**PARTICIPATION REPORT**

**GROUP NO.: 3**

**COURSE TITLE: AIGS 1006 DEEP LEARNING**

**PROJECT TITLE: SNEAKER DATA ANALYSIS**

**GROUP MEMBERS AND CONTRIBUTIONS**

|  |  |  |  |
| --- | --- | --- | --- |
| **S/N** | **NAME** | **CONTRIBUTIONS** | **% (0 – 100)** |
| 1 | Vennila | 1. Explored various machine learning techniques to analyze and forecast prices. Initially tried to implement a Simple Linear Regression model, providing a foundational understanding of price trends. 2. Further, to enhance the predictive accuracy tried utilizing the Gradient Boosting method, known for its effectiveness in handling complex datasets. To compare the evaluation of these models, yield, provided insights towards the most effective strategy for accurate price forecasting. | 50% |
| 2 | Neha Parveen Mohammed | 1. Data Cleaning in Snowflake 2. Performing EDA and plotting different visualizations to understand the relationships between variables. 3. Created new features, Encoded the features using Label Encoder and plotted the feature importance using Random Forest Regressor. 4. Used StandardScaler to scale all the values to a particular scale. Used Feed Forword Neural Network to predict the future price by using the top 5 features. Tried other algorithms like Arima, Gradient Boost but could not complete it. | 100% |
| 3 | Sonali Jammichedu | 1. I have done research on data cleaning, EDA, and some of the algorithms. 2. I started performing comprehensive data cleaning, handling missing values, and performed label encoding, and tried one hot encoding for Categorical variables. 3. During EDA I visualized Random Forest Regressor to visualized feature importance’s. 4. Again, I conducted correlation analysis, calculating and visualizing the correlation matrix of the features. 5. For predictive modelling I applied Linear Regression using the stats model’s library considering Multiple independent variables. I split the data using train test split, evaluated the model. After evaluating the model, high mean square error was observed. 6. For data visualization purpose I have done actual vs predicted values plotting. | 70% |
| 4 | Sonjeet Kaur | 1. Explored the applicability of LSTM in the context of sales prediction, specifically targeting sneaker sales. 2. Studied the architecture, functionality, and underlying principles of LSTM to comprehend its suitability for time-series forecasting. | 50% |
| 5 | Sana Naseem | 1. Worked on ARIMA model with an aim to achieve future price prediction for sneakers, performed Augmented Dickey-Fuller test to check the stability over time for time series. Very high mean square error was observed using this algorithm. 2. Worked on linear regression model, converted categorical variables into dummy variables using one-hot encoding, plotted actual vs predicted prices along with residual plot. | 75% |