A Project Report on

ESTIMATION AND PREDICTION OF

HOSPITALIZATION AND MEDICAL CARE COSTS.

**by**

**Team Leader: KUMMARI SURESH (20AT1A05G7)**

**Student 1 MANGALI SAI THIRUMALESH (20AT1A05I7)**

**Student 2 MEENIGA THARUN (20AT1A05H5)**

**Student 3 EDIGA SUMANTH GOUD(20AT1A05G4)**

**Student 3 BEECHPALLI JITHESH(20AT1A0J2)**

**Under the Guidance of**

**MRS.M. JAYA SUNITHA, MTech**

**Associate Professor**

****

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING**

**G. PULLAIAH COLLEGE OF ENGINEERING AND TECHNOLOGY**

**(Autonomous)**

(Approved by AICTE | NAAC Accreditation with ‘A’ Grade | Accredited by NBA (ECE, CSE, EEE, CE) |

Permanently Affiliated to JNTUA)

ABSTRACT

Medical costs are one of the most common recurring expenses in a person’s life. Based on different research studies, BMI, ageing, smoking, and other factors are all related to greater personal medical care costs. The estimates of the expenditures of health care related to obesity are needed to help create cost-effective obesity prevention strategies. Obesity prevention at a young age is a top concern in global health, clinical practice, and public health. To avoid these restrictions, genetic variants are employed as instrumental variables in this research. Using statistics from public huge datasets, the impact of body mass index (BMI) on overall healthcare expenses is predicted. A multiview learning architecture can be used to leverage BMI information in records, including diagnostic texts, diagnostic IDs, and patient traits. A hierarchy perception structure was suggested to choose significant words, health checks, and diagnoses for training phase informative data representations, because various words, diagnoses, and previous health care have varying significance for expense calculation. In this system model, linear regression analysis, naive Bayes classifier, and random forest algorithms were compared using a business analytic method that applied statistical and machine-learning approaches. According to the results of our forecasting method, linear regression has the maximum accuracy of 97.89 percent in forecasting overall healthcare costs. In terms of financial statistics, our methodology provides a predictive method.

**Contents**

**Abstract** 2

**Contents** 3

**List of Figures and Tables** 4

**CHAPTER 1 INTRODUCTION 4**

**CHAPTER 2 LITERATURE REVIEW 5**

**CHAPTER 3 PROPOSED METHOD 6**

**CHAPTER 4 EXPERIMENTAL RESULTS 7-9**

**CHAPTER 5 APPLICATIONS/ADVANTAGES 10-11**

**CHAPTER 6 CONCLUSIONS & FUTURE SCOPE 12**

**LIST OF FIGURES AND** **TABLES**

Database collected from IBM cogon’s about the ESTIMATION AND

PREDICTION OF HOSPITALIZATION AND MEDICAL CARE COSTS.

**CHAPTER 1**

**INTRODUCTION**

The incidence of overweight and obesity has increased significantly in most countries in recent decades. Excess weight is associated with an increased incidence of many chronic diseases, including vascular disease, respiratory disease, osteoarthritis, some cancer, type 2 diabetes, and premature death. There is consistent evidence that an increased BMI is associated with higher health costs, and these costs are expected to increase as obesity. Modelling uses machine-learning methods, in which the machine learns from the data and uses it to forecast new data [1, 2]. The most commonly predictive analytic model used is regression [3–6]. The proposed model for accurate prediction of future outputs has applications in banking, economics, e-commerce, sports, business, entertainment, etc. A method used to forecast healthcare costs for BMI is based on several factors. Multiple linear regression is one of the statistical techniques for estimating the relationship among the dependent (target) and independent variables. The regression method is commonly used to develop a system based on a number of factors to predict the cost [5–11].

