Project Charter

**Project Name:** car price pridiction

**Department:** computer science and engineering

**Focus Area:** EDA and Prediction Analysis

**Product/Process:** Data Analysis And machine learning

**Prepared By**

| **Document Owner(s)** | **Project/Organization Role** |
| --- | --- |
| Ankit Tiwari | Project Lead |
| Piyush | Team Member |
| Shikhar Garg | Team Member |
|  |  |
|  |  |

**Project Code Version Control**

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Change Description** |
| 1.0 | 16/11/22 | Ankit | * Business problem statement * Import the dataset |
| 2.0 | 16/11/22 | ankit | * Exploratory descriptive analysis |
| 3.0 | 16/11/22 | piyush | * Exploratory descriptive analysis |
| 4.0 | 17/11/22 | shikhar | * Dataset description |
| 5.0 | 17/11/22 | sikhar | * Code review |
| 6.0 | 18/11/22 | ankit | * Text data cleanup using regex * Word cloud generation |
| 7.0 | 18/11/22 | piyush | ● Exploratory descriptive analysis  ● Initial model setup |
| 8.0 | 19/11/22 | ankit | * Model local deployment |
| 9.0 | 19/11/22 | piyush | * Multiple model set up to compare and pick an optimal model |
| 10.0 | 20/11/22 | ankit | * Model hyperparameter tuning |
| 11.0 | 20/12/22 | sikhar | * Code Review |
| 12.0 | 21/12/22 | Ankit tiwari | * Web App deployment |

**TABLE OF CONTENTS**

[**PROJECT PURPOSE**](#_heading=h.39kk8xu) **4**

[**PROJECT EXECUTIVE SUMMARY**](#_heading=h.48pi1tg) **4**

[**PROJECT OVERVIEW**](#_heading=h.184mhaj) **5**

[**PROJECT SCOPE**](#_heading=h.279ka65) **5**

[4.1 Goals and Objectives](#_heading=h.meukdy) 5

[4.2 Project Deliverables](#_heading=h.243i4a2) 6

[4.3 Project Duration (Start date: 01/04/2022 - End date: 01/14/2022 )](#_heading=h.38czs75) 6

[**PROJECT CONDITIONS**](#_heading=h.415t9al) **7**

[5.1 Project Assumptions](#_heading=h.2gb3jie) 7

[5.2 Project Issues](#_heading=h.2tq9fhf) 8

[5.3 Project Constraints](#_heading=h.3rdcrjn) 8

[**PROJECT STRUCTURE APPROACH**](#_heading=h.4bewzdj) **8**

[**PROJECT TEAM ORGANIZATION PLANS**](#_heading=h.15phjt5) **8**

[**PROJECT REFERENCES**](#_heading=h.3jd0qos) **9**

[**APPROVALS**](#_heading=h.2xn8ts7) **10**

# PROJECT PURPOSE

1. This project aims to predict the Price of an used Car by taking it's Company name, it's Model name, Year of Purchase, and other parameters.

# PROJECT EXECUTIVE SUMMARY

* Goals

To sold all old car by predicting prices……………………

* Objectives
* 1. This project takes the parameters of an used car like: Company name, Model name, Year of Purchase, Fuel Type and Number of Kilometers it has been driven.
* 2. It then predicts the possible price of the car. For example, the image below shows the predicted price of our Hyundai Grand i10.
* Scope

Selling the old cars

.

# PROJECT OVERVIEW

Deciding whether a used car is worth the posted price when you see listings online can be difficult. Several factors, including mileage, make, model, year, etc. can influence the actual worth of a car. From the perspective of a seller, it is also a dilemma to price a used car appropriately. Based on existing data, the aim is to use machine learning algorithms to develop models for predicting used car prices.

