Anders E. Nielsen

ML/AI | Computational Biology| Statistics | AWS

I am an ambitious data professional with a background in computational biology and experience applying data technology to the domains of biotech, supply chain, and manufacturing. I have a strong focus on applying data, ML, and modern data tools to solve real-world problems. Visit my website for more details or feel free to send me an email.

Email me at: andellegaard@gmail.com

Visit my website at: anhosu.com

Find me on GitHub as: AnHosu

I am comfortable speaking and writing Danish, English, and Korean

Work Experience

Research Scientist, Computational Biology

Gubra from 2021-04

I am currently contributing to preclinical research and drug discovery at Gubra.
My responsibilities include

- Digitalisation of internal discovery processes and customer facing services, through Shiny applications and robust data-informed processes.
- Data architecture for a fast peptide synthesis and assay pipeline
- Experiment planning and optimisation

Data Scientist

Novo Nordisk 2017-09 to 2021-04

I started out in Supply Chain where I applied data and occasionally machine learning to simplify processes and understand trends while maintaining databases.
Two years ago, I helped create the Manufacturing Intelligence department that I am currently a part of. We are applying modern data technology to manufacturing problems while advocating data science and machine learning within the organisation. My responsibilities are quite diverse ranging from data engineering tasks and coming up with machine learning cases to teaching colleagues and hiring.
Here are some projects I have contributed to:

- Increasing overall equipment efficiency of an assembly line by applying a combination of Design of Experiments and Gaussian Processes Regression and Bayesian optimisation.
  Skills applied: ML, A/B testing, Python, combating the curse of dimensionality, Bayesian statistics.
- Predictive maintenance on brownfield manufacturing equipment.
  Skills used: AWS Greengrass, AWS IoT, CI/CD with Azure DevOps. Check out my guide to industrial IoT and doing inference with Tensorflow models trained in the cloud and deployed to the edge, in my GitHub.
- Digital transformation of 3 factories through fast-paced prototyping and development of data based products. One of these factories is recognised as a Digital Lighthouse Factory by the World Economic Forum. Skill s appied: Agile/Scrum, data visualisation, optimisation.
- Creating dashboards for planning in factories.
  Skills applied: Docker, Tableau, Streamlit, InfluxDB, visualisation, Python.
- We built a regulatory framework for applying machine learning to quality-critical pharma processes.
  Skills applied: GxP, models in production, model validation, model monitoring.
- Involved in hiring 15+ data professionals, personally leading 100+ interviews.

Intern

ICDK, Embassy of Denmark to South Korea 2016-08 to 2017-01

6-month internship at Innovation Centre Denmark at the Embassy of Denmark in South Korea. Collected market intelligence, facilitated Korean-Danish technology collaboration, arranged and executed large events, and had various diplomatic tasks.

Education

M.Sc. Bioinformatics and Systems Biology Engineering

Technical University of Denmark 2015-08 to 2017-08

M.Sc. at the Technical University of Denmark (DTU). Thesis on applying Deep Learning techniques to identify potential enzyme sequences in large metagenomic datasets.
Courses in advanced bioinformatics, protein structure &computational biology, systems biology, metagenomics, cell factory engineering, machine learning, etc.

B.Sc. Biotechnology Engineering

Technical University of Denmark 2012-08 to 2015-07

B.Sc. at DTU. Thesis on analyzing terabyte-scale NGS metagenomic data to discover ancient pathogens. Semester abroad at Seoul National University.

Skill Highlight

Machine Learning

I have a solid theoretical understanding of **ML** and **deep learning** methods, as well as 5+ years of experience applying them to various domains.
I have applied ML to solve complex problems, often with sparse or little data.
Experience with AWS **SageMaker** and the Python data science toolbox: Pandas, **SciKit**, Numpy, Scipy, **Tensorflow**, matplotlib, etc.

Computational Biology

I have worked with genome data and NGS reads in several projects, including my thesis. I have a particular interest in creating repeatable and automatic pipelines for analysis of sequence data.

Programming & Scripting

I have been writing **Python** for 6+ years and I am comfortable with several common data manipulation libraries.
I also write **Perl** and **VBA**, and have previous experience with **R** and **C#**.
I am comfortable working in **Linux** and Windows environment. I am **Unix**-savvy and used to work in command-line environments.

Databases & SQL

I have been developing and using databases for 3+ years. I am comfortable with multiple dialects of SQL, though mostly **MSSQL** and **Postgres**.
I have some **NoSQL** experience too.

Statistics

I have a strong theoretical understanding of **classical statistics** and experience applying **multivariate statistics**, including ML methods like **Gaussian Processes**.

Other Skills

- Metagenomics
- Biochemistry
- CI/CD with Azure DevOps
- Agile/Scrum
- Data streaming (AWS Kinesis)
- Natural language processing
- Design of experiments
- Hypothesis testing
- Git
- Containerisation (Docker)
- AWS certified
- GxP (and ML in GxP)
- IoT and sensor data
- Communication & Presentation

Reach out – I would be happy to elaborate on any of these.