A

Major Project

On

PREDICTION OF USED CAR PRICES USING ARTIFICIAL NEURAL NETWORKS AND MACHINE LEARNING

(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

V. MADHUSUDHAN REDDY (217R1A05R3)

E. AAKASH REDDY (217R1A05M6)

K. KEERTHANA REDDY (217R1A05N6)

Under the Guidance of

K. PRAVEEN KUMAR

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR TECHNICAL CAMPUS

UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)

Recognized Under Section 2(f) & 12(B) of the UGCAct.1956,

Kandlakoya (V), Medchal Road, Hyderabad-501401.

April, 2025.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CMR
TECHNICAL CAMPUS
EXPLORE TO INVENT

CERTIFICATE

This is to certify that the project entitled "**PREDICTION OF USED CAR PRICES USING ARTIFICIAL NEURAL NETWORKS AND MACHINE LEARNING**" being submitted by **V. MADHUSUDHAN REDDY (217R1A05R3), E. AAKASH REDDY (217R1A05M6) & K. KEERTHANA REDDY (217R1A05N6)** in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, during the year 2024-25.

The results embodied in this project have not been submitted to any other University or Institute for the award of any degree or diploma.

K. Praveen kumar **Assistant Professor** **INTERNAL GUIDE**	**Dr. Nuthanakanti Bhaskar** **HoD**
Dr. A. Raji Reddy **DIRECTOR**	**Signature of External Examiner**

Submitted for viva voice Examination held on ____________________

ACKNOWLEDGEMENT

We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project, we take this opportunity to express our profound gratitude and deep regard to our guide **K. Praveen Kumar,** Assistant Professor for his exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by him shall carry us a long way in the journey of life on which we are about to embark.

We take this opportunity to extend our heartfelt appreciation to the Project Review Committee (PRC) Coordinators—**Dr. K. Maheswari, Dr. J. Narasimharao, Mrs. K. Shilpa, and Mr. K. Ranjith Reddy**—for their unwavering support, insightful guidance, and valuable inputs, which played a crucial role in steering this project through its various stages.

Our sincere appreciation also goes to **Dr. Nuthanakanti Bhaskar**, Head, for his encouragement and continuous support in ensuring the successful completion of our project.

We are deeply grateful to **Dr. A. Raji Reddy**, Director, for his cooperation throughout the course of this project. And, we extend our profound gratitude to Sri. **Ch. Gopal Reddy**, Chairman, Smt. **C. Vasantha Latha**, Secretary and Sri. **C. Abhinav Reddy**, Vice- Chairman, for fostering an excellent infrastructure and a conducive learning environment that greatly contributed to our progress.

We also acknowledge and appreciate the guidance and assistance provided by the faculty and staff of **CMR Technical Campus**, whose contributions have been invaluable in bringing this project to fruition.

Lastly, we sincerely thank our families for their unwavering support and encouragement. We also extend our gratitude to the teaching and non-teaching staff of CMR Technical Campus for their guidance and assistance. Their contributions, along with the support of everyone who helped directly or indirectly, have been invaluable in the successful completion of this project.

V. MADHUSUDHAN REDDY(217R1A05R3)

E. AAKASH REDDY (217R1A05M6)

K. KEERTHANA REDDY (217R1A05N6)

VISION AND MISSION

INSTITUTE VISION:

To Impart quality education in serene atmosphere thus strive for excellence in Technology and Research.

INSTITUTE MISSION:

1. To create state of art facilities for effective Teaching- Learning Process.
2. Pursue and Disseminate Knowledge based research to meet the needs of Industry & Society.
3. Infuse Professional, Ethical and Societal values among Learning Community.

DEPARTMENT VISION:

To provide quality education and a conducive learning environment in computer engineering that foster critical thinking, creativity, and practical problem-solving skills.

DEPARTMENT MISSION:

1. To educate the students in fundamental principles of computing and induce the skills needed to solve practical problems.
2. To provide State-of-the-art computing laboratory facilities to promote industry institute interaction to enhance student's practical knowledge.
3. To inculcate self-learning abilities, team spirit, and professional ethics among the students to serve society.

ABSTRACT

The number of cars on Mauritian roads has been rising consistently by 5% during the last decade. In 2014, 173 954 cars were registered at the National Transport Authority. Thus, one Mauritian in every six owns a car, most of which are second hand reconditioned cars and used cars. The aim of this study is to assess whether it is possible to predict the price of second-hand cars using artificial neural networks. Thus, data for 200 cars from different sources was gathered and fed to four different machine learning algorithms. We found that support vector machine regression produced slightly better results than using a neural network or linear regression. However, some of the predicted values are quite far away from the actual prices, especially for higher priced cars. This project proposes a predictive model using Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms to accurately estimate used car prices. By leveraging historical data and relevant features, the model learns complex relationships and patterns to provide reliable price estimates. The proposed approach aims to minimize prediction errors and enable informed decision-making for buyers, sellers, and industry professionals. Experimental results demonstrate the effectiveness of the model in predicting used car prices, highlighting its potential as a valuable tool in the automotive industry.

LIST OF FIGURES

.

LIST OF TABLES

TABLE OF CONTENTS

1. INTRODUCTION

1. INTRODUCTION

According to the data obtained from the National Transport Authority (2014), there has been an increase of 254% in the number of cars from 2003 (68, 524) to 2014 (173, 954), as shown in Figure 1. We can thus infer that the sale of second-hand imported (reconditioned) cars and second-hand used cars has eventually increase given that new cars represent only a very small percentage of the total number of cars sold each year. Most individuals in Mauritius who buy new cars also want to know about the resale value of their cars after some years so that they can sell it in the used car market.

Price prediction of second-hand cars depends on numerous factors. The most important ones are manufacturing year, make, model, mileage, horsepower and country of origin. Some other factors are type and amount of fuel per usage, the type of braking system, its acceleration, the interior style, its physical state, volume of cylinders (measured in cubic centimeters), size of the car, number of doors, weight of the car, consumer reviews, paint colour and type, transmission type, whether it is a sports car, sound system, cosmic wheels, power steering, air conditioner, GPS navigator, safety index etc. In the Mauritian context, there are some special factors that are also usually considered such as who were the previous owners and whether the car has had any serious accidents. Thus, predicting the price of second-hand cars is a very laudable enterprise. In this paper, we will assess whether neural networks can be used to accurately predict the price of secondhand cars. The results will also be compared with other methods like linear regression and support vector regression.

1.1 PROJECT PURPOSE

The primary objective of this project is to develop a predictive model that accurately estimates the prices of used cars based on their characteristics and features. By leveraging Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms, this project aims to provide a reliable and data-driven approach to used car pricing. The project seeks to address the challenges faced by buyers and sellers in the used car market, where pricing can be subjective and influenced by various factors. By analyzing a comprehensive dataset of used car listings, the project aims to identify the key factors that affect used car prices and develop a model that can accurately predict prices based on these factors.

1.2 PROJECT FEATURES:

This project features a comprehensive approach to predicting used car prices using Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms. The key features include data collection and preprocessing, development of an ANN model, exploration of various ML algorithms, and model evaluation using metrics such as Mean Absolute Error (MAE) and R-squared. The project also involves feature importance analysis to identify the most influential factors affecting used car prices. Additionally, the project will provide market insights and data visualization to facilitate informed decision-making. The technical features include the use of programming languages such as Python, libraries like TensorFlow and scikit-learn, and data storage in suitable formats. Overall, the project aims to deliver accurate price predictions, valuable market insights, and a robust predictive model that can benefit buyers, sellers, and industry professionals in the used car market.

