|  |  |
| --- | --- |
| Internship Project Title | TCS ion RIO-125: Image Classification - using Different Machine Learning Algorithms |
| Name of the Company | TCS iON |
| Name of the Industry Mentor | Imran Khan |
| Name of the Institute | Mepco Schlenk Engineering College , Sivakasi |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Start Date | End Date | | Total Effort (hrs.) | | Project Environment | Tools used |
| 08.09.2021 | 07.12.2021 | | 20 | | Google colab | Python3 |
| Milestone # | 1 | Milestone: | | Cat and dogs image classifier | | |

**TABLE OF CONTENT**

* Acknowledgements
* Objective
* Introduction
* Internship Activities
* Approach / Methodology
* Outcome
* Link to code and executable file

**ACKNOWLEDGEMENTS**

I am conveying my sincere gratitude towards my industry mentor, Imran Khan, and academic mentor, Dr. S. MuthuKumar for helping me throughout this project till now and providing me this wonderful platform to complete this project. I am also thankful for answering my queries at every phase of the project. I also want to thank all my friends who helped me with valuable suggestions during this project.

**OBJECTIVE**

The objective of this model is to make a image classification using a Machine Learning Algorithms.

**INTRODUCTION**

From the first 5 days of my project, I have collected the dataset. I also cleaned and sanitized the dataset. Now the dataset is ready for training which shall be used for Image classification and I have tried a simple cat and dog classifier.

**INTERNSHIP ACTIVITIES**

* + Watched the welcome kit videos.
  + Done preparations for RIO – pre-assessment.
  + Attended the RIO – pre-assessment test.
  + Went through the day-wise plan.
  + Read the project reference material.
  + Read the industry project material.
  + Watched webinar 1.
  + Watched webinar 2.
  + Gone through all posts in the digital discussion room.
  + I went through the linear regression YouTube videos.
  + Read the linear regression article.
  + Watched the lectures provided and other videos for further understanding.
  + Created a GitHub account.
  + Searched and found out a proper data set for this project.
  + Wrote activity reports.
  + Checked and clarified the data set whether it has enough data for the project.
  + Read articles and find out how to clean and sanitize the data.
  + Cleaned the data set.
  + Sanitized the data set.
  + Done Exploratory Data Analysis(EDA)
  + Watched videos on model training
  + Used Logistic Regression and trained it
  + Used Random Forest Classifier and trained it

**APPROACH / METHODOLOGY**

The approach I took for the internship project for completing the 1st milestone is firstly understanding the concepts of the requirements. Reading articles and watching videos helped in achieving knowledge about the requirements. Google colab has been used for doing the programming. Jupyter notebook has also been used for much faster execution. A GitHub account has been created for publishing the codes.

**OUTCOME**

After the 1st milestone of this internship project, I have learned about regression models and understood how to clean and sanitize the dataset is provided.I have made a cat dog image classifier using machine learning algorithms.

I have checked all the columns using some graphs and bar plots. This gave an idea about what values were there in the corresponding columns. Then converted the categorical columns into numeric which was needed for model training. I trained two models; Logistic Regression and Random Forest Classifier. Both of their classification reports have been included in the Jupyter notebook. Among these to models, Random forest classifier showed the best score or accuracy in predicting the salary.

**LINK TO CODE AND EXECUTABLE FILE**

* Link to the code:

<https://colab.research.google.com/drive/1LMsuSYHVfZgZaT9AnZNUkPsXfPdcCGWP#scrollTo=4AyMtHw-v-SE>