**ABSTRACT:**

Project1: Forest Fires

This dataset consists of various information of forest coordinates, climatic conditions at these particular coordinates also various index rates of Forest Whether Index(FWI) System, and area of forest burnt at a particular forest coordinate due to these conditions. This data is highly valuable resource to predict forest fires at a particular location in a forest.

Project2: Accent prediction

This dataset consists of MFCC coefficients of a particular accent. In speaker recognition, most of the computation originates from the likelihood computations between feature vectors of the unknown speaker and the models in the database.

Tools/Skills used:

* Python Programming
* Jupyter Notebook
* Pandas
* Numpy
* Matplotlib
* Seaborn
* Exploratory Data Analysis
* Machine Learning

**Problem Statement 1 - Introduction to the project**

Forest Fires:

The task in this project is to predict the amount of forest burnt at any particular coordinates using the features of FWI parameters.

Implementation:

Workflow

* Import data
* Train data
* EDA
* Data Processing
* Train test split
* Ml models
* Cross Validation
* Prediction
* Test data
* Result

Modelling:

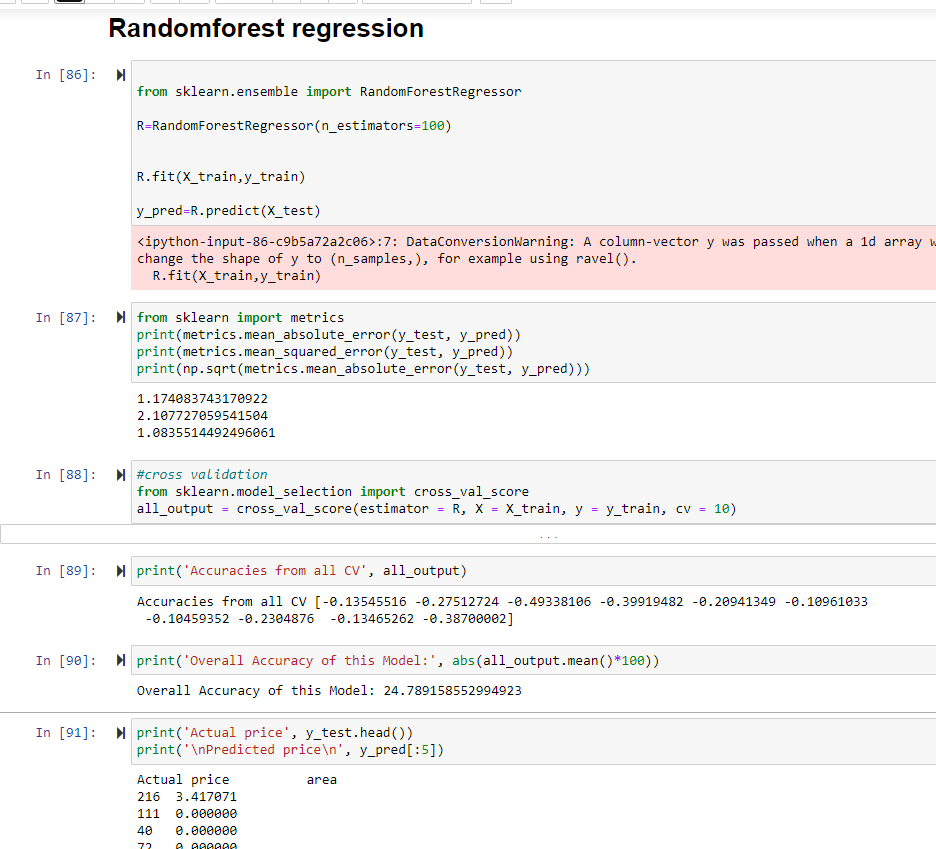
Linear Regression Model:

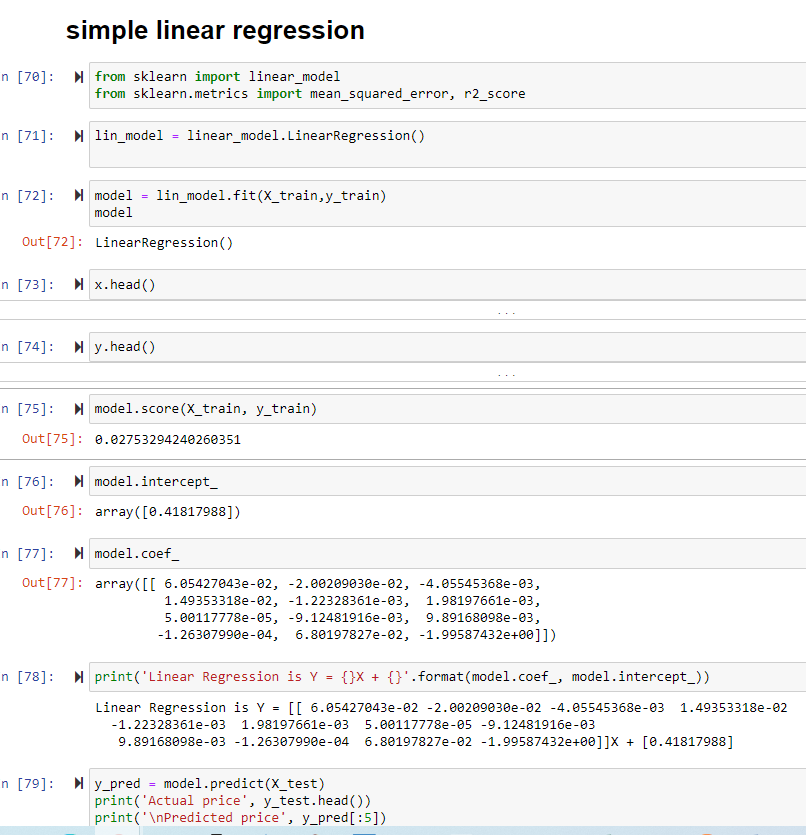
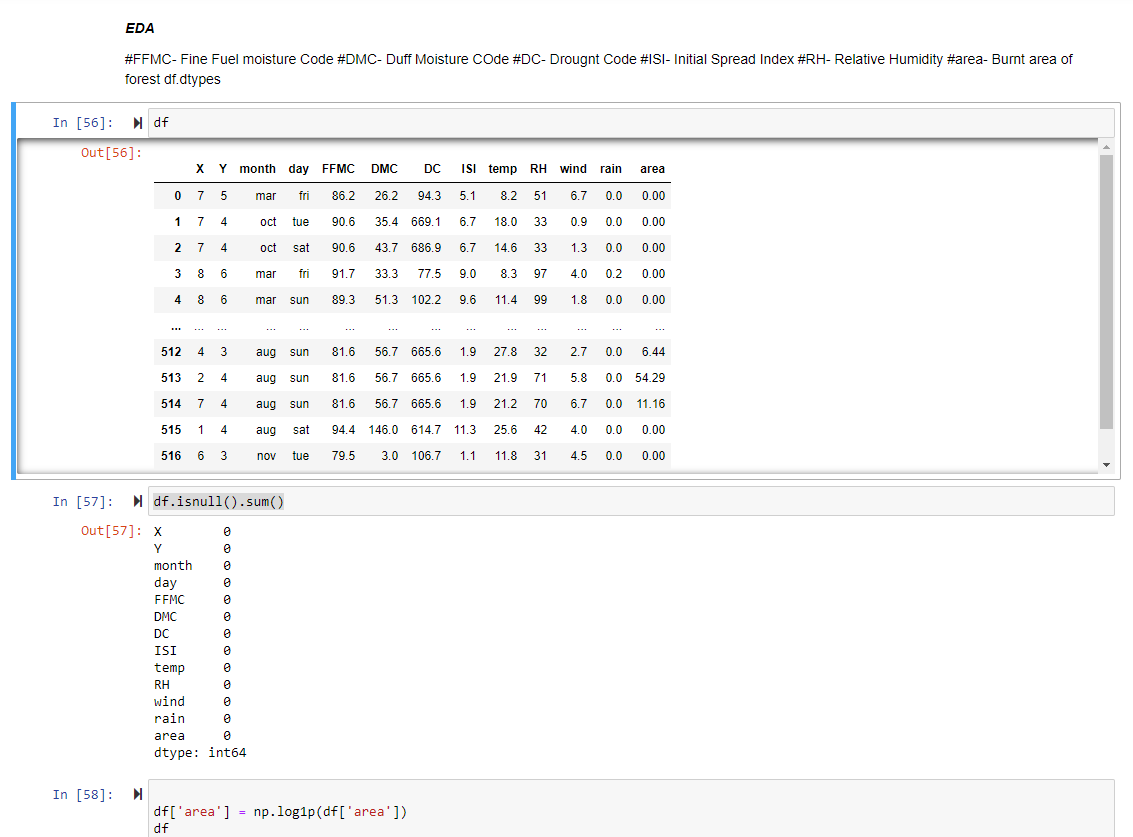
**Linear regression** is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables.