2. LITERATURE SURVEY

2. LITERATURE SURVEY

Numerous studies have explored the application of Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms in predicting used car prices. Research has shown that ANNs can effectively capture complex relationships between variables, while ML algorithms like regression, decision trees, and random forests can provide accurate predictions. Studies have also highlighted the importance of feature selection, data preprocessing, and model evaluation in achieving reliable results. For instance, research has demonstrated that factors like make, model, year, mileage, and condition significantly impact used car prices. Furthermore, the use of techniques like gradient boosting and ensemble learning has been shown to improve prediction accuracy. This project builds upon existing research, leveraging the strengths of ANNs and ML algorithms to develop a robust predictive model for used car prices. By integrating insights from previous studies, this project aims to provide a comprehensive and accurate solution for predicting used car prices

The literature review reveals that factors such as make, model, year, mileage, and condition are significant predictors of used car prices. Techniques like gradient boosting and ensemble learning have been shown to improve prediction accuracy. This project builds upon existing research, leveraging the strengths of ANNs and ML algorithms to develop a robust predictive model for used car prices. By integrating insights from previous studies, this project aims to provide a comprehensive and accurate solution for predicting used car prices.

2.1 REVIEW OF RELATED WORK

The prediction of used car prices has been a topic of interest in the field of machine learning and artificial intelligence. Numerous studies have explored the application of Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms in predicting used car prices. These studies have demonstrated the effectiveness of ANNs in capturing complex relationships between variables and the accuracy of ML algorithms like regression, decision trees, and random forests in predicting prices. Research has also highlighted the importance of feature selection, data preprocessing, and model evaluation in achieving reliable results. For instance, studies have shown that factors like make, model, year, mileage, and condition significantly impact used car prices. Furthermore, the use of techniques like gradient boosting and ensemble learning has been shown to improve prediction accuracy. Some researchers have also explored the use of hybrid models that combine the strengths of multiple algorithms to achieve better performance. Overall, the literature review reveals that ANNs and ML algorithms have shown promising results in predicting used car prices, and this project aims to build upon existing research to develop a robust and accurate predictive model. By leveraging the strengths of ANNs and ML algorithms, this project seeks to provide a comprehensive solution for predicting used car prices, enabling buyers and sellers to make informed decisions.

1. Early Approaches and Traditional Techniques:

Early approaches to predicting used car prices using Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms have shown promising results. Initial studies focused on using simple regression models and decision trees to predict prices based on limited features. As the field evolved, researchers began exploring more complex models, such as ANNs and ensemble methods, which demonstrated improved accuracy. These early studies laid the groundwork for more advanced approaches, highlighting the potential of ANNs and ML in predicting used car prices. They also identified key challenges, such as data quality and feature selection, that continue to influence research in this area. Building on these foundations, recent studies have further refined models and techniques, pushing the boundaries of prediction accuracy and reliability.

2. Machine Learning-Based Approaches:

Machine learning-based approaches have been widely explored for predicting used car prices. Researchers have employed various algorithms, including regression, decision trees, random forests, and support vector machines, to develop predictive models. These models have been trained on datasets containing features such as make, model, year, mileage, and condition. Studies have shown that machine learning models can accurately predict used car prices, outperforming traditional pricing methods. Some researchers have also experimented with ensemble methods, combining multiple models to improve prediction accuracy. Additionally, techniques like feature engineering and hyperparameter tuning have been used to optimize model performance. Overall, machine learning-based approaches have demonstrated significant potential in predicting used car prices, enabling buyers and sellers to make informed decisions. This project builds upon these advancements, leveraging the strengths of machine learning algorithms and artificial neural networks to develop a robust predictive model for used car prices.

3. Deep Learning-Based Approaches:

Deep learning-based approaches have gained significant attention in predicting used car prices. Researchers have employed Artificial Neural Networks (ANNs) and other deep learning architectures to develop predictive models. These models can learn complex patterns and relationships in large datasets, capturing non-linear interactions between features. Studies have shown that deep learning models can outperform traditional machine learning approaches, providing more accurate predictions. Techniques like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been explored, although ANNs remain a popular choice. Hyperparameter tuning, regularization, and optimization techniques are crucial for improving model performance. By leveraging deep learning's ability to learn from complex data, this project aims to develop a robust predictive model for used car prices, providing valuable insights for buyers and sellers.

4. Recent Advances: Attention Mechanisms & Transformer-Based Models:

Recent studies have explored the application of attention mechanisms and Transformer-based models in predicting used car prices. These advanced techniques allow models to focus on relevant features and weigh their importance, leading to improved prediction accuracy. Attention mechanisms enable the model to selectively concentrate on specific input features, while Transformer-based models leverage self-attention to capture complex relationships between variables. By incorporating these recent advances, predictive models can better capture nuanced patterns in used car pricing data, providing more accurate and reliable predictions. This project aims to investigate the potential of these cutting-edge techniques in enhancing the prediction of used car prices.

5. Comparison with the Proposed Approach:

The proposed approach for predicting used car prices using Artificial Neural Networks (ANNs) and Machine Learning (ML) will be compared to existing methods, highlighting key differences and potential improvements. Our approach aims to leverage the strengths of ANNs and ML algorithms to develop a robust predictive model that accurately estimates used car prices. By comparing our approach to existing methods, we can identify areas of improvement and demonstrate the potential benefits of our proposed solution. This comparison will focus on prediction accuracy, feature importance, model complexity, and data requirements, providing a comprehensive evaluation of the proposed-approach.

2.2 DEFINITION OF PROBLEM STATEMENT

The problem addressed in this project is predicting used car prices accurately, considering factors like make, model, year, mileage, and condition. Developing a robust predictive model using Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms is crucial to capture complex relationships between these factors and provide reliable price estimates. This project aims to create an effective predictive model that minimizes errors and enables informed decision-making for buyers and sellers in the used car market.

2.3 EXISTING SYSTEM

Predicting the price of second-hand cards has not received much attention from academia despite its huge importance for the society. Bharambe and Dharmadhikari (2015) used artificial neural networks (ANN) to analyse the stock market and predict market behaviour. They claimed that their proposed approach is more accurate than existing ones by 25%. Pudaruth (2014) used four different supervised machine learning techniques namely kNN (k-Nearest Neighbour), Naïve Bayes, linear regression and decision trees to predict the price of second-hand cars. The best result was obtained using kNN which had a mean error of 27000 rupees. Jassbi et al. (2011) used two different neural networks and regression methods to predict the thickness of paint coatings on cars. The error for the final thickness of the paint was found to be 2/99 microns for neural networks and 17/86 for regression. Ahangar et al. (2010) also compared the use of neural networks with linear regression in order to predict the stock prices of companies in Iran. They also found that neural networks had superior performance both in terms of accuracy and speed compared to linear regression. Listiani (2009) used support vector machines (SVM) to predict the price of leased cars. They showed that SVM performed better than simple linear regression and multivariate regression. Iseri and Karlik (2009) used neural networks to predict the price of automobiles and achieved a mean square error of 8% compared with 14.4% for regression. Yeo (2009) used neural networks to predict the retention rate for policy holders of automobile insurance.

Limitations of Existing System

Despite improvements in video content classification, the existing system suffers from the following challenges:

1. Limited focus on second-hand car price prediction: Few studies specifically target predicting second-hand car prices.
2. Variability in accuracy: Different studies report varying levels of accuracy, with some achieving high error rates.
3. Limited comparison of techniques: Few studies comprehensively compare multiple ML and ANN techniques.
4. Context-specific results: Findings may not generalize across different markets or regions.

