Anza G

Innovative Data Scientist/ Engineer with 7+ years of experience in data analytics and mathematical modeling. Advanced understanding of statistical, algebraic and other analytical techniques for handling data. Eager to bring expertise to growing organization in a challenging role.

Experience

Capgemini.
Senior Data Scientist   Helsinki, Finland
Jan. 2022 – Present
- **Senior Data Engineer/ Scientist:**
  - Working as a Senior Data Consultant for I&D Finland.
  - Working as a Data Engineer in the finance industry for Credit Risk Department.
  - Prototyping for graph database, big data pipelines using Scala and Spark and developing best practices for MLOps.

Keywords: Python, SAS, SQL, Spark, Snowflake, Databricks, Neo4j, Azure IOT hub, PowerBI

Edventures
AI Lead Developer   Stockholm, Sweden
June. 2020 – Mar 2024
- **Data Scientist:**
  - Development of extractive summarization algorithm using cosine distance.
  - Chatbot development using Rasa open source and development of dialogue flow.
  - Research on BERT and Hugging Face for implementation in chatbot for Q&A.

Keywords: Python, Rasa, NLTK, Beautiful Soup, BERT, Open AI.

Fliq Oy
Data Engineer   Vaasa, Finland
July. 2019 – Dec 2022
- **Data Engineer:**
  - Development of Data Pipelines for ETL process using SQL, PotgreSQL & Talend.
  - Developed BI Dashboard using Plotly Dash, Power BI, Qlik Sense andTableau.
  - Feature Engineering for creating Data Marts for different domains withDWH.
  - Integration of Dashboards in the existing Angular frontend.

Keywords: Python, Talend, SQL, PostgreSQL, Power BI, Tableau, Qlik Sense, Plotly, Angular.

University of Vaasa
Project Researcher (Part Time)   Vaasa, Finland
Feb. 2021 – Feb.2022
- **Data Scientist:**
  - Feature development and predictive analytics for atmospheric icing detection.
  - Analysis of Synthetic Aperture Radar data for ICE detection within open fields and forest.
  - Development of machine learning models and data pipelines for weather prediction.

Keywords: Python, Jupyter-Notebooks, SAR, FMI, Data Pipelining.

United Nations (International Atomic Energy Agency)
Data Scientist   Vienna, Austria
Oct. 2019 – August 2020
- **Developer in Bio Informatics:**
  - Understanding and contributing to snakemake workflows for mutation detection.
  - Developing a snakemake workflow for running several genetic mapping tools.
  - Developed R packages for visualizing the location of variants along genomes.
  - Developed a DashBoard prototype for a gene/mutation catalogue.

Keywords: Python, R, Snake Make, Bio Conda, BCF Tools, Data Pipelining, Azure, Plotly, Ggplot.

Infralutions Ltd

Hong Kong (Remotely)

Data Scientist

Aug.2019 – Jan.2020

- **Researcher in Machine Learning:**
  - Integration of Bitmex api with python notebook for real-time data streaming.
  - Feature Engineering to estimate the price of the trade before amending the order.
  - Visualization and sampling of data on different samplings.
  - Implementation of ARIMA model on the time series data.

Keywords: Python, ARIMA, Time Series Analysis, Anomaly Detection, Mathematical Modeling.

EVRY

Oslo, Norway

Data Scientist

Sep. 2017 – Jan. 2019

- **Developer in Artificial Intelligence:**
  - Artificial Intelligence in Anti Money Laundering to Reduce False Positives (Role: Data Scientist)
  - Personalized Recommendation System for bank products (Role: Data Scientist)
  - Intelligent Traffic Monitoring System using Computer Vision (Role: Project Lead)
  - Pre-study for the implementation of Data-Lake Solution (Role: Data Scientist)

Keywords: Python, R, Recommendation Systems, Graph Databases, Churn Analysis, Anomaly Detection

Education

University of Vaasa

Vaasa, Finland

Masters- Wireless Industrial Automation

Sep.2016 - 2020

National University of Computer & Emerging Sciences

Islamabad, Pakistan

Bachelors-Telecommunication Engineering

Aug.2008 –Jul.2012

Skills

Type	Skill	Level	Type	Skill	Level
Programming Languages	Python & R &Matlab	4	Operating Systems	Unix/Linux	4
	C/C++	4		Windows	4
	Angular	3	Microservices	Docker, Conda	3
Embedded Systems	Micro Controller Programming	3	Data Bases	Graph Databases (Neo4j)	3
	IoT Devices (Rasberry Pi)	3		Mango Db, PostgreSQL	3
Cloud Computing	AWS S3, EC2, Athena, Redshift, Lambda...	3	Business Intelligence Tools	Power BI	3
				Tableau	3
	Azure	3		Qlik Sense	4
Data Science Libraries	Numpy, Pandas, Scikit Learn, TensorFlow, Matplotlib, NLTK, SnakeMake, NetworkX, ggplot, seaborn, Dask,Plotly, Rshiny		Languages	English	4
				Urdu	4
				Norwegian	
				Swedish	2