|  |  |
| --- | --- |
| Internship Project Title | TCS iON RIO-125: HR Salary Dashboard - Train the Dataset and Predict Salary |
| Name of the Company | TCS iON |
| Name of the Industry Mentor | Harish Kumar |
| Name of the Institute | ICT Academy of Kerala,Thiruvananthapuram |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Start Date | End Date | | Total Effort (hrs.) | | Project Environment | Tools used |
| 05-02-2023 | 13-02-2023 | |  | | Jupyter | Python3 |
| Milestone # | 1 | Milestone: | | Clean & Sanitize salary dataset | | |

**TABLE OF CONTENT**

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

**ACKNOWLEDGEMENTS**

I am conveying my sincere gratitude towards my industry mentor, Harish Kumar 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 is to clean & sanitize the HR salary dataset.

**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 to Predict Salary.

**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.
  + Watched few of the linear regression YouTube videos.
  + Read the linear regression article.
  + Went through the linear regression & Random Forest YouTube videos.
  + Read the linear regression article.
  + Watched the lectures provided and other videos for further understanding.
  + Searched and found out a proper data set for this project.
  + Wrote activity reports.

**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. Jupyter notebook has been used for coding.

**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 checked all the columns using box plots which gave an idea about what values & outliers in the corresponding columns.

**LINK TO CODE AND EXECUTABLE FILE**

* Executable file: https://github.com/amaniyote/TCS-iON-RIO-125-HR-Salary-Dashboard---Train-the-Dataset-and-Predict-Salary-By-TCS-iON.git