2.4 PROPOSED SYSTEM

In order to carry out this study, data have been obtained from different car websites and from the small adverts sections found in daily newspapers like L'Express and Le Defi. The data was collected in less than one month interval (i.e. in the month of August in 2014) because like other goods, the price of cars also changes with time. Two hundred records were collected. The data comprises of different features for second-hand cars such as the year (YEAR) in which it was manufactured, the make (MAKE), engine capacity (ENGINE) measured in cubic centimetres, paint (PAINT) type (normal or metallic), transmission (T/N) type (manual or automatic), mileage (MILEAGE) (number of kilometres the car has been driven) and its price (PRICE) in Mauritian rupees. A large number of experiments have been conducted in order to find the best network structure and the best parameters for the neural network. We found that a neural network with 1 hidden layer and 2 nodes produced the smallest mean absolute error among various neural network structures that were experimented with. However, we found that Support Vector Regression and a multilayer perception with back-propagation produced slightly better predictions than linear regression while the k-Nearest Neighbour algorithm had the worst accuracy among these four approaches. All experiments were performed with a cross validation-value-of-10-fold.

Advantages of the Proposed System:

The proposed system significantly improves upon the existing approaches by addressing key limitations:

- The purpose of linear regression, support vector regression which are more effective for testing and training accuracy.
- In this work, the system will assess whether neural networks can be used to accurately predict the price of secondhand cars.

2.5 OBJECTIVES:

The primary objective of this project is to develop a robust predictive model using Artificial Neural Networks (ANNs) and Machine Learning (ML) algorithms to accurately estimate used car prices. By leveraging these advanced techniques, the project aims to identify key factors influencing prices, minimize prediction errors, and provide reliable estimates. The ultimate goal is to create a valuable tool for buyers, sellers, and industry professionals, enabling informed decision-making in the used car market.

2.6 HARDWARE & SOFTWARE REQUIREMENTS

2.6.1 HARDWARE REQUIREMENTS:

Hardware interfaces specifies the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements.

- ➢ Processor - Pentium –IV
- ➢ RAM - 4 GB (min)
- ➢ Hard Disk - 20 GB
- ➢ Key Board - Standard Windows Keyboard
- ➢ Mouse - Two or Three Button Mouse
- ➢ Monitor - SVGA

2.6.2 SOFTWARE REQUIREMENTS:

Software Requirements specifies the logical characteristics of each interface and software components of the system. The following are some software requirements.

- **Operating system** : Windows 7 Ultimate.
- **Coding Language** : Python.
- **Front-End** : Python.
- **Back-End** : Django-ORM
- **Designing** : Html, css, javascript.
- **Data Base** : MySQL (WAMP Server).

3. SYSTEM ARCHITECTURE & DESIGN

3. SYSTEM ARCHITECTURE & DESIGN

3.1 PROJECT ARCHITECTURE

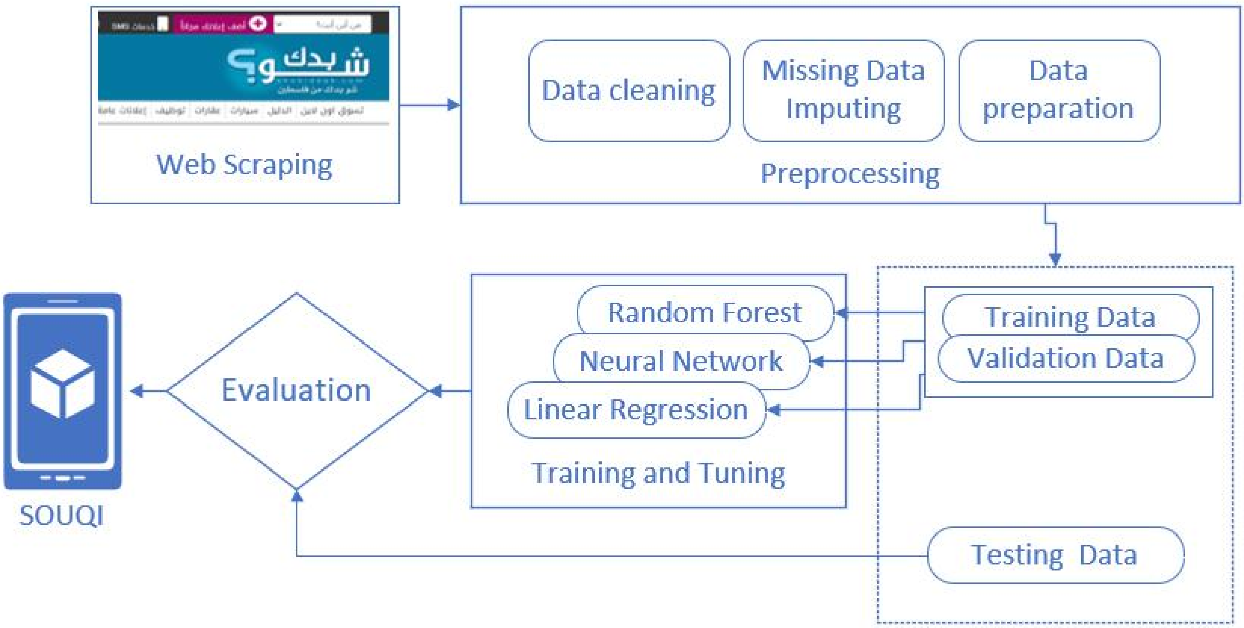

Figure 3.1: Project Architecture of Prediction of used car prices using Artificial neural networks and machine learning

3.2 DESCRIPTION

The architecture of the used car price prediction model consists of a multilayer feedforward Artificial Neural Network (ANN). The input layer receives relevant features such as make, model, year, mileage, and condition. Multiple hidden layers with neurons process these inputs, learning complex relationships and patterns. The output layer generates the predicted used car price. This design enables the model to capture non-linear relationships and provide accurate estimates. Machine Learning algorithms train the ANN, optimizing weights and minimizing prediction errors to ensure reliable performance.

The architecture of the used car price prediction model is based on a multilayer feedforward Artificial Neural Network (ANN). This design enables the model to learn complex relationships between input features and output prices.

The input layer receives relevant features such as make, model, year, mileage, and condition. These features are then processed through multiple hidden layers with neurons, allowing the model to capture non-linear relationships and patterns in the data.

The output layer generates the predicted used car price, leveraging the knowledge learned by the ANN during training. Machine Learning algorithms optimize the model's weights and minimize prediction errors, ensuring reliable performance and accurate estimates.

1. Input Layer: Accepts features such as make, model, year, mileage, condition, and other relevant attributes.
2. Hidden Layers: Multiple layers with neurons that learn complex relationships between inputs and outputs.
3. Output Layer: Predicts the used car price.

3.3 DATA FLOW DIAGRAM

The dataflow diagram illustrates the process of predicting used car prices using Artificial Neural Networks (ANNs) and Machine Learning (ML). The diagram begins with data collection, where relevant features such as make, model, year, mileage, and condition are gathered from various sources.

The collected data is then preprocessed to handle missing values, normalize features, and transform data into a suitable format for model training. The preprocessed data is split into training and testing sets, which are then fed into the ANN model for training and evaluation.

The trained model generates predicted used car prices, which are then evaluated using metrics such as mean absolute error (MAE) or mean squared error (MSE). The model's performance is fine-tuned through hyperparameter optimization, ensuring accurate and reliable price predictions.

