Aiml Mini Project Report
on
Insurance Cost Predictor

by

Manish Shashikant Jadhav
2023301005

Mayur Krishna Solankar
2023301018

Vishesh Bhimji Savani
2022300100

Guide Name:
Prof. Swapnali Kurhade

Computer Engineering Department

Bharatiya Vidya Bhavan's

Sardar Patel Institute of Technology

Munshi Nagar, Andheri(W), Mumbai-400058

University of Mumbai

Nov 2024

1. Introduction

1.1 Problem Statement

To develop a machine learning-based predictive model for accurately estimating health insurance costs by analyzing key variables such as age, BMI, number of children, gender, smoking status, and region, addressing the challenges of manual cost estimation and improving the efficiency of insurance pricing strategies.

Key Challenges:

- Manual insurance cost calculation is time-consuming
- High potential for human error
- Complex interactions between multiple factors affecting insurance charges
- Need for an automated, data-driven approach to cost prediction

1.2 Literature Survey/Market Survey

Research Paper Link:

https://www.researchgate.net/publication/348559741_Predict_Health_Insurance_Cost_by_using _Machine_Learning_and_DNN_Regression_Models

1.3 Research Objectives

- Analyze factors influencing health insurance costs
- Implement linear regression and random forest models
- Generate predictive model for insurance cost estimation

2.Methodology

3.1 Data Source

- **Dataset:** Kaggle Medical Cost Personal Dataset
- **Features:** Age, BMI, number of children, gender, smoking status, region
- **Target Variable:** Medical insurance charges

3.2 Data Preprocessing

- Cleaned and prepared dataset
- Converted categorical variables to numeric values
- Split dataset into training and testing sets
- Created pickle file for random forest model

3.3 Machine Learning Models

1. Linear Regression
2. Random Forest Regression

3.4 Technology Stack

- Backend: Django (Python)
- Frontend: React.js
- Machine Learning: scikit-learn, pickle

3.5 System Architecture

3.5.1 Backend Development (Django)

- Create Django REST Framework for API endpoints
- Implement machine learning model loading from pickle file
- Develop prediction logic for insurance cost estimation
- Handle data validation and preprocessing
- Create secure API for model inference

3.5.2 Frontend Development (React)

- Design responsive user interface
- Create input forms for insurance variables
- Implement state management
- Develop prediction result display
- Connect frontend with Django backend API

3.5.3 Machine Learning Pipeline

- Data preprocessing
- Linear regression model
- Random forest regression model
- Model serialization using pickle
- Performance evaluation metrics

3.5.4 Integration Approach

- Django backend serves machine learning model
- React frontend consumes prediction API
- Seamless communication between frontend and backend
- Secure data transmission

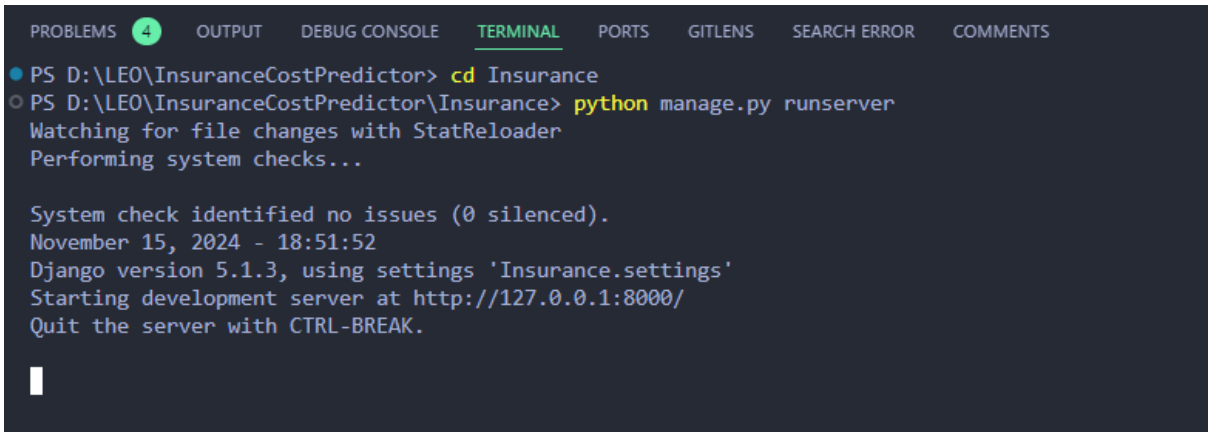
3. Implementation

Github Repo Link:

