Vishwak Balaji

Seeking a full-time Data Scientist role.

An entry-level Data Scientist, interested in analysing , exploring and building models and translate data points into business insights. I'm interested to apply my skills while working on new projects on different scenarios and now eager to apply the same knowledge to real-world business problems.

vishwakjaishankar@gmail.com

8610325779

linkedin.com/in/vishwak-balaji-8a384018a

github.com/Vishwakbalaji

SKILLS

- Python Programming
- Statistics
- Machine Learning
- Data Analysis
- AWS
- Natural Language Processing (NLP)
- Microsoft SQL Server
- Power BI
- Tableau

LANGUAGES

English
Full Professional Proficiency

Tamil
Native or Bilingual Proficiency

SOFT SKILLS

- Critical Thinking
- Research
- Collaboration
- Time-Management
- Communication

EDUCATION

Post Graduation Diploma Program in Data Science
Praxis Business School

02/2021 - 11/2021 — *PERCENTAGE - 67%*

B.E - Mechanical Engineering
SNS College of Technology

04/2016 - 09/2020 — *PERCENTAGE - 79.8%*

CERTIFICATES

Python Complete Course For Python Beginners | Udemy - Horizon Tech

SQL for Data Science | Coursera - University of California

Data Visualisation using Power BI | Great Learning Academy

AWS for Beginners | Great Learning Academy

PERSONAL PROJECTS

1. FLIGHT FARE PREDICTION:

- **Primary Goal:** To predict the price of the flight ticket using the below given models.
- **Solution:** To overcome the effect of imbalance training data. Sample techniques were applied. Linear and Decision Tree Regression were implemented to built models which can predict.
- **Result:** Found that *Decision Tree* performed best with Adj. R-Squared 88%.
- **Project Link :** https://github.com/Vishwakbalaji/Data_Science/tree/Notebook/Flight%20Ticket%20Price%20Prediction

2. CREDIT SCORING AND RISK PREDICTION:

- **Primary Goal:** To use the current loan application data to predict whether or not an applicant will be able to repay a loan according to the set of attributes
- **Solution :** To overcome the effect of imbalance training data. Sample techniques were applied. Logistic Regression, Decision Tree Classification were implemented to built models which can predict.
- **Result :** Found that Logistic Regression performed best with F1_Score of 82%.
- **Project Link:** https://github.com/Vishwakbalaji/Data_Science/tree/Notebook/German%20Credit%20Risk%20Prediction

3. THE HOTEL REVIEW ANALYSIS

- **Primary Goal :** Is to analyse the reviews from customer from different hotels to predict which hotel is best and gives full satisfaction to the customers.
- **Solution :** Web Scraping , Cleaning the Data, NLP pre-processing, Common Nouns extraction, Binary Conversion, Association Rules, Removing the duplicates, Compactness Pruning, Redundancy Pruning, Finding Polarity.
- **Result :** To find the Sentiment analysis for those features that we are deriving from above.
- **Project Link:** https://github.com/Vishwakbalaji/Data_Science/tree/Notebook/Cappstone_Project

4. START-UPS FUNDING ANALYSIS:

- **Primary Goal :** To analyse the Start-up company's performance based on the Industry vertical, Investors and Investment type.
- **Solution : Power BI** was used for data visualization.
- **Project Link :** https://github.com/Vishwakbalaji/Power-BI/tree/main/Start-ups%20Funding%20Analysis