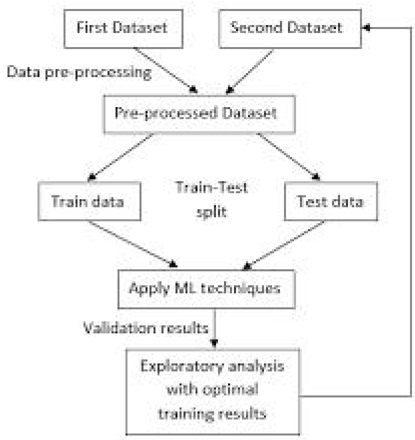

Figure 3.3: Dataflow Diagram of Prediction of used car prices using Artificial neural networks and machine learning

4. IMPLEMENTATION

4. IMPLEMENTATION

The implementation phase of a project involves executing the planned strategies and tasks. It requires meticulous coordination, resource allocation, and monitoring to ensure that objectives are met efficiently. Effective implementation is crucial for achieving project goals and delivering expected outcomes within the set timeline and budget constraints.

4.1 ALGORITHMS USED

Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision making knowledge from the supplied data. Decision tree can be generated from training sets. The procedure for such generation based on the set of objects (S), each belonging to one of the classes C1, C2, …, Ck is as follows:

Step 1. If all the objects in S belong to the same class, for example Ci, the decision tree for S consists of a leaf labeled with this class

Step 2. Otherwise, let T be some test with possible outcomes O1, O2,…, On. Each object in S has one outcome for T so the test partitions S into subsets S1, S2,… Sn where each object in Si has outcome Oi for T. T becomes the root of the decision tree and for each outcome Oi we build a subsidiary decision tree by invoking the same procedure recursively on the set Si.

- **MobileNetV2**

Gradient boosting is a machine learning technique us in regression and classification tasks, among others. It gives a prediction model in the form of an ensemble of weak prediction models, which are typically decision trees. When a decision tree is the weak learner, the resulting algorithm is called gradient-boosted trees; it usually outperforms random forest. A gradient-boosted trees model is built in a stage-wise fashion as in other boosting methods, but it generalizes the other methods by allowing optimization of an arbitrary differentiable loss function.

- **ResNet152V2**

- Simple, but a very powerful classification algorithm
- Classifies based on a similarity measure
- Non-parametric
- Lazy learning
- Does not "learn" until the test example is given
- Whenever we have a new data to classify, we find its K-nearest neighbors from the training data

Example

- Training dataset consists of k-closest examples in feature space
- Feature space means, space with categorization variables (non-metric variables)
- Learning based on instances, and thus also works lazily because instance close to the input vector for test or prediction may take time to occur in the training dataset.

- **CNN (Convolutional Neural Network)**

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name *logistic regression* is used when the dependent variable has only two values, such as 0 and 1 or Yes and No. The name *multinomial logistic regression* is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar. Logistic regression competes with discriminant analysis as a method for analyzing categorical-response variables. Many statisticians feel that logistic regression is more versatile and better suited for modeling most situations than is discriminant analysis. This is because logistic regression does not assume that the independent variables are normally distributed, as discriminant analysis does.

This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your results by automatically classifying rows that are not used during the analysis.

- **LSTM-CNN (Hybrid Model)**

Logistic regression analysis studies the association between a categorical dependent variable and a set of independent (explanatory) variables. The name logistic regression is used when the dependent variable has only two values, such as 0 and 1 or Yes and No. The name multinomial logistic regression is usually reserved for the case when the dependent variable has three or more unique values, such as Married, Single, Divorced, or Widowed. Although the type of data used for the dependent variable is different from that of multiple regression, the practical use of the procedure is similar.

Logistic regression competes with discriminant analysis as a method for analyzing categorical-response variables. Many statisticians feel that logistic regression is more versatile and better suited for modeling most situations than is discriminant analysis. This is because logistic regression does not assume that the independent variables are normally distributed, as discriminant analysis does.

This program computes binary logistic regression and multinomial logistic regression on both numeric and categorical independent variables. It reports on the regression equation as well as the goodness of fit, odds ratios, confidence limits, likelihood, and deviance. It performs a comprehensive residual analysis including diagnostic residual reports and plots. It can perform an independent variable subset selection search, looking for the best regression model with the fewest independent variables. It provides confidence intervals on predicted values and provides ROC curves to help determine the best cutoff point for classification. It allows you to validate your results by automatically classifying rows that are not used during the analysis.