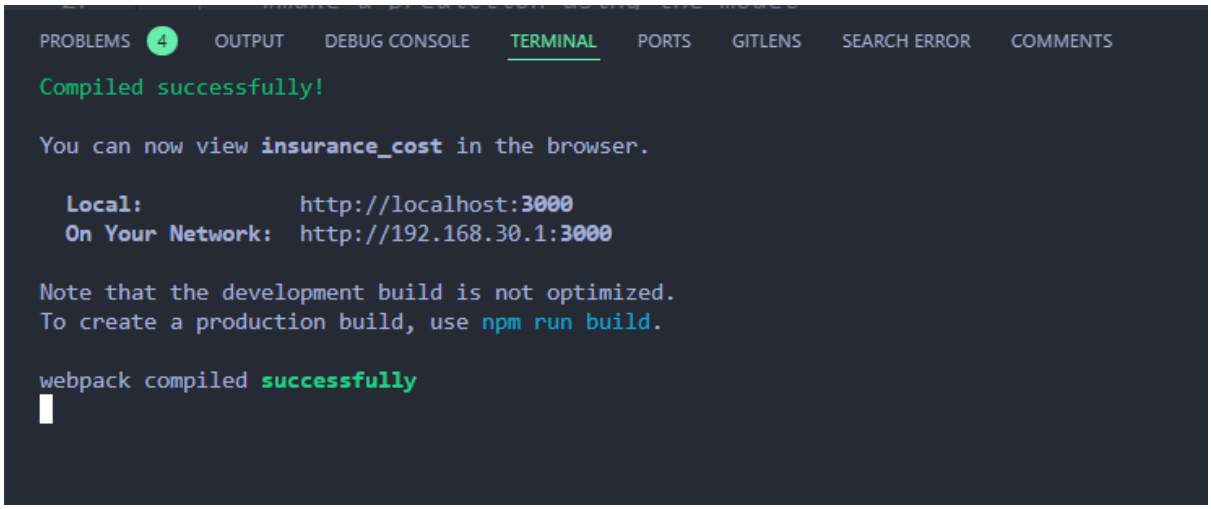
https://github.com/manishjadhav9/InsuranceCostPredictor

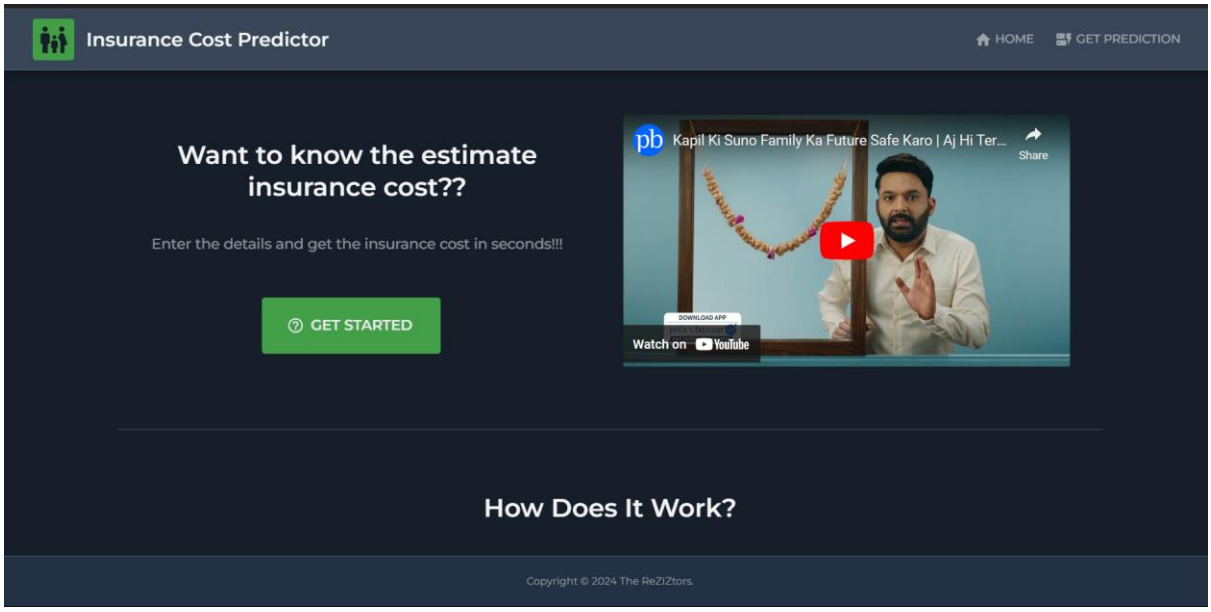
Collab Link:
https://colab.research.google.com/drive/1KF87J81hl5vbbIUWB_9pvp9Vj8L0BMGC

