+91 9013806373
jubinben@gmail.com
https://jubinben98.github.io/Portfolio/
https://www.linkedin.com/in/jubinben/
https://github.com/jubinben98

JUBIN BEN

SUMMARY

Data Scientist with over 3+ years of experience, skilled in predictive modeling, data processing, and data mining. Proficient in scripting languages and passionate about AI and computer vision. Expertise in building ML models and developing Python-based software.

KEY SKILLS

Machine Learning Algorithm
Deep Learning Models
Recommendation Systems
Object Detection Models
Statistical Modeling
Optimization Techniques
Data Analysis
Programming
Debugging
Model Development

TECHNICAL SKILLS

DeepLearning: ANN, CNN, RNN, LSTM
Object Detection: YOLO, RCNN, FCN, SSD
Statistics/ML: Linear/Logistic Regression, SVM, Trees, Ensemble Trees Models
Cloud & Production: AWS, Azure, Docker
Database: SQL, No-SQL, MongoDB
Languages: Python, CPP, HTML, CSS, JS
VCS & Tools: GitHub, GitLab, MlFlow, CVAT

SIDE-PROJECTS

Object Detection Model Evaluation
A user-friendly and customizable software tool for evaluating object detection models' performance using mAP and inference speed.
Tech Stacks: OpenCV, Numpy, TF

iPhone Price Prediction
This project uses Linear Regression model to predict the price drop of iPhones over time with a front-end & backend support.
Tech Stacks: Flask, Angular, Sklearn and Pandas

K-Means Clustering
This is a Python implementation of the K-means clustering algorithm that groups similar data points into a given number of clusters, providing methods to fit, predict, and return cluster centroids.
Tech Stacks: Python, Numpy, Matplotlib.

Altitude Mountain Sickness Analysis
Predict possibility of Altitude Mountain Sickness by analysing a person's features and develop an application to help people plan their trip to high altitude places safely.
Tech Stacks: Numpy, Pandas, Sklearn

PROFESSIONAL EXPERIENCE

NEC India

Data Scientist — Mar '20 – Present

- Enhancement of software pipeline used for Information Extraction. Utilising diverse NLP techniques, rule-based systems and models.
- Integrated different modules into the pipeline like Entity Linking, Co-reference resolution, Named Entity Recognition, CompactIE, GenIE, Entity Canonicalization etc.
- Led team in developing an automated forklift system using 6D Pose Detection using YOLO-V2 & YOLO-V3 and PnP algorithm.
- 3D Coordinates prediction using RGBD image and FCN-2x Segmentation model.
- Development of Python libraries to productize the AI workflow for the end customer along with the CICD pipelines for development efficiency and to maintain code quality.
- Deployed the model into AWS Server and created restful API using Flask.
- Developed Deeplearning models such as CNN-2D and Siamese Network to detect fiber optic cable anomalies using DAS sensor data.
- Possess expertise in managing and preprocessing complex datasets, such as vibration/intensity and frequency data, for model training.
- Developed a ML-Library and conducted its evaluation with Sklearn and Pyspark.
- Development of chatbot for client's customer care agents using Microsoft Azure and Python. The end product helped reducing query resolution time and improving the end customer's experience.
- Creation of quantized and Intermediate Representation models using Intel Open Vino.

DRDO – Government of India

Internship — Mar '19 – July '19

- Developed a website to enable researchers across India to submit research papers using submission forms using JSP, Servlets, Eclipse DB, HTML, CSS, and JavaScript.
- Created a user-friendly interface that uses the latest web technologies to ensure speed and security. It promoted academic excellence and innovation in the organization by providing a simple and efficient way for researchers to share their work.

ACHIEVEMENTS

- Developed and trained YOLO-V2 model which attained inference time of 40 FPS with 73 mAP score.
- Developed and implemented GitHub CICD pipeline for multiple modules, automating testing on pull requests to "dev" branch, improving efficiency by over 70%.
- Used multiprocessing techniques to reduce data preprocessing time by 16x.
- Recognized for developing advanced models on complex datasets, leading to 6-month on-site assignment in Japan.
- Upgraded post-processing script for 6D-pose detection, reducing processing time by 88%, improving efficiency for detecting multiple objects simultaneously with Non-maximal-suppression algorithm.
- Created CPP wrapper for 6D-pose detection Python library, allowing for CPP-based predictions.

EDUCATION

Bachelors in Computer Science

Amity University Noida, India — Aug'16 – June'20
Courses: DBMS, DS & Algorithms, Operating System, Computer Network, ML, Big Data

Certified Courses

Advanced Computer Vision	DeepLearning.AI
Custom models, layers, & loss function	DeepLearning.AI
Custom and Distributed training	DeepLearning.AI
Machine Learning	Stanford Online
Data Analytics	NPTEL