The regression analysis is performed to determine the relationship among two or more variables with cause-effect relationships and to make predictions for the topic using the relationships [12]. If regression used one independent variable, then it is known as univariate regression analysis, or else if it used more than two independent variables then it is known as multivariate regression analysis. Linear regression involves initially uploading the data and then analysing the data. Subsequently, the data are cut, and then, the data are trained and separated to create the model. At last, it will evaluate the accuracy. The main aim of regression is to develop an efficient technique for predicting dependent properties from a set of characteristic variables. A regression problem is the actual or continuous value of the output variables, that is, area, salary, and weight. Regression can be defined as a statistical method used in applications such as predicting the healthcare costs. Regression is used to predict the relationship among the dependent variable and set of independent variables. There are various types of regression techniques available namely simple linear regression, multiple linear regression, polynomial regression, support vector regression, and random forest regression [13].

Fast-growing healthcare costs have become a significant challenge in several developed countries. Existing evidence suggests that healthcare costs have accumulated among a large number of BMI. Even though experiments have attempted to develop accurate models for predicting healthcare costs for BMI, their effectiveness is excellent due to the lack of detailed clinical information in the data used to create complex intervals and prognostic models. Numerous studies on more costs for obesity patient prognostic models have relied on self-report data and electronic health data from claims [14]. Data from laboratory tests are defined—these, more granular and detailed clinical information, lead to improvements in the prognostic model. A recent survey by health research programand claim data shows that there is an improvement in the performance of the machine-learning-based predictive model for health costs for obesity. Still, many insurers and providers worldwide are actively seeking an approach that can accurately predict obesity BMI [15].

However, despite the potential value of advanced machine-learning approaches for risk prediction, payers and providers still rely heavily on linear regression to manage and adapt their patient population [16, 17]. The slow adoption of advanced machine-learning techniques may be partly explained by the lack of familiarity with risk stabilization analysts with such techniques and the combination of complex interpretation and results required in practice. Machine-learning regression models are within the framework of standard linear regression and perform some sophisticated but less explicit machine-learning techniques [18, 19]. This study focused on fine linear regression models, which conducted a complete comparison of penalty regression with linear regression in forecasting overall health costs, which was not reported in the previously published literature. The major focus of this study is to estimate the health costs incurred due to obesity in the population.

Fast-growing healthcare costs have become a significant challenge in several developed countries. Existing evidence suggests that healthcare costs have accumulated among a large number of BMI. Even though experiments have attempted to develop accurate models for predicting healthcare costs for BMI, their effectiveness is excellent due to the lack of detailed clinical information in the data used to create complex intervals and prognostic models. Numerous studies on more costs for obesity patient prognostic models have relied on self-report data and electronic health data from claims [14]. Data from laboratory tests are defined—these, more granular and detailed clinical information, lead to improvements in the prognostic model. A recent survey by health research program and claim data shows that there is an improvement in the performance of the machine-learning-based predictive model for health costs for obesity. Still, many insurers and providers worldwide are actively seeking an approach that can accurately predict obesity BMI [15].

However, despite the potential value of advanced machine-learning approaches for risk prediction, payers and providers still rely heavily on linear regression to manage and adapt their patient population [16, 17]. The slow adoption of advanced machine-learning techniques may be partly explained by the lack of familiarity with risk stabilization analysts with such techniques and the combination of complex interpretation and results required in practice. Machine-learning regression models are within the framework of standard linear regression and perform some sophisticated but less explicit machine-learning techniques [18, 19]. This study focused on fine linear regression models, which conducted a complete comparison of penalty regression with linear regression in forecasting overall health costs, which was not reported in the previously published literature. The major focus of this study is to estimate thehealth costs incurred due to obesity in the population**.**