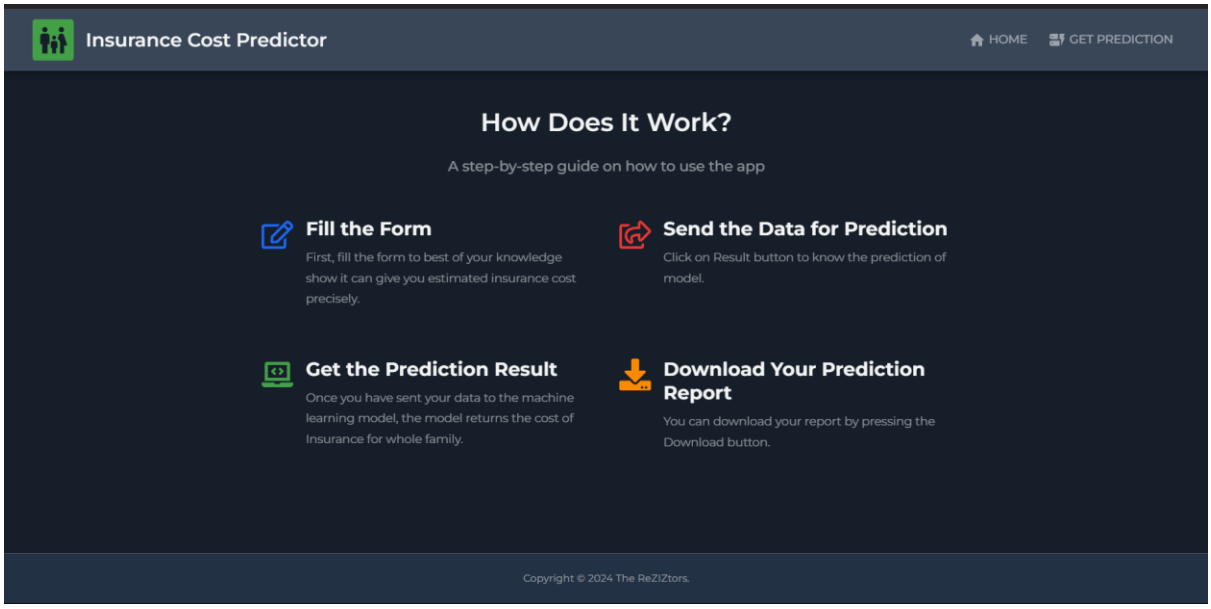
1. **Start Server:**

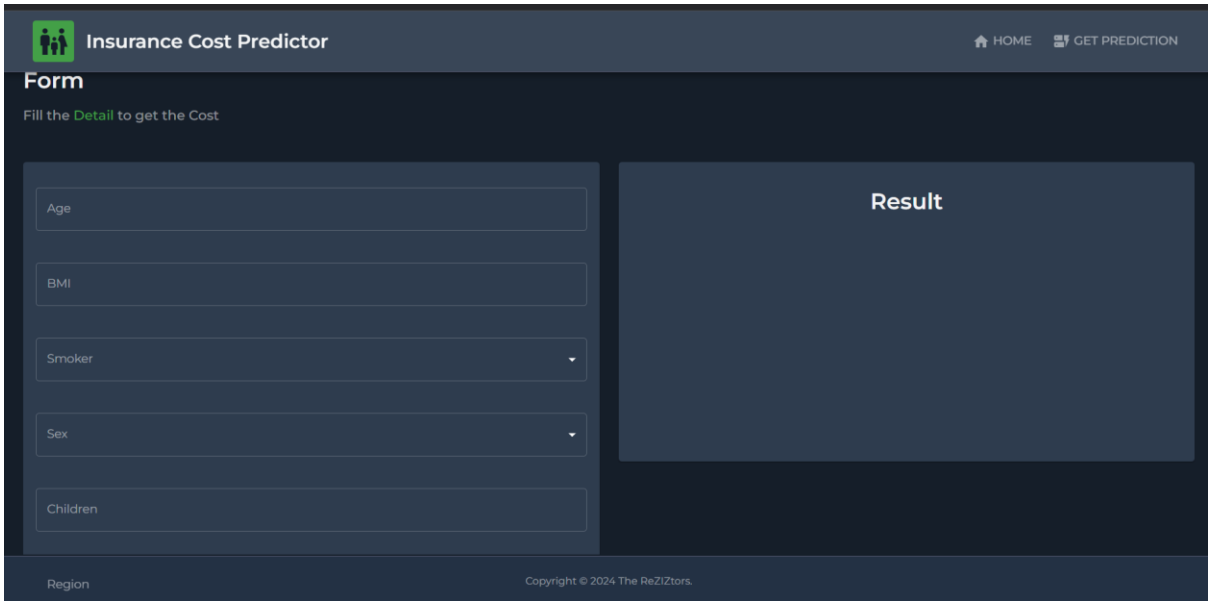


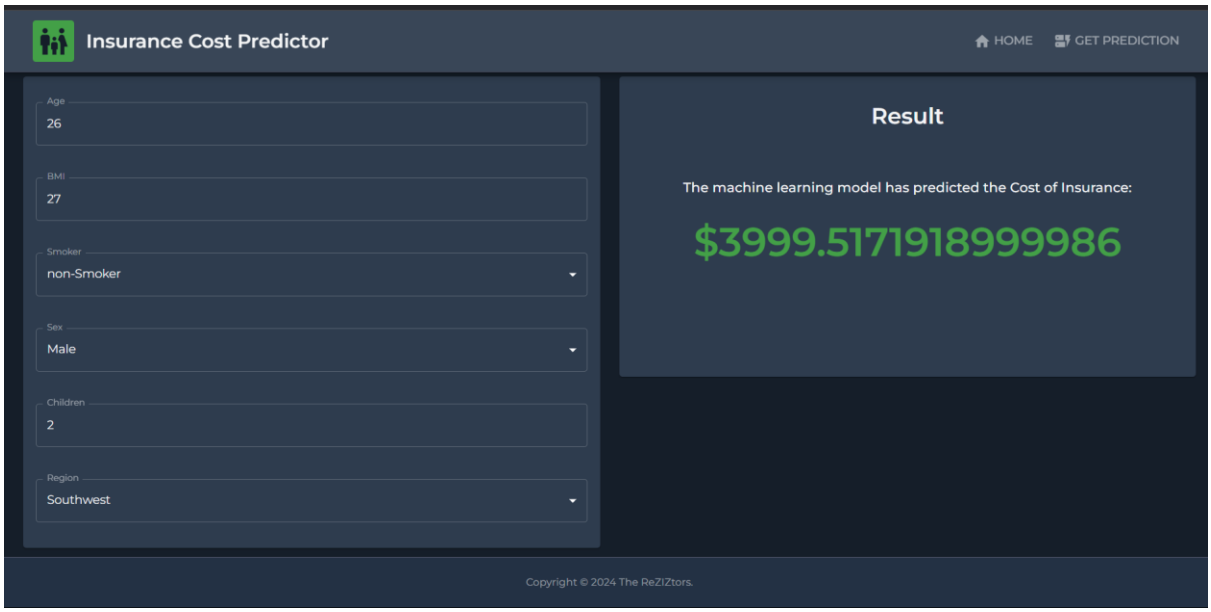
2. **Start frontend:**

