**Main Object:**

The research focuses on developing a **machine learning-powered mobile app** for predicting the prices of used cars. The website pages use several machine learning models to generate an approximate value based on car features and advertisements.

**Dataset:**

The dataset includes over **200,000 used car records** collected from a popular Palestinian website for used car advertisements. The dataset includes 25 features, such as car model, make, year, mileage, and more.

**Methodology:**

1. **Web Scraping**: Data was scraped from a car sales website over a period of several years.
2. **Preprocessing**: Data cleaning involved removing noise and outliers, imputing missing values, and scaling features.
3. **Feature Engineering**: Some features were transformed or reduced in complexity.
4. **Model Training**: The researchers tested five machine learning algorithms:
   * Random Forest
   * Gradient Boosted Decision Trees
   * Artificial Neural Network
   * K-Nearest Neighbors
   * Support Vector Regression These models were evaluated based on regression metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

**Conclusion:**

The Random Forest model achieved the highest prediction accuracy (around **90%**), and it was integrated into the website service to help users estimate car prices. The study aims to automate data collection and model training in the future for real-time market price predictions.

Research 2

**Main Object:**

The research focuses on **predicting used car prices** using machine learning techniques. The goal is to build models that accurately estimate car prices based on various features such as age, kilometers driven, and fuel type. The system aims to aid buyers and sellers in making informed decisions.

**Dataset:**

The dataset was sourced from Kaggle, containing features such as:

* Car name
* Year of manufacture
* Selling price
* Current price
* Kilometers driven
* Fuel type (diesel, petrol, or CNG)
* Seller type (dealer or individual)
* Transmission type (automatic or manual)
* Number of previous owners

**Methodology:**

1. **Data Gathering**: The dataset was collected from Kaggle.
2. **Data Preprocessing**: Features were cleaned, renamed, and irrelevant ones removed. Exploratory Data Analysis (EDA) was performed to visualize relationships between features. One Hot Encoding was used for categorical variables, and a correlation matrix was generated.
3. **Model Building**: The dataset was split into training and testing sets (80% and 20%, respectively). Various regression algorithms were tested:
   * Linear Regression
   * Ridge and Lasso Regression
   * Bayesian Ridge Regression
   * Random Forest
   * XGBoost
   * Gradient Boosting
   * Decision Tree
4. **Evaluation**: The models were evaluated using metrics like R² score, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

**Conclusion:**

The **Decision Tree Algorithm** performed the best, achieving the highest R² score (0.9544), indicating accurate predictions. It also had the lowest MSE and RMSE values, making it the most reliable model for predicting used car prices in the study.