**Aniket Rawat**

+917014053591

aniketrawat97@gmail.com

[linkedin.com/in/aniketrawat97](https://www.linkedin.com/in/aniketrawat97/)

Data Scientist

Data Scientist with 2.3 years of experience of using **ML** algorithms and **Computer Vision**. Experienced in gathering, cleaning and organizing data for use by technical and non-technical personnel. Advanced understanding of statistical, algebraic and other analytical techniques with significant background in **Machine Learning** and experience using data insights to drive business growth. Used **Computer Vision** skills to successfully deliver **Automated Projects.** Highly organized, team leader , critical thinker and excellent communicator with 1.8 years of working in aminimal assistance and supervision environment.

**SKILLS**

**Machine learning** : Linear Regression , Clustering , Logistic Regression , SVM, Decision tree , Random Forest , Timeseries Data , ARIMA , Data Modelling , Python, SQL , Rest API ( FastAPI ), Modelling .

**Deep Learning** : CNN , Computer Vision , Transfer Learning , RNN , LSTM.

**Computer Vision**: Object detection , YOLOv3, Image classification , Transfer Learning , OCR , Tensor Flow

**Database** : MySQL, PostgreSQL.

**AWS** : EMR, S3 bucket , SageMaker .

**Tools & Libraries** : PyTorch , Keras , Pandas, Numpy, PySpark , Scikit-learn , OpenCV, Tableau.

**Exploratory Data Analysis (EDA)** : SAS , Pandas-Profiling , Sweetviz , D-Tale, Box-plots , violin plots , IQR, Multiple indexing plots.

**Data Workflow/MLOps** : Apache Airflow , Apache MLFlow , AWS Sagemaker.

**Data Visualization** : Tableau, Dash, Flask, Django , Plotly , Matplotlib, TNSE.

**Big Data** : PySpark(Spark) , Dask , Delta Lake , Delta Table , parquet.

**Cloud** : AWS EMR, AWS S3 bucket, AWS SageMaker, AWS Lambda.

**AutoML** : Auto-Sklearn , AutoKeras , TPOT

Quantitative Analysis , Probability , Mathematics , Debugging , Programming, Statistics.

**Soft Skills** : Critical Thinking , Research , Collaboration .

**PROFESSIONAL EXPERIENCE**

**Data Scientist Jul 2020 - Apr 2022**

Secure Meters, Udaipur

**Automated Energy Meter Reading using Computer Vision (Deep Learning)**

To create an automated system that reads electrical meter reading from meter images accurately and reduces validation costs.

* Designed and implemented a **low latency** **end-to-end** **classification model** for electricity meter consumption reading with **working accuracy of 0.95** on field test dataset using **CNN, YOLO**, and **Tensorflow** in object detection , image classification and character recognition stages which helped reduce the validation operation cost by **40%** till date**.**
* Trained a smaller version using **TFLite** to run on mobile devices with an **accuracy of 0.87** .
* Achieved a **30% faster predicting model** by preprocessing using image filters . Analysed data of **80,000+** custom meter images.
* **Deployed** the deep Learning model on **AWS Lambda** and integrated with multiple services to be consumed on **mobile devices**.
* Managed and trained a team of service staff to **annotate images with** **bounding boxes** using python with **high quality output**.

**Dynamic Consumer Indexing**

To establish and dynamically map electricity consumers with their feeding nodes and maintain a hierarchy to be used to detect theft.

* Proposed and developed a **K-means clustering model** on **time series data** for consumer indexing and energy accounting using **PySpark** which helped to reduce operating budget by **70%.**
* Developed a **scheduled web scraping** application using a dash UI to scrape weather data for **warehousing**.
* **Deployed** machine learning model with production-grade code, integrated with **ETL pipelines** and multiple services to be consumed at **web-scale**.
* Collaborated with **on-site field staff** and **project managers** to map electrical grid configuration and **generated insight** to propose a solution while keeping the **operating budget low**.

**Assisted Living - Anomalous behaviour detection**

To create a system that identifies daily routine of an occupant and detect anomalous days and deviations from the routine and safeguard the occupant from any mishaps by raising alarms. Created ML based / statistical models that feed on sensor data in the assisted living area.