# PROJECT SCOPE

## 4.1 Goals and Objectives

|  |  |
| --- | --- |
| **Goals** | **Objectives** |
| * Explore the need for hate speech detection | * Research about the various reasons for hate detection in tweets |
| * Explore various algorithms that can be used to detect hate speech | * Research about the various models that are optimal for text data |
| * Explore the dataset | * Research about the features needed for applying such models |
| * Perform descriptive analysis | * Research various ways to visualize the data |
| * Explore ways to preprocess tweets to fit in a model | * Research various ways to preprocess the data |
| * Perform predictive analysis on various models * Choose best model based on model metrics | * Research about the best model metric to be used to compare models |
| * Tune the hyperparameters of the model | * Research about the various ways to perform hyperparameter tuning |
| * Model deployment | * Research about how to deploy the Web App |

## 

## 4.2 Project Deliverables

| **Milestone** | **Deliverable** |
| --- | --- |
| 1. Explore the domain | Research all the aspects of hate speech detection and about how the experience of using the social platform can be affected by detecting and predicting hate speech |
| 2. Explore the dataset | Clear documentation on data and data visualization |
| 3. Data cleansing | Clear documentation on the data EDA and sentiment analysis |
| 4. Data preprocessing | Clear documentation on how the tweet text is altered before tokenization |
| 5. Model selection | Clear documentation on how the best fit model was chosen and tuned |
| 6. Model deployment | Web App deployment for client’s access |

## 4.3 Project Duration (Start date: 18/11/22 - End date: 25/11/2022 )

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Milestone** | **Date Estimate** | **Deliverable(s) Included** | **Confidence Level** |
| Understanding the problem statement | 16/11/22 | * Prepare task lists and notes on problem understanding | High |
| Preparing the project charter | 22/11/22 | * Goals and Objectives * Project deliverables * Deliverables out of scope * Project duration | High |
| Understanding the dataset | 16/11/22 | * Dataset inspection * Features and label identification * Perform exploratory descriptive analysis | High |
| Data preparation | 16/11/22 to  17/11/22 | * Prepare the initial dataset for modeling using feature engineering | Medium |
| Getting the best model | 20/11/22 | * Perform different modeling techniques such as Multinomial Logistic Regression, Naïve Bayes Classifier, Random Forest, Decision Tree * Compare metrics to choose the best model among all | Medium |
| Model evaluation | 22/11/22 | * Evaluate results obtained on the test data by the final model chosen * Tune the model to obtain better results | Medium |
| Model deployment | 23/11/22 | * Create a User Manual for the client to test | Medium |
| Evaluation | 24/11/22 | * Evaluation of the User Manual by the Project Manager, Stakeholders and Client | Medium |

# PROJECT CONDITIONS

## 5.1 Project Assumptions

* Spelling has to be correct in the detection of words
* Words for example @ss or C0ck would be hard to police since certain characters would be filtered out in preprocessing
* Although we found that the training data is not well rounded, an assumption could be made that the data should be well rounded unless certain processing and sampling techniques need to be deployed in association with our own modeling
* We would be specifically looking at tweets’ classification and not looking to ban users

| **#** | **Date** | **Priority** | **Owner** | **Description** | **Status & Resolution** |
| --- | --- | --- | --- | --- | --- |
| 1 | 01/11/22 | 4 | ankit | Imbalanced dataset | Logistic regression was applied on 3 kinds of the dataset (imbalanced, over-sampled, under-sampled). The dataset of the best fit model was chosen to be the updated dataset. |

## 5.3 Project Constraints

## The project constraint is to minimize false positives (wrong classify a tweet as hate speech).

## 

# PROJECT TEAM ORGANIZATION PLANS

|  |  |  |
| --- | --- | --- |
| **Project Team Role** | **Project Team Member(s)** | **Responsibilities** |
| Project Management | ankit | * Project Charter Review * Code Review * Documents Review |
| Data Handling | piyush | * Data Research * Data Understanding |
| Data Preparation / EDA | Piyush  Ankit  Shikhar garg | * Data Cleansing * Data Visualization * Data Balancing |
| Model Building | ankit | * Model Research * Model Building * Model Tuning |
| Model Deployment | piyush | * Model Deployment |
| Documentation | Ankit  shikhar | * PowerPoint Presentation * Error Document |
| ankit | * Project Charter |
| piyush | * Meeting Agenda Document * Project Progress Document * Technical Document |
| piyusg | * User Manual |

# PROJECT REFERENCES

<https://medium.com/analytics-vidhya/car-price-prediction-end-to-end-machine-learning-web-application-8e9e1fcbd8b3>

<https://www.kaggle.com/code/cagkanbay/car-price-prediction>

# APPROVALS

**Prepared by** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Manager

**Approved by** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Project Sponsor