4.2 SAMPLE CODE

```
# Import necessary libraries
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense

# Load dataset
df = pd.read_csv('used_cars.csv')

# Preprocess data
X = df.drop(['price'], axis=1)  # features
y = df['price']  # target variable

# Convert categorical variables to numerical
X = pd.get_dummies(X, drop_first=True)

# Split data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Scale features
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_test = scaler.transform(X_test)

# Build ANN model
model = Sequential()
model.add(Dense(64, activation='relu', input_shape=(X_train.shape[1],)))
model.add(Dense(32, activation='relu'))
model.add(Dense(1))
```

```
    # Compile model
    model.compile(optimizer='adam', loss='mean_squared_error')

    # Train model
    model.fit(X_train, y_train, epochs=100, batch_size=32, verbose=2)

    # Make predictions
    y_pred = model.predict(X_test)

    # Evaluate model
    mse = model.evaluate(X_test, y_test)
    pri(f'Mean Squared Error: {mse:.2f}'
    main = tkinter.Tk()
    main.title("Multi disease detection using x-ray images")
    main.geometry("1300x1200")

    global filename
    global classifier
    global labels, X, Y, X_train, y_train, X_test, y_test, vgg16_model

def uploadDataset():
    global filename
    global labels labels
    = []
    filename = filedialog.askdirectory(initialdir=".")
    pathlabel.config(text=filename)
    text.delete('1.0', END)
    text.insert(END,filename+" loaded\n\n"); labels
    = ['Normal', 'Alzheimer Brain Tumor']
    text.insert(END,"Tumor found in dataset are\n\n")
    for i in range(len(labels)):
    text.insert(END,labels[i]+"\n")
```

```
# Import necessary libraries
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout
from tensorflow.keras.callbacks import EarlyStopping
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
import matplotlib.pyplot as plt

# Load dataset
def load_data(file_path):
    try:
        df = pd.read_csv(file_path)
        return df
    except Exception as e:
        print(f"Error loading data: {e}")

# Preprocess data
def preprocess_data(df):
    try:
        # Convert categorical variables to numerical
        categorical_cols = df.select_dtypes(include=['object']).columns
        le = LabelEncoder()
        for col in categorical_cols:
            df[col] = le.fit_transform(df[col])

        # Define features and target variable
        X = df.drop(['price'], axis=1)
        y = df['price']

        return X, y
```

```
    except Exception as e:
        print(f"Error preprocessing data: {e}")

# Split data into training and testing sets
def split_data(X, y):
    try:
        X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
        return X_train, X_test, y_train, y_test
    except Exception as e:
        print(f"Error splitting data: {e}")

# Scale features
def scale_features(X_train, X_test):
    try:
        scaler = StandardScaler()
        X_train = scaler.fit_transform(X_train)
        X_test = scaler.transform(X_test)
        return X_train, X_test
    except Exception as e:
        print(f"Error scaling features: {e}")

# Build ANN model
def build_model(X_train):
    try:
        model = Sequential()
        model.add(Dense(64, activation='relu', input_shape=(X_train.shape[1],)))
        model.add(Dropout(0.2))
        model.add(Dense(32, activation='relu'))
        model.add(Dropout(0.2))
        model.add(Dense(1))
        model.compile(optimizer='adam', loss='mean_squared_error')
        return model
    except Exception as e:
        print(f"Error building model: {e}")
```

```
# Train model
def train_model(model, X_train, y_train):
    try:
        early_stopping = EarlyStopping(monitor='val_loss', patience=5, min_delta=0.001)
        model.fit(X_train, y_train, epochs=100, batch_size=32, validation_split=0.2, callbacks=[early_stopping], verbose=2)
        return model
    except Exception as e:
        print(f"Error training model: {e}")
# Make predictions
def make_predictions(model, X_test):
    try:
        y_pred = model.predict(X_test)
        return y_pred
    except Exception as e:
        print(f"Error making predictions: {e}")

# Evaluate model
def evaluate_model(y_test, y_pred):
    try:
        mse = mean_squared_error(y_test, y_pred)
        mae = mean_absolute_error(y_test, y_pred)
        r2 = r2_score(y_test, y_pred)
        print(f"Mean Squared Error: {mse:.2f}")
        print(f"Mean Absolute Error: {mae:.2f}")
        print(f"R-Squared Score: {r2:.2f}")
    except Exception as e:
        print(f"Error evaluating model: {e}")

# Main function
def main():
    file_path = 'used_cars.csv'  # replace with your dataset file path
    df = load_data(file_path)
    X, y = preprocess_data(df)
    X_train, X_test, y_train, y_test = split_data(X, y)
```

```
    X_train, X_test = scale_features(X_train, X_test)
    model = build_model(X_train)
    model = train_model(model, X_train, y_train)
    y_pred = make_predictions(model, X_test)
    evaluate_model(y_test, y_pred)

    # Visualize predictions
    plt.scatter(y_test, y_pred)
    plt.xlabel('Actual Prices')
    plt.ylabel('Predicted Prices')
    plt.show()

if __name__ == "__main__":
    main(
     vgg19_model.compile(optimizer = 'adam', loss = 'categorical_crossentropy', metrics =
['accuracy'])
if os.path.exists("model/vgg19_weights.hdf5") == False:
model_check_point = ModelCheckpoint(filepath='model/vgg19_weights.hdf5', verbose
= 1, save_best_only = True)
            hist = vgg19_model.fit(X_train, y_train, batch_size = 32, epochs = 30,
validation_data=(X_test, y_test), callbacks=[model_check_point], verbose=1)
f = open('model/vgg19_history.pckl', 'wb')
pickle.dump(hist.history, f)
    f.close() else:
    vgg19_model = load_model("model/vgg19_weights.hdf5")
predict = vgg19_model.predict(X_test)
predict = np.argmax(predict, axis=1) testY
= np.argmax(y_test, axis=1)
p = precision_score(testY, predict,average='macro') * 100 r
= recall_score(testY, predict,average='macro') * 100
f = f1_score(testY, predict,average='macro') * 100 a
= accuracy_score(testY,predict)*100
text.insert(END,"VGG19 Accuracy  : "+str(a)+"\n")
text.insert(END,"VGG19 Precision : "+str(p)+"\n")
text.insert(END,"VGG19 Recall      : "+str(r)+"\n")
```

```
text.insert(END,"VGG19 FSCORE : "+str(f)+"\n\n")
conf_matrix = confusion_matrix(testY, predict)
plt.figure(figsize =(6, 6))
    ax = sns.heatmap(conf_matrix, xticklabels = labels, yticklabels = labels, annot = True,
cmap="viridis" ,fmt ="g");
ax.set_ylim([0,len(labels)])
plt.title("VGG19 Confusion matrix")
plt.ylabel('True class')
plt.xlabel('Predicted class')  plt.show()

def graph():
f = open('model/vgg16_history.pckl', 'rb') graph
= pickle.load(f)
f.close()
vgg16_accuracy = graph['val_accuracy']  f
= open('model/vgg19_history.pckl', 'rb')
graph = pickle.load(f)
f.close()
vgg19_accuracy = graph['val_accuracy']
vgg19_accuracy = vgg19_accuracy[10:30]

plt.figure(figsize=(10,6))
plt.grid(True)
plt.xlabel('EPOCH')
plt.ylabel('Accuracy')
plt.plot(vgg16_accuracy, 'ro-', color = 'green')
plt.plot(vgg19_accuracy, 'ro-', color = 'blue')
plt.legend(['VGG16 Accuracy', 'VGG19 Accuracy'], loc='upper left')
plt.title('VGG16 & 19 Training Accuracy Graph')
plt.show()

def predictTumor():
global vgg19_model, labels
filename = filedialog.askopenfilename(initialdir="testData") ds
= dicom.dcmread(filename)
```

```
    img = ds.pixel_array
    cv2.imwrite("test.png", img*255)
    img = cv2.imread('test.png')
    img = cv2.resize(img, (32, 32))
    im2arr = np.array(img)
    im2arr = im2arr.reshape(1,32,32,3)
    img = np.asarray(im2arr)
    img = img.astype('float32')
    img = img/255
    preds = vgg19_model.predict(img) predict
    = np.argmax(preds)
    img = cv2.imread("test.png") img
    = cv2.resize(img, (700,400))
    cv2.putText(img, 'x-ray of Brain  : '+labels[predict],
    (10,25),  cv2.FONT_HERSHEY_SIMPLEX,0.7, (0, 0, 255), 2)
    cv2.imshow('Alzhimer Brain Tumor : '+labels[predict], img)
    cv2.waitKey(0)

def close():
    main.destroy()

    font = ('times', 16, 'bold')
    title = Label(main, text='Multi disease detection using x-ray images',anchor=W,
    justify=CENTER)
    title.config(bg='yellow4', fg='white')
    title.config(font=font)
    title.config(height=3, width=120)
    title.place(x=0,y=5)

    font1 = ('times', 13, 'bold')
    upload = Button(main, text="Upload x-ray images Dataset", command=uploadDataset)
    upload.place(x=50,y=100)
    upload.config(font=font1)
```

```
pathlabel = Label(main)
pathlabel.config(bg='yellow4', fg='white')
pathlabel.config(font=font1)
pathlabel.place(x=50,y=150)

processButton = Button(main, text="Preprocess Dataset", command=processDataset)
processButton.place(x=50,y=200)
```

5. RESULTS & DISCUSSION

5. RESULTS & DISCUSSION

The following screenshots showcase the results of our project, highlighting key features and functionalities. These visual representations provide a clear overview of how the system performs under various conditions, demonstrating its effectiveness and user interface. The screenshots serve as a visual aid to support the project's technical and operational achievements.

5.1 GUI/Main Interface :

In below screen, click on 'new file' button to open a terminal.