* Researched and designed a **clustering model** to classify various types of days to detect days of anomalous human behaviour in an assisted living setup, achieved an **accuracy of 0.88** and **sensitivity of 0.92** . Used **TSNE** for visualization.
* Proposed , implemented and deployed a **statistical model** , analysed using **SAS** to generalize human activity and behaviour, achieving **accuracy of 0.83** on live anomaly detection, deployed on a **mobile application using FastAPI** in python.
* Developed inventive graphic **visualization**, high performance **dashboards** and **reports** to visualize activity patterns using **dash framework** and **Tableau**.
* Prepared and presented **weekly stakeholder report** which represented a snapshot **of key advancements and limitations.**

**POC – Presence detection and Pose Estimation using mmWave Radar & CNN**

A Proof of concept to detect human presence and to estimate their pose using multiple mmWave Radars & CNN (based on research papers).

**Roles and Responsibilities**

* Researched feasible techniques to detect presence and estimate pose using radar hardware.
* Analysed data generated by two point **mmWave Radar Sensors** using **FFT** and other signal processing techniques.
* Tested a **CNN** model that successfully **detected** human presence in an empty room using classification on ambient signals.
* Used pretrained RF-Pose CNN model for pose estimation with satisfactory results.

**Last Semester Trainee Jan 2020 - Jun 2020**

Secure Meters, Udaipur

**Analysis and Prediction of Daily Activity Patterns**

To research , analyse and predict the daily routine of an occupant in a house using ambient sensor data by forming and visualizing activity patterns. Created a statistical model to generalize the activities into a routine by making it a multiclass classification problem.

* Proposed and developed a **statistical model** to predict activity at a time of day with average class **F1 score of 0.78** with **5 classes.**
* Researched and Analysed **ADL (Activities of Daily Living)** of live experiments and **reported helpful insights**.
* Cleaned and processed sensor **timeseries data** , developed innovative **visualizations** ( heatmaps and gantt charts ) using **plotly** to visualize human activity patterns.
* Deployed a processing pipeline using python scheduler to automate raw data processing to be kept in a postgre SQL database.

**Sangrah - A digital library for Sanskrit manuscript**

* Assisted in developing an **android application** to scan Sanskrit Manuscript.
* Implemented edge detection, image thresholding, DPI Adjustment, and OCR (through **OpenCV** python library).

**PERSONAL PROJECTS**

**Smart Glasses - Object detection for visually impaired**

* Implemented a **YOLO CNN** architecture to detect objects in front of camera connected to **Raspberry Pi** using Python.
* Achieved **high accuracy of 92%** using **transfer learning technique** on a model pre-trained on **COCO dataset**.
* Achieved **low latency** and a **processing** **framerate of 2 frames per second** using **RTSP streaming data pipeline** between Raspberry Pi and cloud instance
* Deployed on a **GPU-based AWS instance** as cloud processor. Implemented **logics** for vehicle distance, speed and alert priority.

**Spam Classification using BERT**

* Fine-tuned pre-trained bert-base-uncased to classify whether a message is spam or not with an f1-score of 0.90 on minority class using Python.
* Deployed the model using Flask API.

**EDUCATION**

**Bachelor of Technology (B.Tech.) - Computer Science 7.27 GPA Jun 2016 - Jul 2020**

College of Technology and Engineering (MPUAT), Udaipur

**12th PCM – 90.0% (CBSE) 2016**

Central Public School, Udaipur

**10th – CGPA : 9.2/10 (CBSE) 2014**

St Paul Sr. Sec. School, Udaipur

**CERTIFICATIONS**

* **Deep Learning Specialization** **(5 courses) –** *Coursera* - DeepLearning.ai **2020**
* **Analyse Datasets and Train ML Models using AutoML** - *Coursera*  - DeepLearning.ai **2021**

**ACHIEVEMENTS AND VOLUNTEER WORK**

**Club President, Member** **Jun 2017 - Sep 2020**

*Toastmasters International*

Member of Toastmasters club for 3 years, served as President for 2 tenures, and a persistent executive committee member.

**Founder and Lead** **Sep 2018 - Apr 2020**

*College Machine Learning Club*

Started and led the machine learning club in college with over 80 members.

**Got 150/360 in JEE-MAINS and a percentile of 99.2**