**CHAPTER 2**

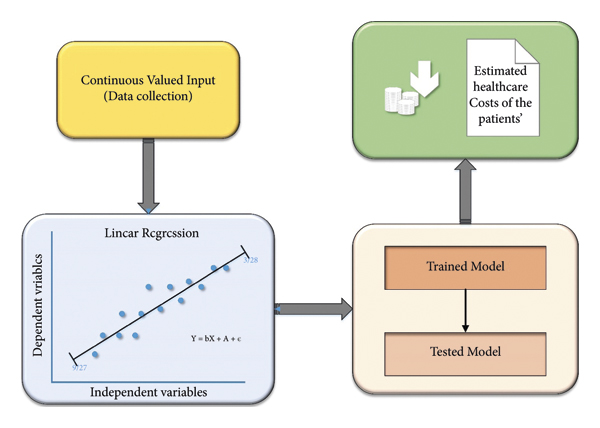
**LITERATURE REVIEW**

1. **Introduction to Healthcare Cost Prediction**:
   * The speaker started by introducing the challenges faced in healthcare cost management and the importance of employing predictive modeling techniques.
   * The audience was given a clear understanding of how increasing healthcare expenses necessitate the use of data-driven approaches for reliable cost predictions.
2. **Advantages of Cost Estimation**:
   * The lecture effectively highlighted the numerous advantages that accurate cost estimation models bring to the healthcare industry.
   * It elaborated on how precise cost predictions can aid in better resource planning, financial management, and budget allocation for hospitals and healthcare facilities.
   * The benefits of fair pricing, transparent patient billing, and effective insurance planning were emphasized, establishing a connection between cost estimations and patient satisfaction.
3. **Applications in Healthcare Management**:
   * The lecture provided practical examples of how cost estimation and prediction can be applied in various aspects of healthcare management.
   * It covered areas such as healthcare financial planning, insurance strategies, patient billing processes, and clinical decision support systems.
   * The speaker skillfully demonstrated how incorporating cost predictions in policymaking, regulation, and resource allocation can lead to efficient healthcare practices, especially during public health emergencies.
4. **Disadvantages and Limitations**:
   * A balanced viewpoint was presented by discussing the potential disadvantages and limitations of cost estimation models.
   * Ethical concerns related to balancing cost considerations with patient well-being were addressed, making the audience aware of the importance of maintaining an ethical stance in cost management.
   * The challenges associated with handling complex data and accounting for unpredictable events were explained, underscoring the need for a cautious approach in relying solely on predictive models.
5. **Population Health and Research Implications**:
   * The lecture delved into the implications of cost estimation in population health management and research endeavors.
   * It emphasized how cost predictions can be used to identify high-risk individuals, enabling proactive preventive care strategies.
   * The speaker illustrated how predictive models are valuable in healthcare research, including clinical trials and the evaluation of new medical technoloies.

