

N Ashwin Siddhartha

Data Science Enthusiast

✉ ashwinsiddhartha000@gmail.com ☎ 9502092033

📍 Hyderabad, Telangana

🌐 github.com/ashwin2353?tab=repositories

in www.linkedin.com/in/ashwin-siddhartha

👤 PROFILE

Experienced data scientist with expertise in Python libraries, deep learning models, and Tableau dashboards. Passionate about problem-solving and seeking challenging opportunities to drive impactful results.

🧩 SKILLS

KEY SKILLS

- Data Visualization
- Predictive Analysis
- Statistical Modeling
- Web Scraping
- Data preprocessing
- Clustering & Classification
- Machine Learning & Deep Learning
- Model Development

TECHNICAL SKILLS

Tools:

- Python
- MySQL
- Tableau
- HTML
- CSS
- Jupyter Notebook
- Google colab
- VS Code

PACKAGES

- Scikit-Learn
- Tensorflow
- Keras
- Numpy
- Pandas
- NLTK
- OpenCV
- Seaborn
- Matplotlib
- AutoScraper

MACHINE LEARNING ALGORITHMS

- Linear Regression
- Logistic Regression
- K Nearest Neighbor
- Support Vector Machine

🧩 EXPERIENCE

Ai Variant ☑

Data Science Intern

Jul 2022 – Apr 2023

Covered the complete Data Science lifecycle concepts such as Data Collection, Data Extraction, Data Cleansing, Data Exploration, Data Transformation, Feature Engineering, Data Integration, Data Mining, building Prediction models and Data Visualization. Acquired skills like Statistical Analysis, Text Mining, Regression Modelling, Hypothesis Testing, Predictive Analytics, Predictive Modelling, Machine Learning, Deep Learning, Neural Networks, Natural Language Processing.

Pennar Industries Limited

Mechanical Maintenance Engineer

Nov 2021 – Jul 2022

Mechanical Maintenance Engineer at Pennar Industries Limited, responsible for maintaining hydraulic and mechanical presses. Planned and executed scheduled maintenance, diagnosed faults, supervised technical staff, managed budgets, and maintained records.

📁 INTERNSHIP PROJECTS

Bankruptcy-Prediction ☑

Mar 2023 – Apr 2023

Business need: This is a classification project aimed at predicting business bankruptcy using distinct features and modeling the probability of bankruptcy.

Data Science Techniques used:

- Cleaned data by removing duplicates, missing values, and outliers. Converted categorical variables to numerical and scaled features.
- Explored data using visualizations and statistical analysis to identify patterns and correlations.
- Chose appropriate ML algorithms: logistic regression, KNN, decision tree, random forest, and SVM to develop the classification model.
- Evaluated model performance using accuracy, precision, and recall metrics and selected random forest as the final model.

Conclusion:

Our classification model accurately predicts business bankruptcy based on distinct features with high evaluation metrics and has been deployed using Streamlit for easy access and probability estimation. This provides businesses with a valuable tool to assess financial risk and take appropriate measures to avoid bankruptcy.

Resume Screening and Classification ☑

Feb 2023 – Mar 2023

Business need: The document classification solution should significantly reduce the manual human effort in HRM. It should achieve a higher level of accuracy and automation with minimal human intervention.

Data Science Techniques used:

- Naive Bayes
- Decision Tree
- Random Forest
- XGBOOST

DEEP LEARNING ALGORITHMS

ANN, CNN, RNN, LSTM

NATURAL LANGUAGE PROCESSING

Word Embedding, Tf-idf, Bag-of-words, Sentiment Analysis

RECOMMENDER SYSTEMS

Collaborative, Content-based

WEB DEVELOPMENT

Streamlit, Flask



EDUCATION

Malla Reddy Engineering College

M Tech in Thermal Engineering

2017 – 2020 | Hyderabad

Grade - 7.94 CGPA

Lords Institute of Engineering And Technology

B-Tech in Mechanical Engineering

2013 – 2017 | Hyderabad

Percentage - 60%



INTERESTS

Playing Cricket | Watching Movies | Lessening Music



LANGUAGES

English | Hindi | Telugu

- Data cleaning techniques to remove unnecessary words, punctuations, URLs, stopwords, and emojis.
- Stemming and lemmatization to convert resumes into clean resumes.
- Exploratory data analysis to check for null values, duplicates, and data information.
- Utilization of multiple machine learning models, including SVC, KNN, Naive Bayes, Decision Tree, Random Forest, Adaboost, Gradient Boosting, XGBoost, LGBM Boost, and CNN.
- Overall, the project aimed to clean, analyze, and test resumes data using various data science techniques.

Conclusion:

The Random Forest model was chosen to predict resume category and deployed using Streamlit.

Oil Price Prediction

Nov 2022 – Jan 2023

Business Need: The business needs to predict the future oil prices to make informed decisions regarding buying and selling of oil.

Data science techniques used:

- Data analysis: Explored the data using various visualizations like line plot, scatter plot, histogram, box plot, etc.
- Outlier elimination: Dropped the records of years showing outliers and records with zscore>2.5.
- Data preparation: Scaled the data using MinMaxScaler.
- Model building: Forecasted the oil prices using Facebook Prophet, a time series forecasting model.

Conclusion:

In summary, the project involved analyzing and preparing the oil prices data, removing outliers, scaling the data, and building a time series forecasting model using Facebook Prophet to predict future oil prices.



CERTIFICATES

Data Science Course Completion

From ExcelR Institute

Machine Learning For Python

From IBM

Ai - Variant

Internship certificate

Food Delivery App Data Analysis

From Hicounselor

Analysing User Behaviour on Instagram using Python and SQL