

ANUJ KUMAR

DATA SCIENTIST

PROFILE

Working professional with 4+ years of experience in Data Science and AI. Strong knowledge of **Python, Machine Learning, Deep Learning, and Statistical Modelling**. Seeking a challenging position in your reputed organization that will enhance my skills and knowledge

WORK EXPERIENCE

SENIOR ENGINEER – PRODUCT DEVELOPMENT

Harman Connected Services, Bangalore — Aug 2022 – Present

Responsibilities:

- Developing a Turn Around Time (TAT) forecasting model using various time series algorithms

SENIOR DATA ANALYST

Ganit Business Solutions, Bangalore — May 2021 – Aug 2022

Responsibilities:

- Developed a job and candidates recommendation model that improve the recommendation by 10 times
- Lead a team of 4 to analyze a retail business, perform various diagnostics and prescriptive analysis
- Developed an optimization model that will generate a monthly beat plan (PJP) for salesmen and automated various reports using python

Achievements:

- Trailblazer Award – Oct 2021

SENIOR PROJECT ASSOCIATE

Indian Institute of Technology, Kanpur — Nov 2018 - May 2021

Responsibilities:

- Developed a python library that converts IEC C37 protocol to IEEE 104 so that we can directly get the data from PMU
- Developed a day-ahead and hour-ahead load prediction model. Developed flask API to deploy the model
- Developed a tableau dashboard that shows the load consumption history, number of active meters, solar power generation
- Extracted data from master's and Ph.D. applications profiles using python and developed a code for automatic information validation
- Developed image classification model using CNN, interacted with IOT developer to build an image capturing device under UV light exposure

Achievements:

- Recognition Award – June 2020

PERSONAL DETAILS

+91 6387595466
anujk949@gmail.com
Kanpur, Uttar Pradesh

github.com/anujk949
linkedin.com/in/anujk949

CORE COMPETENCIES

Artificial Intelligence | Data Science | Business Analytics

Machine Learning | Supervised & Unsupervised Learning | Deep Learning (ANN, CNN)

Regression | Classification | Time Series | NLP | Time Series

Statistical Techniques | Data Wrangling | Data Visualization

TECHNICAL SKILLS

Language

Python (NumPy, Pandas, TensorFlow 2.0+, Keras, Sci-kit Learn, Matplotlib, OR-Tools, etc.), R, C

Databases

PostgreSQL | MySQL | MongoDB

Tools

Tableau | Power BI | Postman | Git | MATLAB | MS Office (Word, Advance Excel, PowerPoint)

Cloud

AWS S3, EC2 Instance, etc.

PROJECT DETAILS

JOB AND CANDIDATE RECOMMENDER SYSTEM

FEB 21 TO
AUG 22

ROLE: DATA SCIENTIST

- Extracted data from MongoDB, preprocessed using Python (pandas, re, nltk)
- Building a **recommendation engine** using **TensorFlow trfs** also incorporating team dynamics features
- **Developed FastAPI** to deploy the recommendation model

PJP OPTIMIZATION

MAY'21 TO
DEC 22

ROLE: DATA SCIENTIST

- Preprocess the customer data (Excel sheets) and validate the geocodes through **Google Maps API** (address to geocode).
- Clustered the stores based on the geocodes using **K-means clustering**, and balanced the clusters using an optimization algorithm. Calculate the road distance using **here map API**
- Developed an **optimization model** to generate the monthly beat plan (PJP) for the given number of salesmen using Google or-tools
- Developed end-to-end fully **customizable module** that will pre-process, validate, and generate the beat plan
- Developed a model that will add new stores to the existing beat plan
- Developed an exception analytics model that compares the planned and actual traveled paths, saves the result in MongoDB
- Collaborated with a full-stack developer to deploy our exception analytics model

PROTOCOL CONVERTER

JAN'21 TO
APR'21

ROLE: PYTHON DEVELOPER AND RESEARCH

- Developed **encoder** and **decoder** for both the protocols that can read the **streaming data** and convert it into human-readable form
- Developed a module that will continuously maintain the connection with the device and convert from IEC C37 to IEEE 104 protocol

LOAD FORECASTING

NOV'18 TO
DEC'20

ROLE: DATA SCIENTIST

- Extracted data from SQL server, POSTGRES SQL, and **DARK SKY API (for weather data)**
- Analyzed, preprocess, and combined the data, treat null values and outliers
- Used **regression-based approach** and tested out various Machine learning models and compare the prediction result
- **Developed flask API** with the best model and write down a daily scheduler that will run at 3:00 am and give predictions for the next 24 hours. Stored the output in the database

EDUCATION

DUAL MASTER CERTIFICATION IN DATA SCIENCE AND AI

*Simplilearn associated with IBM —
2019-2020*

Completed 2 capstone projects and passed with distinction. Covered machine learning, deep learning, Big Data, Tableau, NLP, Python, R, etc.

B.TECH.

*Dr. APJ Abdul Kalam Technical
University, Lucknow — 2014-
2018*

Completed engineering in Electrical and Electronics Engineering with 80.68%, also qualified Gate 2019 with 91 percentiles

CERTIFICATES

Python for Data Science and Machine Learning Bootcamp –
Udemy [click here](#)

AWS Cloud Practioner
Certification