**CHAPTER 3**

**PROPOSED METHOD**

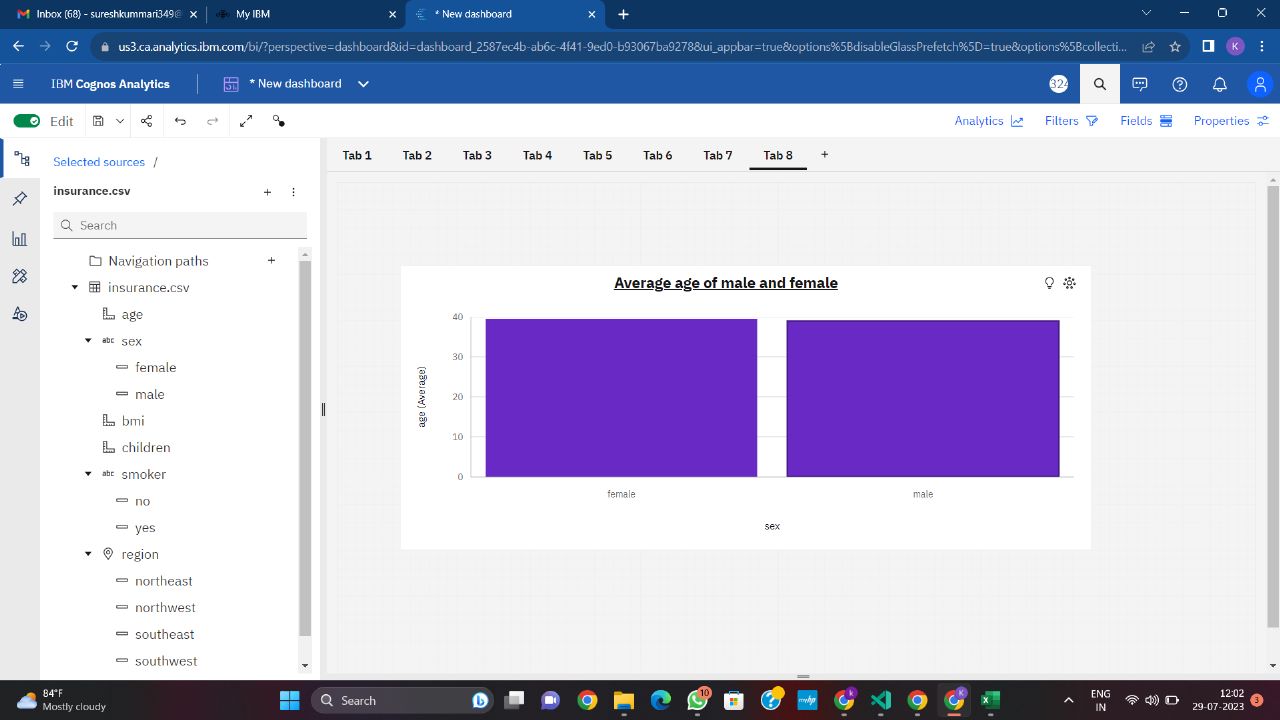
Linear regression is one of the most common supervisory machine learning statistical analysis techniques [26]. It is commonly used to find linear correlations between two or more responses and predictive variables. The technique is divided into two types depending on the number of variables in the model such as simple linear regression and multiple linear regression. A response variable corresponding to a predictive variable is simple linear regression. Whether more than two response variables correspond to predictive variables is known as multiple linear regression as shown in Figure [1](https://www.hindawi.com/journals/jhe/2022/7969220/fig1/). This work used linear regression to study the relationship among total maintenance and other properties in datasets to obtain the properties most affected by the total cost of maintenance. 75% of the data in the dataset were trained, and 25% of the data were tested. Then, Pearson’s correlation coefficient (PCC) for each simple linear regression sample was calculated. The PCC is determined and calculated by the following equation to find the parallel variability and strength of a linear regression relationship between two factors:

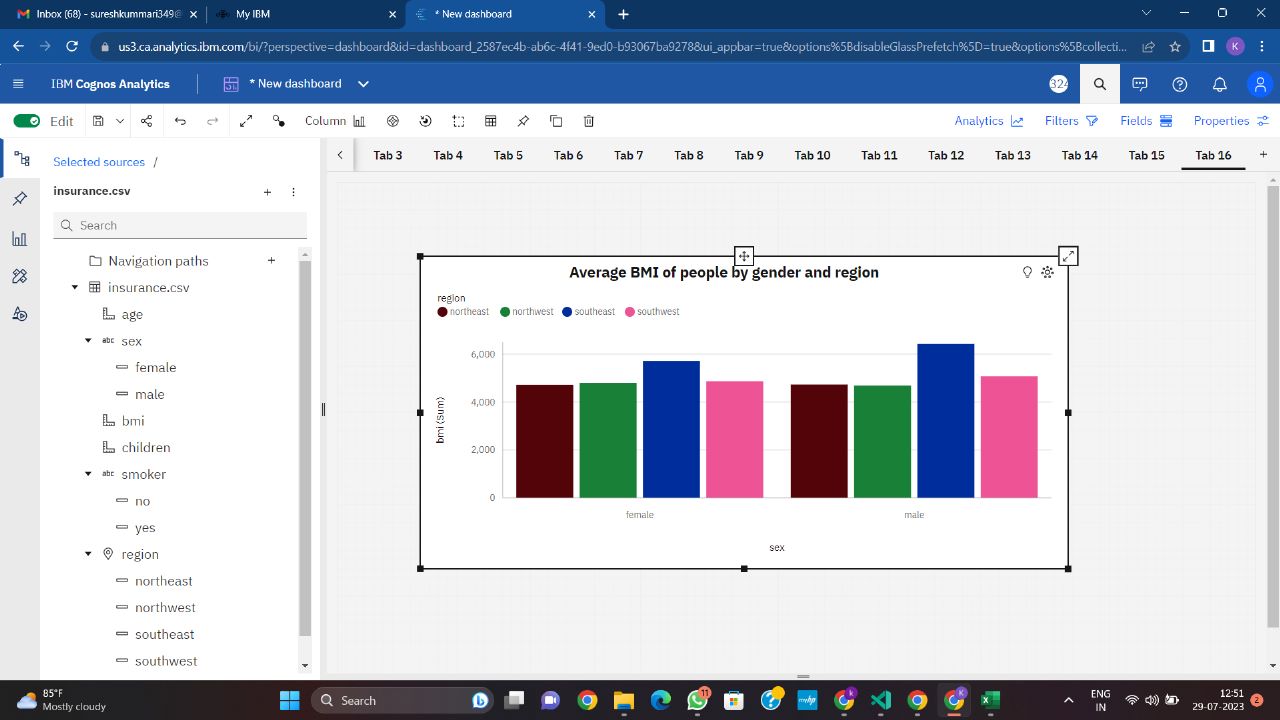


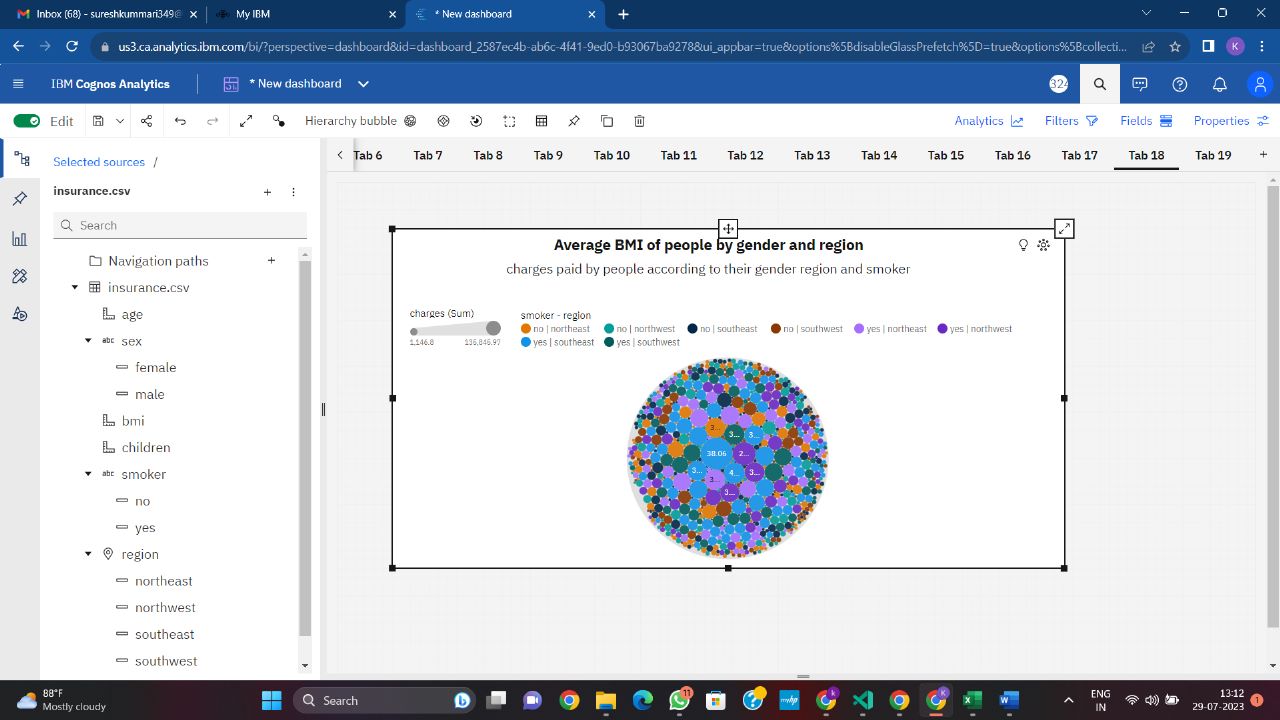
We provided a new linear regression that can easily demonstrate the reasons for producing a certain forecast regarding potential healthcare expenses, which is a useful capacity in the healthcare area. The linear regression algorithm is used to estimate the healthcare costs of the patients such as obesity (BMI) using certain devices such as smartphones and smart devices. For estimation, by the use of linear regression, supervised learning performs more accurately. By providing comprehensive evidence, regression methodology can be effectively used for prognosis in conjunction with the dataset. The domain and time accuracy will determine the prediction model and the estimation of healthcare expenses. The proposed method reduces the risk of overfitting, and also, training time is less. This method is effective in estimating the healthcare costs of patients with an accuracy rate of 97.89%. The extensive tests on a real-time world database have confirmed the efficiency of our method.

