

Shakib Ibna Shameem Polock

Cell: +358469386869 | Email: polock.official@gmail.com

Location: Helsinki, Finland | LinkedIn: <https://www.linkedin.com/in/polock11/>

Github: <https://github.com/polock11>



About

Experienced in Python and Machine Learning. Proficient in data processing and analysis, predictive modeling, capable of solving complex analytical data problems. Enthusiastic to build end to end analytical solutions using machine learning and AI that addresses complex business challenges. I am an active communicator who can collaborate with business stakeholders and can work closely with other teams.

Professional Experiences

Neste

Espoo, Finland

Machine Learning Engineer – Summer Trainee

May 2024 – Sept 2024

- Enhanced Allegro Deal Recap AI Tool by developing and implementing new features for verifying deal information reducing human workload.
- Fined Tuned and improved accuracy of BERT model for custom named entity recognition task.
- Successfully monitored machine learning and business process (Celonis) production pipelines, ensuring seamless operations and minimal downtime.
- Provided troubleshooting and support for machine learning training and inference pipelines, resolving technical issues to maintain uninterrupted operations.

BRAC IT Services Limited

Dhaka, Bangladesh

Data Engineer

Dec 2022 - May 2023

- Prepared dashboards and created visualizations to support management by providing meaningful insights.
- Developed, tested, and validated data processing pipelines to provide analytics ready data.
- Contributed to report automation services from a financial Teradata database.
- Executed data migration and data archiving tasks for optimized database operations.
- Performed data cleaning & validation tasks to ensure excellent data quality.
- Responded to ad-hoc queries and provided data support to management.

Data Engineer - Internship

Sep 2022 – Nov 2022

- Executed queries in database, validated data for reports.
- Performed data cleaning tasks.
- Attended business solution and technical sessions.
- Prepared technical & user documentation from legacy scripts and systems.

Skills

- Programming: Python
- AI/ML: Scikit-learn, PyTorch, Huggingface
- Database: PostgreSQL, MySQL
- Data visualization: Power BI, Seaborn, Matplotlib
- Cloud: Microsoft Azure (Data Factory, SQL DB, Synapse, Databricks)
- Data Processing: Pandas, NumPy
- Version control: Git, Github
- Orchestration: SAP Data Intelligence, Azure Data Factory
- Other tools: Docker, Jira

Courses & Certifications

- [Machine Learning Specialization by DeepLearning.AI](#)
Courses: Supervised Machine Learning: Regression and Classification, Advanced Learning Algorithms, Unsupervised Learning, Recommenders, Reinforcement Learning
- [IBM Data Engineering Foundation Specialization](#)
Courses: Intro to Data Engineering, Python for Data Science & AI, Python Project for Data Engineering, Intro to RDBMS, Databases & SQL with Python.
- [BI Foundations with SQL, ETL and Data Warehousing](#)
- [Version Control](#)
- Deep Learning (UniOulu)
- Software for Intelligent systems and AI (UniOulu)
- Towards Data Mining (UniOulu)
- Big Data Processing (UniOulu)
- Data Analysis & Visualization (OAMK)

Personal Projects

- [Web based RAG App with open source LLM](#)
 - Developed a Streamlit web app using LangChain and Ollama for PDF-based retrieval-augmented generation, enabling document embedding, vector search, and interactive querying of PDF document.
- [Custom Named Entity Recognition model using BERT](#)
 - Trained a custom named entity recognition model using BERT from huggingfaces transformer library with higher token classification accuracy.
- [Eco Mart sales campaign performance report](#)
 - Combined, processed and cleaned data of multiple files. Created dashboard to provide convey data insights super easily for non-technical individuals and upper management.
- [Formula One Statistical on Azure Databricks](#)
 - Implementing a lakehouse architecture using Azure and Databricks to collect, process, and analyze Formula One historical data, with orchestration in Azure Data Factory for incremental data loading and regular scheduling.
- [Tokyo Olympic Analytics on Azure Cloud](#)
 - Collecting data from multiple sources using Azure Data Factory, storing them in ADLS, applied transformation in Databricks using PySpark, showing in PowerBI for reporting.
- [Exploratory Data Analysis of the Zomato dataset](#)
 - Developed a visualization notebook using Python, Pandas and Seaborn. Created various visual plots for easier understanding of the data with appropriate findings.

Education

Master of Science in Business Analytics (CSE)

[University of Oulu](#)

Oulu, Finland | Completed 85 out of 120 credits | Sept 2023 - Present

Bachelor of Science in Computer Science

[BRAC University](#)

CGPA: 3.42 out of 4
Dhaka, Bangladesh | Jan 2017 - Apr 2022

Languages

- English (Full working proficiency – IELTS 7)
- Bengali (Native)

Reference

Simone Heideman (Character & Competence)

Lead, Process AI, Neste | Email: simone.heideman@neste.com