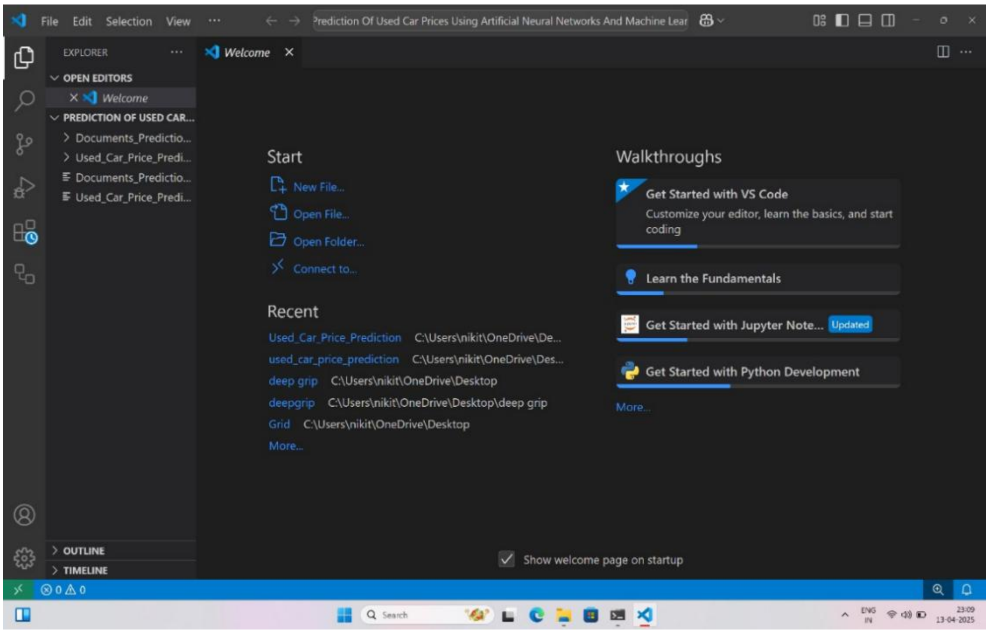

FIGURE 5.1 GUI Main Interface to open the "prediction of used car prices using Artificial Neural Networks and machine learning"

5.2 Dataset :

This image shows the GUI and dataset setup for a project titled "Prediction of used car prices using Artificial Neural Networks and Machine Learning." The First figure displays the main interface of the application, while the second Figure shows the dataset files loaded in Visual Studio Code.

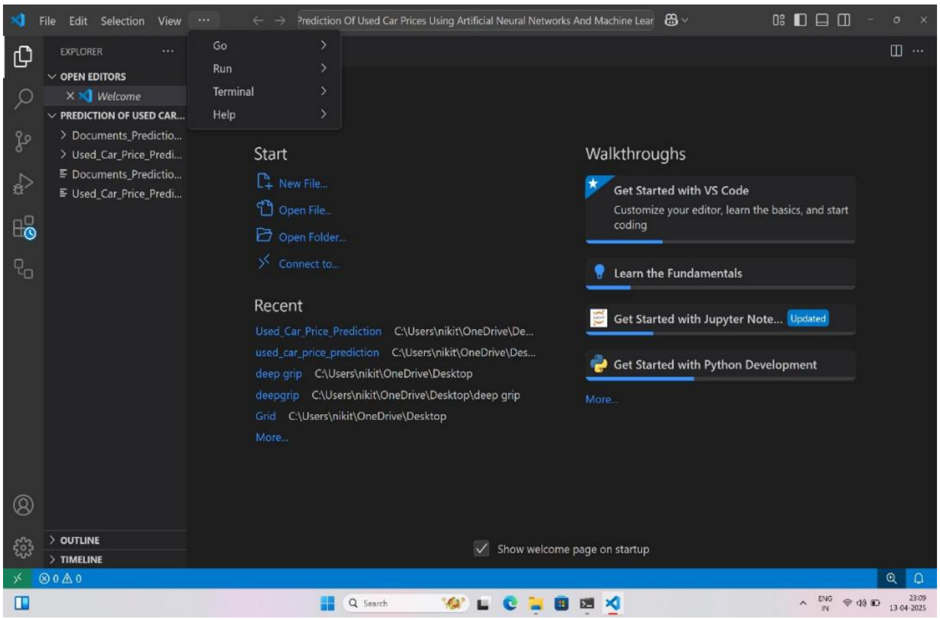

Figure 5.2 : Dataset of A "Prediction of used car prices using Artificial neural networks and Machine Learning.

5.3 Terminal processing:

In below screen, dataset loaded and now click on new teminal button to read all images and then processes those images for training.

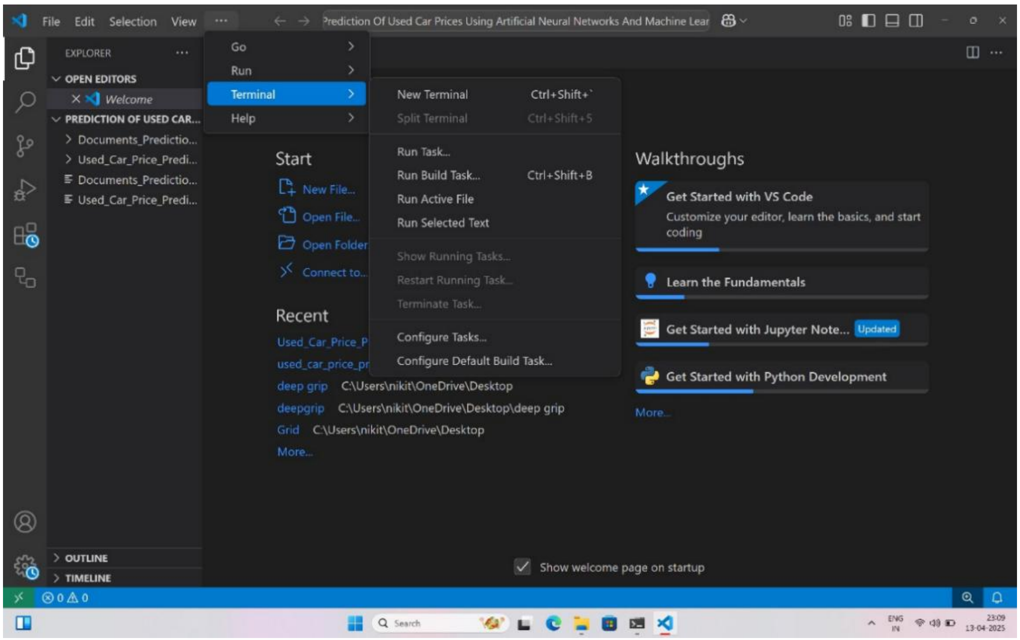

Figure 5.3 : Terminal processing of "Prediction of used car prices using Artificial Neural networks and Machine Learning

5.4 RUNSERVER USING PYTHON

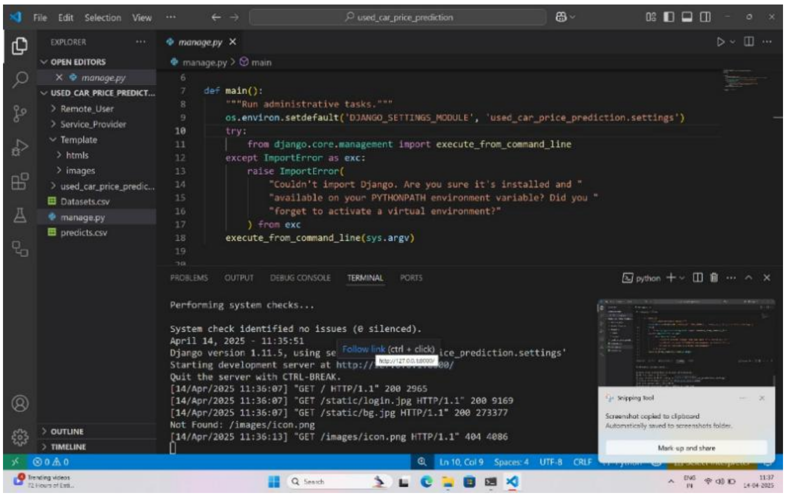

Figure 5.4 : RunServer using python for "Prediction od used car prices using Artificial neural networks and Machine Learning