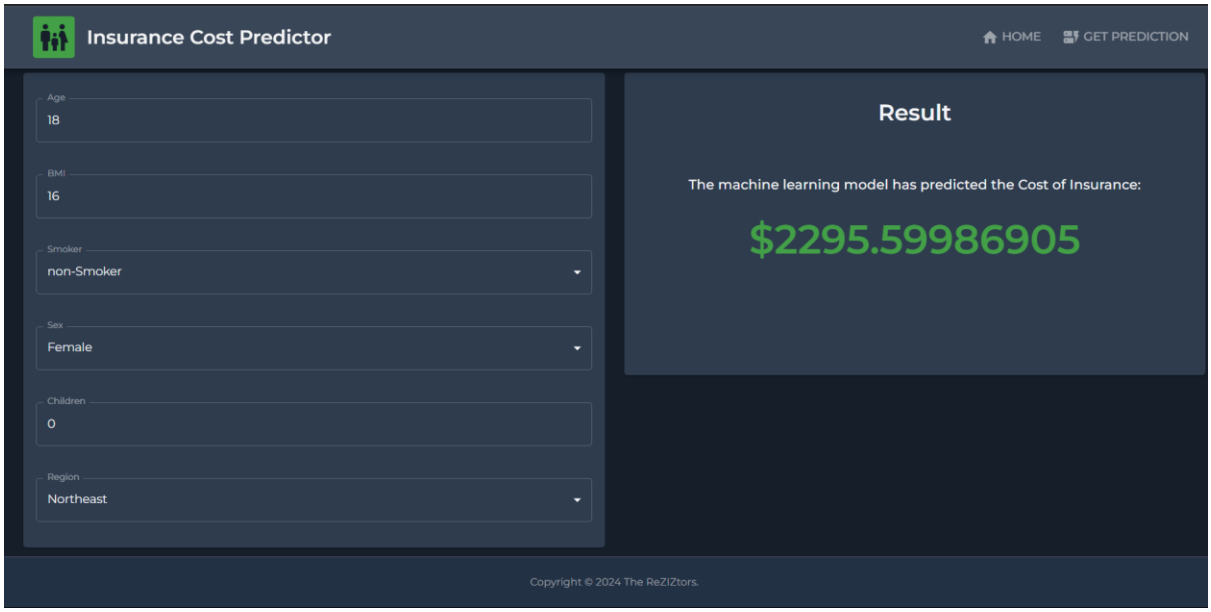



Insurance Cost Predictor
HOME
GET PREDICTION
Want to know the estimate insurance cost??
Enter the details and get the insurance cost in seconds!!!
GET STARTED
pb
Kapil Ki Suno Family Ka Future Safe Karo | Aj Hi Ter...
Share
Watch on YouTube
How Does It Work?
Copyright © 2024 The ReZIZtors.