Random Forest Regression Model:

**Random forests** or **random decision forests** are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees.

Code Snippets:





Conclusion/ Results As a result we can say, Linear Regression is performing very well. Accuracy is almost 61%.

Problem Statement 2 - Introduction to the project:

Speaker’s Accent recognition:

In this project task is to predict the accent of speaker using given MFCC coefficients. This is a classification problem

Implementation:

Workflow

* Import data
* Train data
* EDA
* Data Processing
* Train test split
* Ml models
* Cross validation
* Prediction
* Test data
* Result

Modelling:

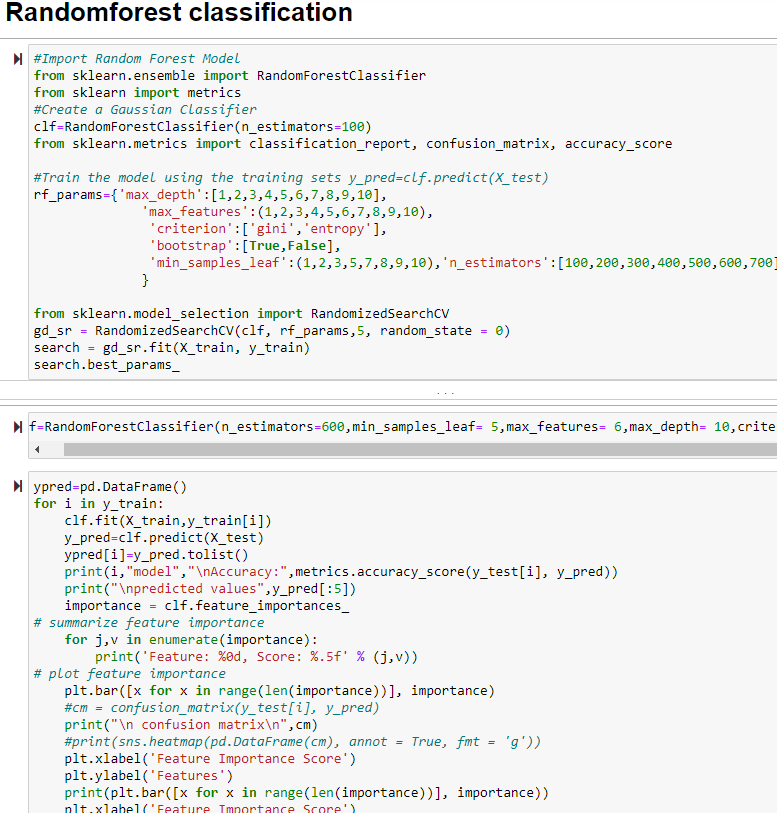
