Avinash Niroula

Karl-Kautsky-Weg 2D, 60439, Frankfurt am Main
+49 1577 283 1334 aniroula41@gmail.com
31.01.1997

Experience

Mar. 2022 – present **Data Engineer: Raisin Bank AG, Frankfurt am Main**

- Maintained and developed the Data Vault 2.0 model on Snowflake to deliver a quality Regulatory Reporting Pipeline for the bank. Consolidated various data sources of bank for internal stakeholders and provided insights in a Looker as a self-service data platform.
- Maintained the Infrastructure with Terraform and used Airflow for orchestration. Compute were pushed down to either Snowflake, or AWS Fargate in a serveless design.
- Implemented CI/CD using GitLab CI, JFrog for artifactory, and Harness for CD.
- Implemented monitoring and alerts to Slack for any inconsistencies in data or failure of pipelines.
- Initiated and implemented GreatExpectations tool for validating data quality.
- Initiated, designed and implemented local testing framework which allowed each developers to have an isolated local testing environment. Each developers could clone the Snowflake database, run a local Airflow environment in a docker, mock services in a docker container locally, and to test end-to-end data pipeline.

Experience (continued)

Aug. 2020 – Feb. 2022 **Data Scientist: Emma Sleep GmbH, Frankfurt am Main**

- Built and maintained ETL/ELT pipelines to consolidate data to AWS Datawarehouse from sources like E-commerce backend, Competitor's websites, ERP, Third-party API endpoints, CRM, Google Analytics, Google Trends using Infrastructure as a Code model and CI/CD.
- Created and maintained dashboards to track various KPIs in Tableau Server to deliver data transparency across the company and enforce business decision based on data-driven insights.
- Created ML and Data Engineering pipelines to forecast Return rates from customers, Warranty provision analysis, Analysis of impact of Marketing Campaigns, Competitor Tracking, Automation of NPS surveys and invites in various Online Reputation platforms like Trustpilot, Trustedshop.
- Jr. Team Lead for Data Science as of Sept. 2021.

Jun. 2019 – Aug. 2019 **Data Science Internship: Roche Diagnostics, Penzberg**
I worked with data of biopharmaceutical products. The project involved cleaning the data collected in the lab, building pipelines, setting up database, performing data analysis, building machine learning models to predict potency of the products, simulating the Mass Spectrum, and finally contributing in building of a web-application. **Toolsused:** Python (Libraries like Numpy, Pandas, Keras, SkLearn were used for Data Analysis. Django was used for setting up web-application), and MySQL (for database design and set-up).

Jun. 2016 – Jul. 2016 **Physics Internship: Deutsches Elektronen-Synchrotron (DESY), Hamburg**
I worked on simulation of physics inside a particle accelerator. I used GEANT4 toolkit in C++ for simulation and created a minimalist desktop application using Qt. MATLAB and Python scripts were used for further analysis of simulation.

Education

Sep. 2018 – Jun. 2020 **M.Sc. Data Engineering, Jacobs University Bremen, Germany**
Thesis: *Proteoform Profiling of Complex Biopharmaceuticals*
GPA: 1.4
Teaching Assistant for the courses Renewable Energy
Award: Roche Cooperative Study Program (Merit-based Full Tuition Scholarship by Roche Diagnostics GmbH)

Education (continued)

Sep. 2014 – Jun. 2017

- **B.Sc. Physics, Jacobs University Bremen, Germany**
  Thesis: *Superconductor-Insulator Transition in Doped Semiconductors*
  GPA: 1.6
  Teaching Assistant for undergraduate course Quantum Mechanics
  Awards: President's List 16/17, Full Tuition Scholarship (Merit-based)

Projects

- Wind Speed and Energy Output prediction. Semester-long project where I applied various Machine Learning Models in order to predict the wind speed using data from National Renewable Energy Laboratory.

- Semester project on Automated batch Analysis of LC-MS (Liquid Chromatography coupled with Mass Spectrometry) data of cocoa beans.
- Bremen Big Data Challenge 2019 (Top 10% result), Bremen Big Data Challenge 2020 (Top 10% result).
- SMS Digital Data Challenge 3.0. Built a Machine Learning model that detects defects in steel production (4th position).

Skills

Programming	Proficient in Python. Good knowledge in Bash. Learning Rust. Have used C, C++ for smaller projects.
Data Engineering	dbt, Airflow, PySpark, Meltano/Singer, Airbyte, Matillion, Fivetran
Infrastructure as Code	AWS SAM, Docker, Terraform, Pulumi
CI/CD	GitLab CI, JFrog, Harness, BitBucket Pipelines
Data Warehouses/ Databases	Snowflake, Redshift, MySQL, Postgres
Data Visualization	Tableau, Looker, Streamlit, Apache Superset
Data Quality	Great Expectations
Cloud Services	AWS (Kinesis, Firehose, SNS, SQS for Stream Processing. Lambda, Fargate, EC2, for compute. S3, Redshift, Athena, Glue for DataLake architecture.)
Machine Learning	AWS Sagemaker, MLflow for Infra. PyTorch, Keras, Scikit-learn for framework.
Other tools	LaTeX, Git
Languages	Proficient in English, Nepali, Hindi, and Urdu. B1 Level competency in German.
Interests	Data Science/Engineering, Machine Learning, Deep Learning, Physics

Research Publications

Journal Articles

1. Niroula, A., Rai, G., Haas, S. & Kettemann, S. (2020). Spatial bcs-bec crossover in superconducting $p-n$ junctions. *Phys. Rev. B*, *101*, 094514. https://doi.org/10.1103/PhysRevB.101.094514

2. Kononenko, O., Bohlen, S., Dale, J., Darcy, R., Dinter, M., Erbe, J.-H., Horbatiuk, T., Indorf, G., Di Lucchio, L., Goldberg, L., Gruse, J.-N., Karstensen, S., Libov, V., Ludwig, K., Martinez de la Ossa, A., Marutzky, F., Niroula, A., Osterhoff, J., Quast, M., … Weichert, S. (2016). Investigation of advanced electron bunch generation and diagnostics in the BOND laboratory at DESY. https://doi.org/10.3204/PUBDB-2016-05712

References

Prof. Dr. Stefan Kettemann
Jacobs University,
Campus Ring 1, 28759, Bremen
s.kettemann@jacobs-university.de

Dr. Alexander Büttner
Roche Diagnostics,
Nonnenwald 2, 82377 Penzberg
alexander.buettner@roche.com