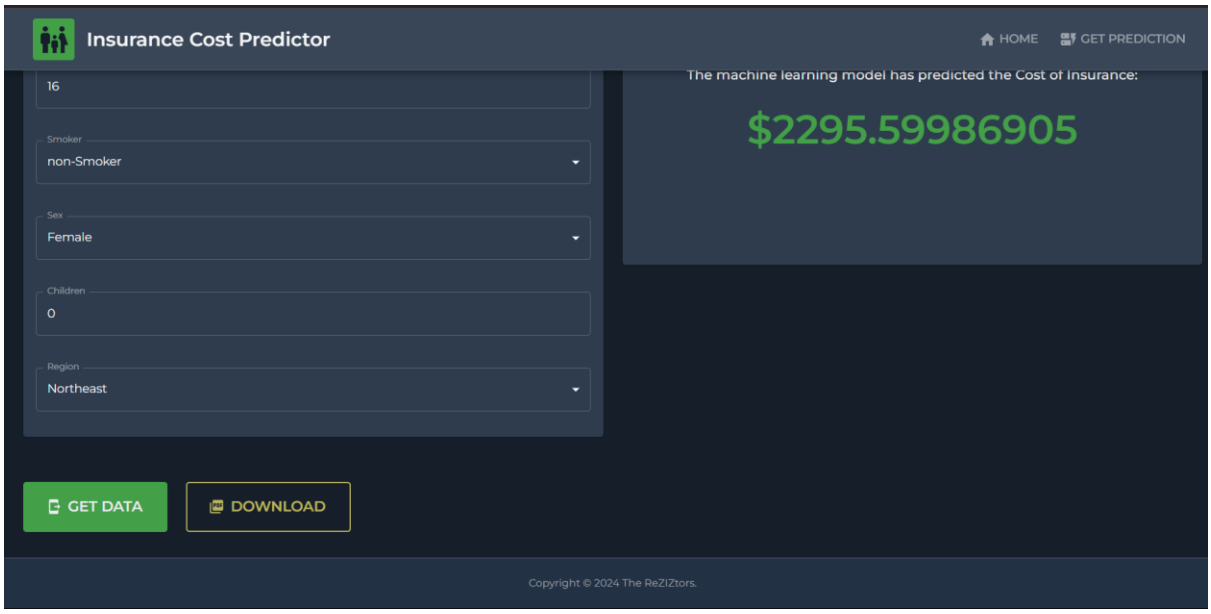

Insurance Cost Predictor
HOME
GET PREDICTION
How Does It Work?
A step-by-step guide on how to use the app
Fill the Form
First, fill the form to best of your knowledge show it can give you estimated insurance cost precisely.
Send the Data for Prediction
Click on Result button to know the prediction of model.
Get the Prediction Result
Once you have sent your data to the machine learning model, the model returns the cost of Insurance for whole family.
Download Your Prediction Report
You can download your report by pressing the Download button.
Copyright © 2024 The ReZIZtors.

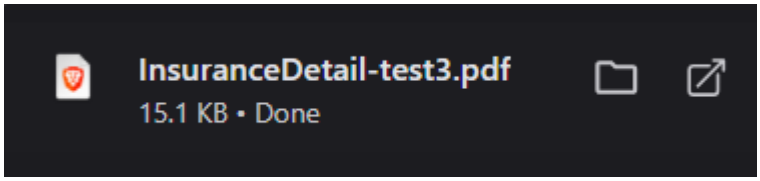

Insurance Cost Predictor
HOME
GET PREDICTION
Form
Fill the Detail to get the Cost
Age
BMI
Smoker
Sex
Children
Region
Result
Copyright © 2024 The ReZIZtors.


Insurance Cost Predictor
HOME
GET PREDICTION
Age
26
BMI
27
Smoker
non-Smoker
Sex
Male
Children
2
Region
Southwest
Result
The machine learning model has predicted the Cost of Insurance:
$3999.5171918999986
Copyright © 2024 The ReZIZtors.



3. Download the result:





4. Conclusion

The **Insurance Cost Predictor Project** is a testament to how technology can streamline financial planning and decision-making in the insurance sector. By leveraging machine learning algorithms, this project enables accurate predictions of insurance costs based on key factors such as age, gender, BMI, smoking habits, and more.

This tool not only benefits customers by offering transparency in cost estimation but also aids insurance companies in optimizing their pricing strategies, enhancing customer satisfaction, and promoting fairness. The project showcases the immense potential of data science in solving real-world problems, emphasizing the importance of predictive analytics in improving efficiency and decision-making across industries.

5. References

- https://www.kaggle.com/datasets/mirichoi0218/insurance?select=insurance.csv
- https://www.researchgate.net/publication/348559741_Predict_Health_Insurance_Cost_by_using _Machine_Learning_and_DNN_Regression_Models