

## SUMMARY

Highly motivated Data Scientist with hands-on experience in data analysis, machine learning, statistical modelling, creating and deploying APIs. Proficient in extracting actionable insights from complex datasets through data pre-processing, visualisation, and predictive modelling. A strong problem solver and communicator, good at collaborating with multidisciplinary teams, and conveying technical findings to non-technical stakeholders. Eager to apply my skills and passion for data-driven solutions to drive organizational success.

## PROFESSIONAL EXPERIENCE

### Data Scientist

Jul '22 - Present

#### Onpassive Technologies Pvt Ltd.

Hyderabad, IN

*Product Based AI Technology Company developing fully autonomous Software as a Service (SaaS) products for global clientele.*

#### Projects

- Worked on various projects like Speech translation modules, Sentence Correction and Smart Compose feature in Grammar check using state-of-the art pre-trained transformers, and Image to document conversion - Text extraction using OCR by training object detection module using YOLOv5. Currently working on EKYC feature which involves face recognition, Liveness detection, and OCR modules

#### Overview

- Close to around 15 months of experience in Data Science, Machine Learning with large datasets of Structured and Unstructured data, Data Acquisition, Data Validation, Data pipelines, Predictive Modelling, Data Visualisation, Testing and Deployment.
- Working knowledge of Python for data extraction and data manipulation, and widely used python libraries like Scikit-Learn, NumPy, Pandas, NLTK, Matplotlib, Seaborn, SciPy, statsmodels.
- Extensive Research on architecture for the above POCs and features and developed end-to-end architecture for Speech Translation and Image to Document Conversion POCs
- Good knowledge in creation and retrieval of Data from databases like SQL, MySQL and MongoDB
- Good knowledge of implementing machine learning algorithms and proficient with developing neural networks such as CNNs, RNNs, attention and Transformer Models using TensorFlow and PyTorch.
- Gained knowledge on CI/CD pipelines for efficient and automated model deployment.
- Hands-on experience in creating APIs using FlaskAPI, FastAPI.

### Data Science Intern

Jan '22 - Apr '22

#### Onpassive Technologies Pvt Ltd.

Hyderabad, IN

*Product Based AI Technology Company developing fully autonomous Software as a Service (SaaS) products for global clientele.*

#### Projects

- Email Category Classification** - This feature is used to classify an email into one of the following categories - Social, Primary, Updates and Promotions.

#### Overview

- Worked extensively on cleaning and analyzing around 3 lakh rows of email data to build a classification model to predict the category of an email
- Developed a classification model using Naive Bayes classifier to classify the category of an email
- Designing Streamlit application to demonstrate the project to stakeholders

EDUCATION

MSc. in Data Science	Oct '20 - Jul '22
GITAM University	Hyderabad, IN
CGPA - 8.89	
Key Courses - Python, Machine Learning, Statistics, SQL, NoSQL, Cloud Computing	
BSc. in Computer Science	Jun '17 - Sep '20
Bhavan's Vivekananda College	Hyderabad, IN
CGPA - 8.81	
Specialization- Mathematics, Statistics and Computer Science	

DATA SCIENCE PROJECTS

- **Master Thesis - Malicious URL Detection | Tech Stack: Python | March '22**
  - Developed a model to classify URLs as Malicious or Benign using a public dataset with 450,176 rows.
  - Extracted 17 features, including Lexical, Count-based, and Binary features.
  - Trained models using AdaBoost and Random Forest classifiers, optimizing results through the Voting Classifier ensemble method.
- **Exploratory Data Analysis - Loan Default Dataset | Tech Stack: Python | July '22**
  - Conducted comprehensive EDA on a Lending Club Dataset with 111 columns and 39,000 unique customer records.
  - Implemented advanced data cleaning techniques and analytical methods for meaningful insights.
  - Executed Univariate, Segmented Univariate, and Multivariate Analysis to uncover relationships and patterns among variables.
  - Demonstrated proficiency in Python, data preprocessing, statistical analysis, and data-driven decision-making.

KEY SKILLS

• Data Analysis	• Machine Learning	• Deep Learning	• Statistical Modelling
• Programming	• Data Visualisation	• Data Wrangling	• Databases
• Version Control	• Statistical Methods	• Data Manipulation	• Modelling Techniques
• Computer Vision	• Natural Language Processing	• Quantitative Analysis	• Debugging
• Cross functional Collaboration	• Problem solving	• Continuous Learning	• Team Coordination

TECHNICAL SKILLS

Programming Languages : Python, C, C++, Java, HTML
Python-toolboxes : NumPy, Pandas, Scikit-learn, Tensorflow, PyTorch, Keras, Matplotlib, OCR, Easy-OCR, Seaborn, Plotly, Statsmodels
Data Visualization Tools : Microsoft Excel, Power BI
Databases : SQL, NOSQL, MONGODB
IDEs : Android Studio, Anaconda, Jupyter Notebook, Google Colab
Version Control : Git - GitHub, Bitbucket
Additional : Streamlit, FastAPI, FlaskAPI, Docker, AWS S3 storage Service, CI/CD Pipeline, Latex, Microsoft Office Suite

COURSES and CERTIFICATIONS

- Advanced Certification Program- ML and DL | IIIT Bangalore | Jul '23 - Present
- Android Development —Android Oreo Developer Course| Udemy | Dec '18

ADDITIONAL ACHIEVEMENTS & VOLUNTEER EXPERIENCE

- Received second prize for my android Application - Onestop at a College-level Tech competition during my Bachelor's.
- Honored with "Mr. Computer genius 2019" award for my achievements and projects in Computer Science.
- The app which I have developed "Bhavan's Vivekananda College App" is being officially used by the college with over 2000+ users.
- Worked as Vice President of Bhavan's Core Division in an organization called Street Cause which is the largest student-run NGO.