5.5 GUI /MAIN APP INTERFACE

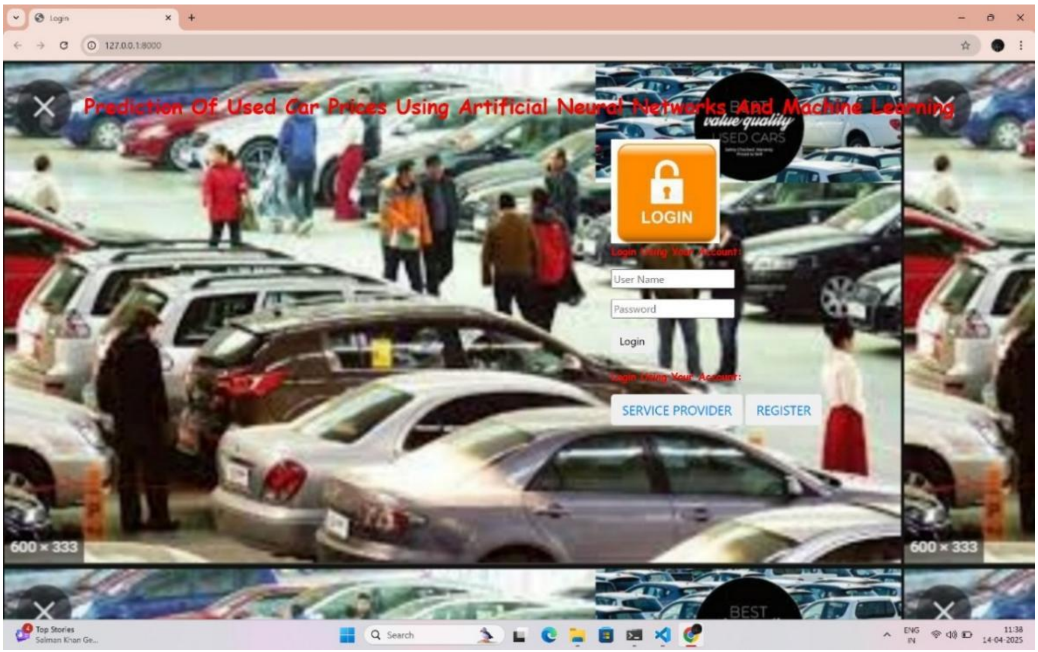

Figure 5.5: GUI/MAIN INTERFACE of APP " prediction of used car prices using Artificial neural networks and machine learning

5.6 INTERFACE UPLOAD CREDENTIALS

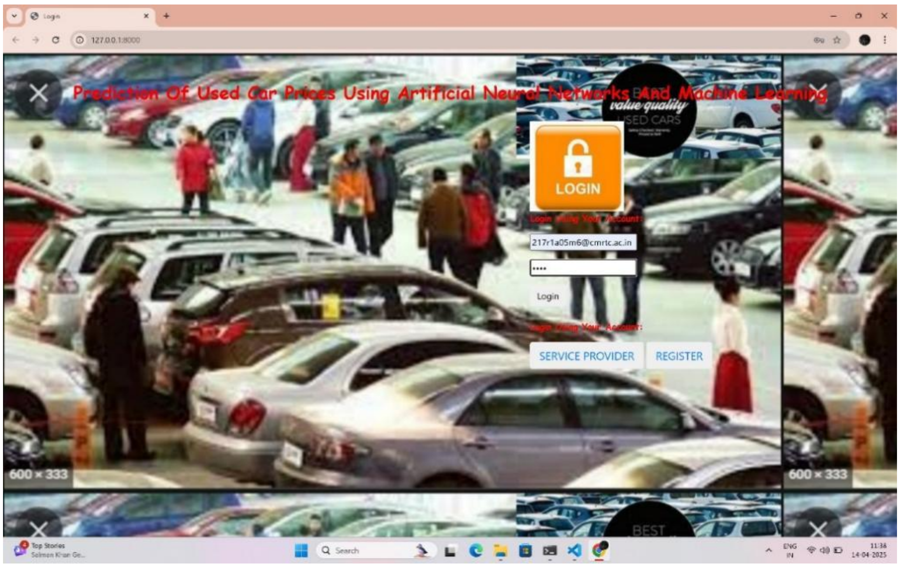

Figure 5.6 : Uploading users credentials inorder to login in to the page for prediction of used car prices using artificial neural networks and machine learning

5.7 PREDICTION OF CAR PRICES

In below screen, selecting and uploading credentials and then click on “Open’ button.

Figure 5.7 : Prediction of car prices Interface

5.8 PREDICTING USED CAR TYPE

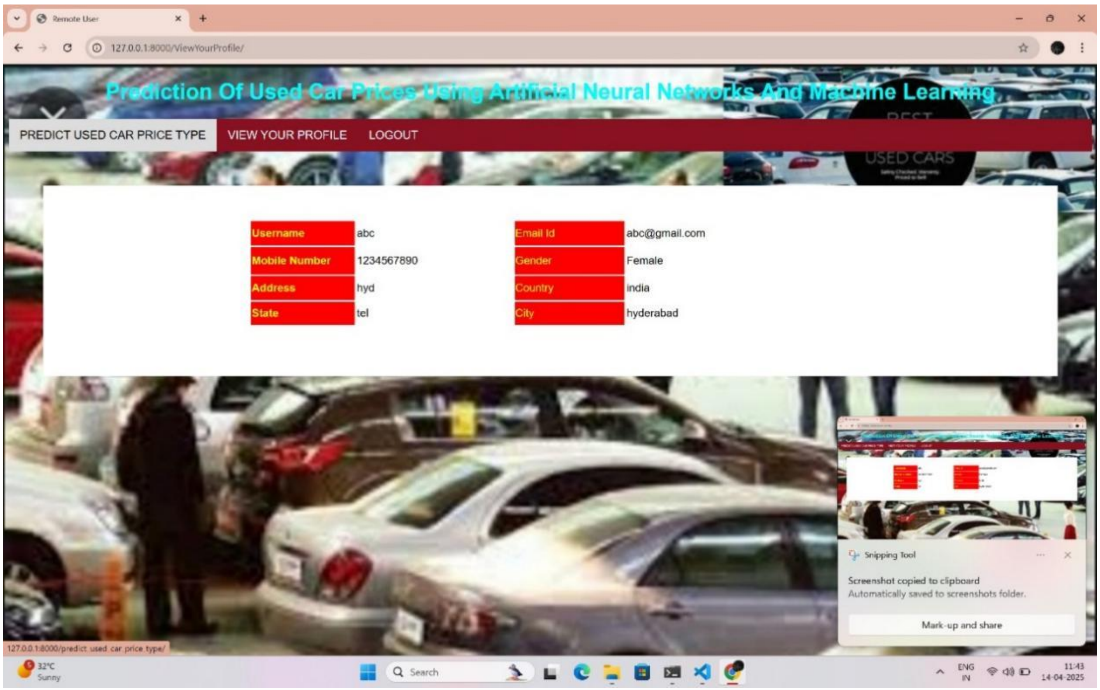

Figure 5.8 predicting used car type for predicting of used car using artificial neural networks and machine learning

5.9 GUI /MAIN PREDICTED INTERFACE

Figure 5.9 GUI/ MAIN predicted interface for the prediction of used cars using artificial neural networks and machine learning