**CHAPTER 4**

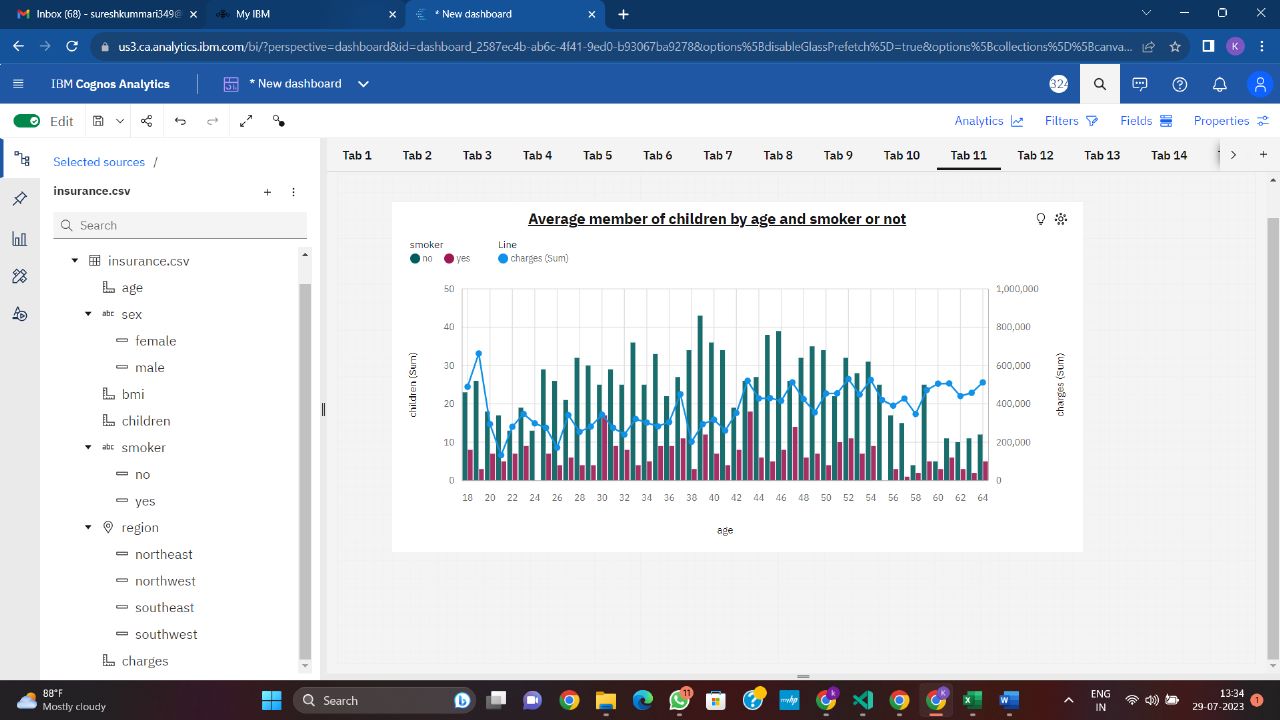
**EXPERIMENTAL RESULTS**

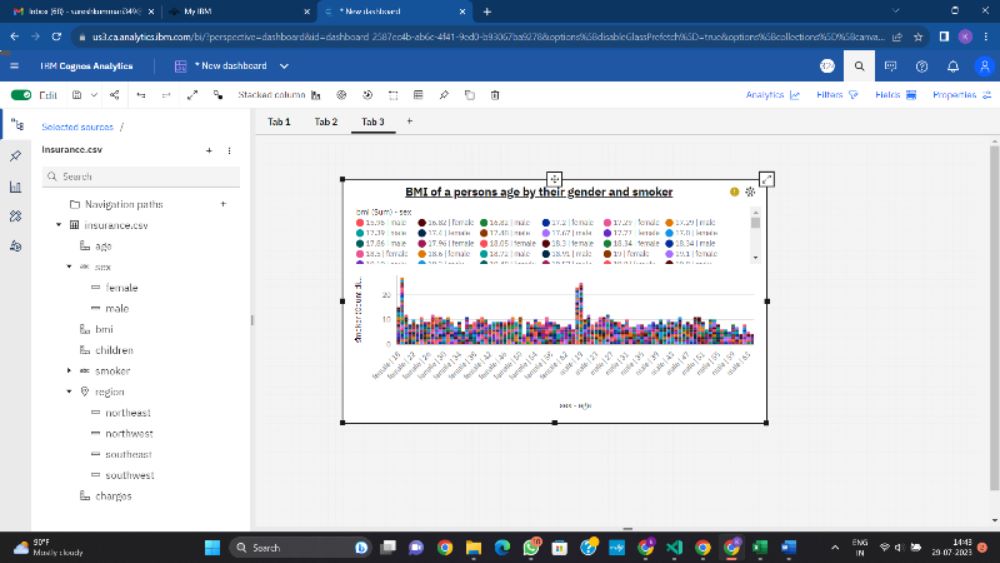












**CHAPTER 5**

**APPLICATIONS/ADVANTAGES**

**ADVANTAGES**

1. **Resource Planning**: Accurate cost estimation helps hospitals and healthcare providers allocate resources effectively. They can anticipate the number of patients, required medical supplies, and staffing needs, ensuring better preparedness.
2. **Financial Management**: Predicting costs in advance allows healthcare organizations to create more accurate budgets and financial plans. This, in turn, helps in optimizing spending and identifying areas for cost-saving measures.
3. **Pricing and Billing**: Hospitals can set appropriate prices for their services and treatments based on estimated costs. This ensures that patients are billed reasonably and reduces the likelihood of undercharging or overcharging.
4. **Risk Management**: Estimating and predicting costs can aid in risk assessment and insurance planning for healthcare providers. They can better prepare for unexpected expenses and potential financial challenges.
5. **Comparative Analysis**: Healthcare organizations can compare their costs with industry benchmarks and best practices, identifying areas where they can improve efficiency and reduce expenses.

