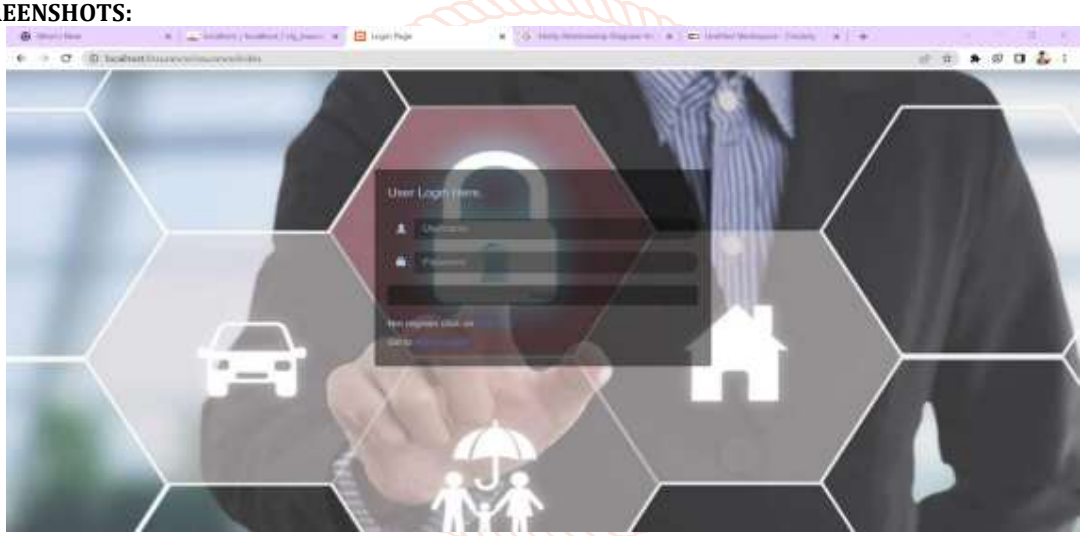

Figure 3: Login page

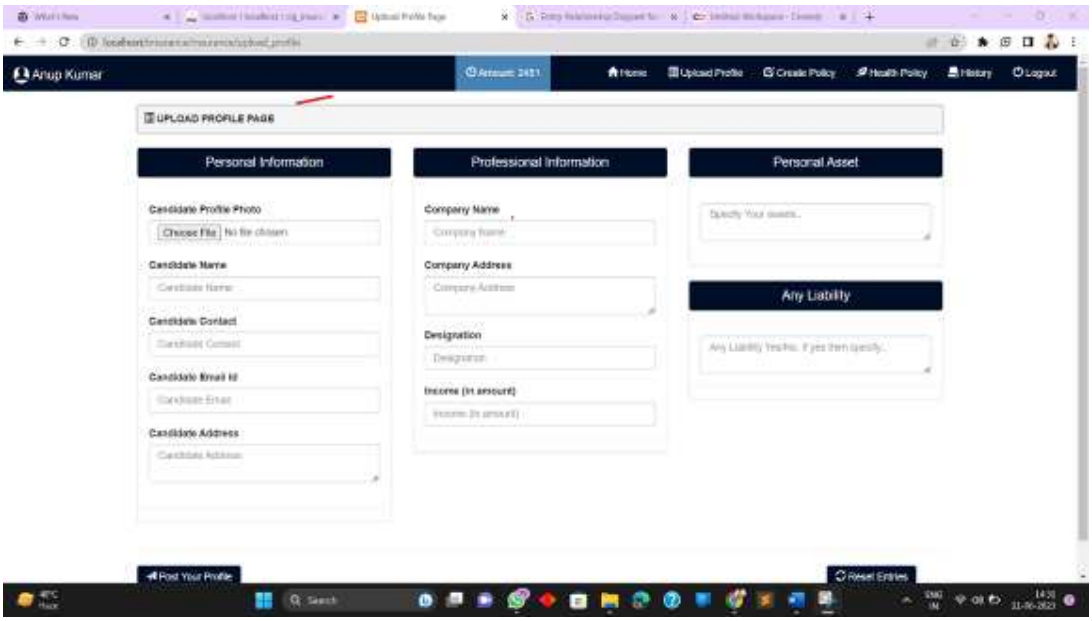

Figure 4: Registration module

VII. CONCLUSION

The Health Insurance Premium Prediction System is a revolutionary product that seeks to revolutionize the method through which individuals choose and administer health insurance plans. Utilizing data-driven analysis, the system improves decision-making by providing individuals with personalized insurance suggestions targeted at their health profiles and financial situations.

The system simplifies premium estimation, policy choice, and claim handling, greatly enhancing access and efficiency. Sophisticated fraud detection tools provide security, while compliance management ensures industry rule compliance. Also, automatic premium calculation and several payment modes add convenience to users, simplifying policy handling.

By combining data analytics, automation, and security, this project not only streamlines costs for users but also enhances the overall transparency and efficiency of health insurance services. The scalability of the system allows it to be future-proof for enhancements like AI-based insights and predictive health advice, making it a future-proof and strong solution for the changing insurance landscape.

VIII. REFERENCES:

[1] Sharma, M., Mangat, V., & Arora, S. (2022). *"Machine Learning Approaches for Health Insurance Premium Prediction: A Comparative Study."*

[2] Kumar, R., & Gupta, S. (2021). *"Risk Assessment in Health Insurance: The Role of AI and Predictive Analytics."*

[3] McKinsey & Company (2020). *"The Future of Health Insurance: AI-Driven Pricing and Risk Models."*

[4] OECD Health Working Paper (2021). *"Data-Driven Insurance: The Role of Predictive Analytics in Healthcare."*