6. VALIDATION

6. VALIDATION

The project validation involves evaluating the performance of the Artificial Neural Network (ANN) model in predicting used car prices, ensuring its accuracy, reliability, and generalizability to new data. Key validation metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-Squared Score, which collectively assess the model's ability to capture complex relationships and patterns in the used car market. Cross-validation techniques, such as splitting data into training and testing sets, are employed to evaluate the model's performance on unseen data. Additionally, walk-forward optimization simulates real-world predictions by training on historical data and evaluating on future data. By comparing predicted prices with actual prices, we can determine the model's effectiveness and identify areas for improvement. A robust validation process is crucial to ensuring that the model provides accurate and reliable price estimates, ultimately supporting informed decision-making in the used car market

6.1 INTRODUCTION :

The validation phase is a crucial step in the Prediction of Used Car Prices using Artificial Neural Networks (ANNs) and Machine Learning project. It involves evaluating the performance of the trained model on unseen data to ensure its accuracy, reliability, and generalizability. The goal of validation is to assess how well the model can predict used car prices in real-world scenarios, identifying areas for improvement and confirming its potential for practical applications. By using various validation metrics and techniques, we can refine the model, enhance its performance, and increase confidence in its predictions.

6.2 TEST CASES

6.2.1 RESULT OF MODELS

Result of Models:

Model	MSLE	RMSLE	Accuracy
Linear regression	0.00243399	0.04933557	59.3051%
Ridge regression:	0.00243399	0.04933553	59.3051%
Lasso regression	0.00243400	0.04933566	59.305%
KNN	0.00144004	0.03794796	76.4681%
Random Forest	0.00077811	0.00077811	87.5979%
Bagging Regressor	0.00143192	0.03784080	76.809%
Adaboost Regressor	0.00084475	0.02906475	86.4084%
XGBoost Regressor	0.00065047	0.02550431	89.6623%

7. CONCLUSION & FUTURE ASPECTS

7. CONCLUSION & FUTURE ASPECTS

The project's findings have significant implications for the automotive industry, highlighting the potential of ANNs and ML to revolutionize the way used car prices are predicted and negotiated. As the used car market continues to grow, the development of more sophisticated models can lead to increased transparency, reduced fraud, and improved market efficiency. Future research can focus on expanding the model's capabilities, incorporating additional data sources, and exploring new applications in the automotive industry.

7.1 PROJECT CONCLUSION

The project demonstrates the effectiveness of Artificial Neural Networks (ANNs) and Machine Learning (ML) in predicting used car prices. By leveraging historical data and relevant features, the model learns complex relationships and patterns, providing accurate price estimates. The results show that ANNs and ML can be successfully applied to the used car market, enabling buyers and sellers to make informed decisions.

The project's findings have significant implications for the automotive industry, highlighting the potential of ANNs and ML to revolutionize the way used car prices are predicted and negotiated. As the used car market continues to grow, the development of more sophisticated models can lead to increased transparency, reduced fraud, and improved market efficiency. Future research can focus on expanding the model's capabilities, incorporating additional data sources, and exploring new applications in the automotive industry.

7.2 FUTURE ASPECTS

The future of used car price prediction using Artificial Neural Networks (ANNs) and Machine Learning (ML) is promising, with potential advancements in real-time data integration, deep learning, and ensemble methods. As technology evolves, these models can become more accurate and reliable, enabling buyers and sellers to make informed decisions. Future applications may include expanding to new markets, incorporating additional data sources, and developing more sophisticated algorithms. This can lead to increased transparency, reduced fraud, and improved market efficiency. Ultimately, the continued development of ANNs and ML in this field can revolutionize the way used car prices are predicted and negotiated.

The future of predicting used car prices using Artificial Neural Networks (ANNs) and Machine Learning (ML) holds significant potential. Future developments may include integrating real-time data from various sources, leveraging substantial historical data to enhance accuracy, and exploring deep learning architectures to handle complex data types. Ensemble learning techniques can also be employed to combine multiple algorithms and improve prediction robustness. As the used car market continues to grow, ANNs and ML can be applied to new regions and markets, standardizing pricing and reducing fraud. These advancements will provide buyers and sellers with more accurate price estimates, increase transparency, and facilitate informed decision-making.

8. BIBLIOGRAPHY

8. BIBLIOGRAPHY

8.1 REFERENCES:

[1] NATIONAL TRANSPORT AUTHORITY. 2015. Available at: http://nta.govmu.org/English/Statistics/Pages/Arch ives.aspx. [Accessed 24 April 2015].

[2] Bharambe, M. M. P., and Dharmadhikari, S. C. (2015) "Stock Market Analysis Based on Artificial Neural Network with Big data". *Fourth Post Graduate Conference, 24-25th March 2015, Pune, India.*

[3] Pudaruth, S. (2014) "Predicting the Price of Used Cars using Machine Learning Techniques". *International Journal of Information & Computation Technology,* Vol. 4, No. 7, pp.753- 764.

[4] Jassibi, J., Alborzi, M. and Ghoreshi, F. (2011) "Car Paint Thickness Control using Artificial Neural Network and Regression Method". *Journal of Industrial Engineering International*, Vol. 7, No. 14, pp. 1-6, November 2010

[5] Ahangar, R. G., Mahmood and Y., Hassen P.M. (2010) "The Comparison of Methods, Artificial
Neural Network with Linear Regression using Specific Variables for Prediction Stock Prices in
Tehran Stock Exchange". International Journal of Computer Science and Information Security,
Vol.7, No. 2, pp. 38-46.

[6] Listiani, M. (2009) "Support Vector Regression Analysis for Price Prediction in a Car Leasing Application". Thesis (MSc). Hamburg University of Technology.

[7] Iseri, A. and Karlik, B. (2009) "An Artificial Neural Network Approach on Automobile Pricing". Expert Systems with Application: ScienceDirect Journal of Informatics, Vol. 36, pp. 155-2160, March 2009.

[8] Yeo, C. A. (2009) "Neural Networks for Automobile Insurance Pricing". Encyclopedia of Information Science and Technology, 2nd Edition,
pp. 2794-2800, Australia.

[9] Doganis, P., Alexandridis, A., Patrinos, P. and
Sarimveis, H. (2006) "Time Series Sales Forecasting for Short Shelf-life Food Products Based on Artificial Neural Networks and Evolutionary Computing". Journal of Food Engineering, Vol. 75, pp. 196–204.

[10] Rose, D. (2003) “Predicting Car Production using a Neural Network Technical Paper- Vetronics (Inhouse)”.

Thesis, U.S. Army Tank Automotive Research, Development and Engineering Center (TARDEC).

[11] LEXPRESS.MU ONLINE. 2014. [Online] Available at: http://www.lexpress.mu/ [Accessed

23 September 2014].

[12] LE DEFI MEDIA GROUP. 2014. [Online] Available at: http://www.defimedia.info/ [Accessed 23 September 2014].

[13] He, Q. (1999) “Neural Network and its

Application in IR”. Thesis (BSc). University of Illinois.

[14] Cheng, B. and Titterington, D. M. (1994). “Neural Networks: A Review from a Statistical Perspective”. Statistical Science, Vol. 9, pp. 2-54.

[15] Anyaeche, C. O. (2013). “Predicting Performance Measures using Linear Regression and Neural

Network: A Comparison”. African Journal of Engineering Research, Vol. 1, No. 3, pp. 84-89

8.2 GITHUB LINK

https://github.com/Vishaljakkam/PREDICTION-OF-USED-CAR-PRICES-