DIS ADVANTAGES

1. **Complexity**: Healthcare costs can be influenced by numerous factors, making accurate predictions challenging. Variables like patient demographics, medical conditions, treatment plans, and reimbursement policies can significantly impact the final cost.
2. **Data Quality**: Reliable predictions depend on the availability and quality of data. Inaccurate or incomplete data can lead to flawed estimations, reducing the reliability of the forecasts.
3. **Unpredictable Events**: Unexpected events, such as disease outbreaks or natural disasters, can disrupt cost predictions. These events are difficult to account for in traditional forecasting models.
4. **Ethical Concerns**: Overemphasis on cost predictions might inadvertently prioritize financial gains over patient care. Healthcare decisions should always prioritize patient well-being.
5. **Overlooking Quality of Care**: Focusing solely on cost estimation can lead to a potential compromise in the quality of healthcare services. Providers may cut corners or opt for cheaper alternatives, affecting patient outcomes.
6. **Lack of Human Element**: Cost estimation models may not consider individual patient circumstances or account for the nuances of each case. This can result in generalized predictions that do not reflect the actual needs of patients.

**APPLICATIONS**

1. **Healthcare Financial Planning**: Healthcare organizations can use cost estimation and prediction to create accurate budgets, allocate resources effectively, and plan for future financial needs. This enables them to make informed decisions about investments in equipment, staffing, and infrastructure.
2. **Insurance and Payer Planning**: Insurance companies and payers can benefit from cost predictions to set appropriate premiums, design insurance plans, and anticipate future claim expenses. This helps in ensuring the sustainability and profitability of insurance products.
3. **Patient Billing and Financial Counseling**: Hospitals can use cost estimation to provide patients with upfront cost estimates for medical procedures and treatments. This transparency allows patients to plan for their healthcare expenses and explore financial assistance options if needed.
4. **Clinical Decision Support**: Cost predictions integrated into clinical decision support systems can help healthcare providers choose cost-effective treatment options without compromising patient outcomes. This can lead to better value-based care.
5. **Healthcare Policy and Regulation**: Governments and regulatory bodies can use cost estimation to assess the financial impact of healthcare policies, reimbursement rates, and changes in healthcare systems. This information guides policymaking and helps in the development of sustainable healthcare programs.
6. **Resource Allocation in Public Health**: During disease outbreaks or public health emergencies, accurate cost predictions are essential for allocating resources such as medical supplies, personnel, and hospital beds efficiently.
7. **Research and Clinical Trials**: Cost estimation can aid researchers in designing cost-effective clinical trials and assessing the potential economic impact of new medical interventions or technologies.
8. **Healthcare Vendor Management**: Healthcare organizations can use cost predictions when negotiating contracts with vendors for medical supplies, pharmaceuticals, and other services, ensuring they get the best value for their spending.
9. **Population Health Management**: Estimating future healthcare costs for specific patient populations helps in identifying high-risk individuals and implementing preventive care strategies, reducing overall healthcare expenses.
10. **Healthcare System Optimization**: Cost prediction models can be used to analyze and optimize various aspects of the healthcare system, such as patient flow, bed utilization, and appointment scheduling, leading to increased efficiency and cost savings.

**CHAPTER 6**

**CONCLUSIONS & FUTURE SCOPE**

**CONCLUSION :**

We provided a new linear regression that can easily demonstrate the reasons for producing a certain forecast regarding potential healthcare expenses, which is a useful capacity in the healthcare area. The linear regression algorithm is used to estimate the healthcare costs of the patients such as obesity (BMI) using certain devices such as smartphones and smart devices. For estimation, by the use of linear regression, supervised learning performs more accurately. By providing comprehensive evidence, regression methodology can be effectively used for prognosis in conjunction with the dataset. The domain and time accuracy will determine the prediction model and the estimation of healthcare expenses. The proposed method reduces the risk of overfitting, and also, training time is less. This method is effective in estimating the healthcare costs of patients with an accuracy rate of 97.89%. The extensive tests on a real-time world database have confirmed the efficiency of our method.

Inference

IBM Database extracted from smart internz.

Google.