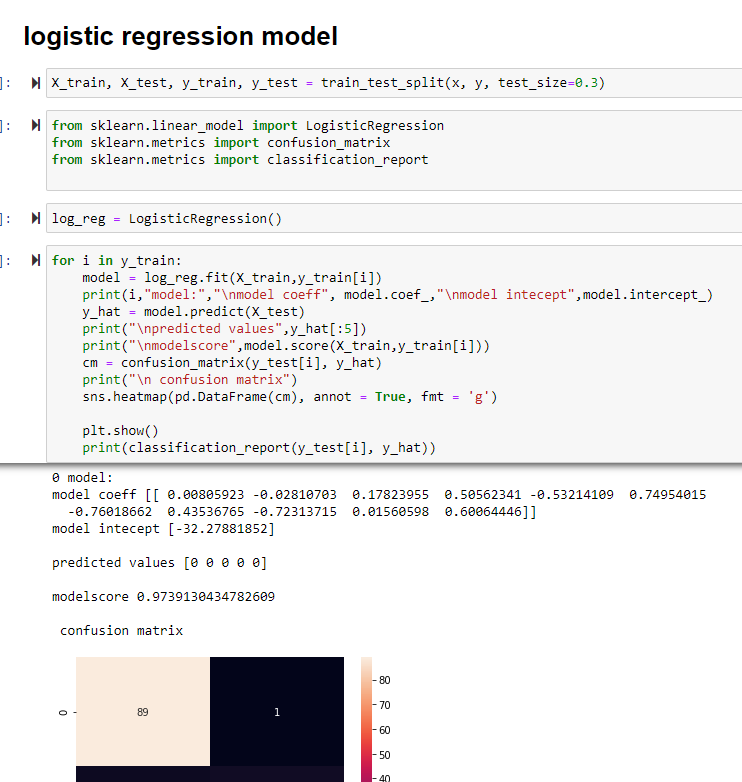
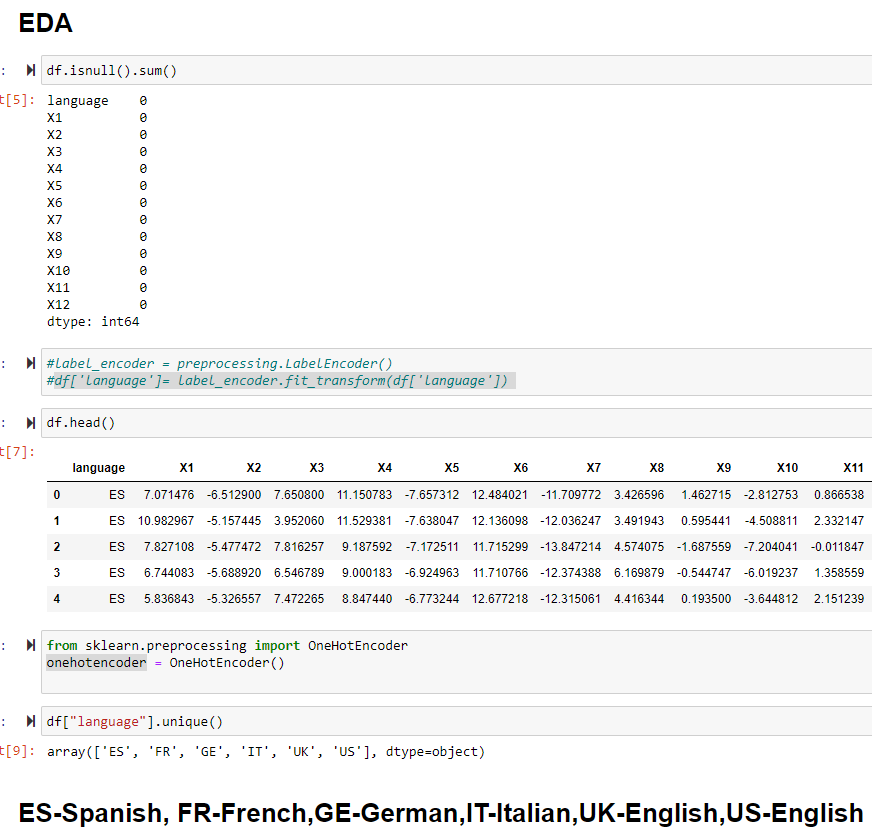
Logistic Regression:

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. It is used for both classification and regression model. In this task it is used for classification problem.

Random Forest classification Model:

**Random forests** or **random decision forests** are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean/average prediction (regression) of the individual trees. In this task we are using it for classification problem.

Code Snippets:



Conclusion/ Results As a result we can say, both Logistical Regression and random forest classifier are performing very well. Accuracy is almost 92%.