****

**BENAZIR BHUTTO SHAHEED UNIVERSITY LYARI**

**FINAL PROJECT REPORT**

**(Artificial Intelligence)**

**Group Members:**

* Salman Abdul Rahim (BCS-616)
* Muhammad Saad (BCS-597)
* Zafar Hayat Khan (BCS-635)

**Submitted To:**

Ma’am Nadiya Mirbahar

**INSURANCE CHARGES PREDICTION USING**

**MACHINE LEARNING**

**INTRODUCTION:**

It has become a necessity for almost every person nowadays to sign up for the health insurance facility. These perks are being granted by many of the government as well as private insurance companies all across the globe. Concerning the value of insurance in the lives of individuals, it becomes important for the insurance companies to remain sufficiently precise while quantifying the amount covered by their policy of depicting the insurance charges over the customers.

Our project goal is to provide an idea about the insurance charges that a person require in order to fulfill the terms and conditions of an insurance company. According to our project, the criteria of insurance charges does not rely on any specific insurance company as we have taken out general estimations depending on the health status of an individual. It is purely a machine learning problem justifying multiple linear regression problem. The insurance dataset that we are using in this project is taken from kaggle while it contains seven features i.e. age, gender, bmi, children, smoker, region, and charges. In addition to this, the Machine Learning algorithm that we are up to use for our dataset is LinearSVR (Linear Support Vector Regressor).

**PROBLEM STATEMENT:**

There are various factors that influence the cost of insurance. These considerations contribute in the origination of some insurance policies. Each of the factors carry an important role when the amounts are being calculated. Ignorance of any factor might cause the policy to change from top to bottom. Hereby, it becomes critical to perform this task with high accuracy. Machine Learning plays an essential part to solve this issue. The technique of supervised learning pulls through the goal to automate the insurance price prediction. The linear regression model learns the insurance data from the past and can give out accurate insurance charges for the new set of data. This on the one hand reduces human effort to assume manual calculations, and on the other hand, can improve insurance company’s expediency through automation.

**TOOLS AND LIBRARIES:**

In our project, we have demonstrated the insurance charges prediction through the most robust programming language, Python. It is considered as one of the best and efficient tool for Artificial Intelligence, especially to compute and deal the Machine Learning problems. There are various significant packages and libraries with predefined methods which help in solving any of the Machine Learning query in a faster as well as systematic way. All we need is to know about these packages and libraries in order to deal with our data perfectly. Here in our project, we are using the following python libraries for data analysis, preprocessing and building our Machine learning model:

1. Pandas
2. Matplotlib
3. Seaborn
4. Sci-kit Learn

Let’s have a brief look on these libraries and what utilities they provide for the Python developers and A.I enthusiasts.

**METHODOLOGY:**

**MODEL EVALUATION AND RESULTS:**

**CONCLUSION:**