Avneet Kaur

Cloud Engineer | Novo Nordisk

Kobenhavn Ø, Capital Region

+4531833156

avneetreen@gmail.com

avneetreen.github.io

linkedin.com/in/avneetkaur97

SUMMARY

- Currently working as a Cloud Engineer, leveraging expertise in data management and analysis to support Novo Nordisk's mission of improving healthcare outcomes.
- Passionate about machine learning, data science, LLMs.
- Hold an M.Sc. in Computer Science from the University of Copenhagen, equipped with a strong foundation in machine learning, data science principles, and methodologies.

SKILLS

Programming Languages

- Python
- Typescript
- Java
- Spark, pySpark
- SQL
- Shell Scripting
- C/C++

Technical Skills

- Github Actions
- AWS Services
- CI/CD Pipelines
- Azure Dev Ops
- Azure Cloud Services
- Databricks

Databases

- MySQL
- Oracle
- SQL Server
- AWS Redshift
- MSSQL

TECHNICAL COURSEWORK

- Algorithms & Data Structures
- Design and Analysis of Algorithms
- Large Scale Data Analytics
- Machine Learning
- Data Mining
- Artificial Intelligence

EDUCATION

Sept 2019 - Jan 2022

Masters of Science, Computer Science

University of Copenhagen

Copenhagen, Denmark

Aug 2014 - Aug 2018

Bachelor of Technology, Computer Science and Engineering

Indraprastha Institute of Information Technology

Delhi, India

WORK EXPERIENCE

November 2024 - Present

Cloud Engineer

Novo Nordisk, [Research and Early Development]

- Designed and implemented platform architecture for ingesting data from contract research organizations, enhancing data integration and accessibility.
- Developed back-end applications for data model migration in the data warehouse, ensuring smooth and efficient data engineering processes.
- Collaborated with cross-functional teams to gather requirements and design data solutions that met business needs, resulting in improved operational efficiency.
- Implemented data governance practices and standards, ensuring compliance with regulatory requirements and maintaining data security.

February 2022 - Present

Data Engineer

Novo Nordisk, [Global Data and Artificial Intelligence]

- Leading the development of robust data pipelines using Python, Typescript, and AWS services, efficiently processing and preparing data for analysis. Improved data quality and accessibility, ensuring GXP compliance.
- Onboarding data onto NNEDL (Novo Nordisk Enterprise Data Lake.)
- Contributed to the development of AutoGDPR, an automated system for handling data subject requests, streamlining the process and ensuring compliance with data privacy regulations from data in NNEDL.
- Implemented several pipelines for digital platforms like WegovyCare and Abbott, SoGrow, enabling seamless data integration and analysis, for our Digital Health Partners.
- Created a Data as a Product service, providing internal stakeholders with curated data for improved decision-making and operational efficiency.

Aug 2021 - Jan 2022

M.Sc. Thesis Student - Development of Positron Emission Tomography pipeline

Neurobiology Research Unit

- Developed an efficient data pipeline for the pre-processing of Positron Emission Tomography (PET) images, improving the accuracy and reliability of subsequent analysis.
- Conducted in-depth research and experimentation to investigate and test methods for testing the robustness of PET image analysis algorithms.
- Contributed to the development of Nipype, an open-source Python library for analyzing MRI/PET data, by implementing new features and resolving issues.

July 2021 - August 2021

Summer Technology Analyst (Machine Learning & Natural Language Processing)

BlackRock, Inc.

- Developed a machine learning classifier using natural language processing techniques to classify bond reports as environmentally friendly, improving the

- Database Management Systems
- Signal and Image Processing
- Natural Language Processing
- Probability and Statistics
- Information Retrieval
- Network Science

AREA OF INTERESTS

- Machine Learning
- Natural Language Processing
- Artificial Intelligence for social good
- Data & Cloud Engineering
- Software Engineering

PUBLICATIONS

- Multidimensional Analysis of Trust in News Articles, AAAI , Paper Link
- That's Interesting, Tell Me More! Finding Descriptive Support Passages for Knowledge Graph Relationships, **ISWC, Paper Link**
- FlavorDB: a database of flavor molecules, Nucleic Acid Research, Paper Link

AWARDS & ACHIEVEMENTS

- Grace Hopper Celebration India Scholar 2018
- Best Paper Award 2018, International Semantic Web Conference
- One of the 8 out of 200 teams selected for the Rails Girls Summer of Code Program 2018
- Gave a talk at PyCon Thailand 2019: Demystifying Conversational AI with Python
- Gave a talk at PyCon Portugal 2023: on Building Dainty Dashboards
- Gave a talk at Pycon Spain 2023: On Automated Testing of data pipelines using BDD.

efficiency and accuracy of the analysis process.
- Played an active role in enhancing the in-house NLP library for analyzing bond reports, contributing to the development of more robust and effective tools for data analysis.
- Collaborated with a team of technology analysts to identify and implement innovative solutions for improving the efficiency and effectiveness of bond report analysis.
- Demonstrated strong problem-solving skills and a deep understanding of machine learning and NLP concepts, contributing to the overall success of the team's projects.

July 2022 - Jan 2021

Student Assistant (Deep Learning & Computer Vision)

Statumanu Aps

- Developed and tested a deep learning model for detecting optic disk and retinal vein pulsations
- Conducted a comprehensive literature survey to gather relevant information and insights, facilitating quick prototyping of models and ensuring alignment with industry best practices.
- Documented all development activities in compliance with Medical Device Regulation, ensuring accuracy, traceability, and adherence to regulatory standards.

NOTABLE PROJECTS

Multidimensional Analysis of Trust in Indian News Articles (NLP)

IBM Research, India

- Conducted multidimensional analysis of trust in Indian news articles, utilizing advanced natural language processing techniques to assess selectivity of facts, journalistic assessment, and topic selectivity.
- Published research findings at the Association of Advancement of Artificial Intelligence (AAAI) 2020, contributing to the field of trust analysis in news articles.
- Leveraged textual content analysis to evaluate the selectivity of facts presented in news articles, providing insights into potential biases and misrepresentations.
- Analysed journalists' perspectives of news articles to assess their tendency to misrepresent facts, shedding light on the trustworthiness of journalistic assessments.
- Examined the temporal diversity of issues covered by various media houses, offering insights into the trustworthiness of their topic selection.

Delineating field boundaries using sentinel-2 imagery (Remote Sensing & Computer Vision, Machine Learning)

DHI Gras, Copenhagen

- Implemented a deep learning model using convolutional networks to accurately delineate field boundaries from Sentinel-2 imagery, reducing human intervention and potential errors in agricultural records.
- Conducted extensive data pre-processing and augmentation to ensure the model's robustness and generalizability across different agricultural landscapes.
- Validated the model's performance through rigorous testing and evaluation, achieving a precision rate of 90% and significantly reducing the labor-intensive nature of field boundary delineation.