# CONTACT

Helsingborg, Sweden 073 71 68 622

anas-mofleh@hotmail.com

ANAS MOFLEH

https://anasmofleh.github.io/

https://[www.linkedin.com/in/an](http://www.linkedin.com/in/an) as-mofleh

https://github.com/AnasMofleh

# SKILLS

AI/ML:

Image processing, NLP, Transformers, Cloud technologies.

Data engineering: Data modelling, Data

Warehousing, Data Pipelines, Data Quality, Business intelligence.

Software developing: Project Management, Agile Programming, Multi-Thread programming.

DATA ENGINEER WITH AI/ML BACKGROUND

A key player in a BI unit, acting the role of a Data Engineer to create and monitor pipelines, to construct data flows with high quality data and provide BI insights. Will make it easier for stakeholders to take business decisions and direct the business in the right direction. Kan even utilize the power of AI as well, to ultimately turn data to business value.

# WORK EXPERIENCE

## DATA ENGINEER

Region Skåne | 2023-06 - Present, Lund

Datalake for healthcare system that holds patients.

## DATA ENGINEER / BI DEVELOPER

Sinch AB | 2021-10 - 2023-06, Malmö

Create and manage pipeline and create and maintain dashboards in a BI unit, that help sales and financial teams with their decisions.

## MACHINE LEARNING ENGINEER

Sinch AB | 2021-05 - 2021-10, Malmö

Implement and maintain state-of-the-art transformers-based models for sentiment classification.

# EDUCATION

## COMPUTER SCIENCE AND ENGINEERING

Lund faculty of engineering | 2017 - 2021, lund

M. SC. in Artificial intelligence and Machine learning

# CERTIFICATIONS

2025-02

2024-10

2024-08

2024-07

2024-06

2024-05

FABRIC DATA ENGINEER ASSOCIATE

Microsoft

GITHUB ACTIONS

GitHub

DATABRICKS CERTIFIED DATA ENGINEER ASSOCIATE

Databricks

GITHUB FOUNDATIONS

GitHub

DATABRICKS LAKEHOUSE FUNDAMENTALS

Databricks

FABRIC ANALYTICS ENGINEER ASSOCIATE

Microsoft

# TOOLING

Excellent Knowledge: Python, Java, Scala, Git, MySQL, PostgreSQL, PySpark, Hadoop, Terraform, Azure, MS Fabric, Azure Data Factory, Databricks, Snowflake, DBT, Airflow, Luigi, AWS S3, Gradle, Linux, REST API, MATLAB, Jira,

Confluence, Excel, network protocols IP, Tableau, PowerBI.

Basic Knowledge: Snowpipe, Adruino, HTML JavaScript, React Native, AWS Lambda. Kafka.

# LANGUAGES

Arabic: Native. English: Professional Swedish: Professional

Good Knowledge: Docker, Kubernetes, GRPC, Haskell, R, MiniZinc, C#, C, C++.

2024-02

2023-11

2023-09

2023-08

2023-03

2023-03

2022-11

2022-04

2022-03

2022-03

2022-02

2022-01

2021-12

2016

AZURE DATA ENGINEER ASSOCIATE

Microsoft

AZURE DATA FUNDAMENTALS

Microsoft

AZURE AI FUNDAMENTALS

Microsoft

AZURE FUNDAMENTALS

Microsoft

HANDS ON ESSENTIALS - DATA ENGINEERING

Snowflake

HANDS ON ESSENTIALS - DATA LAKE

Snowflake

GETTING STARTED WITH AWS STORAGE

AWS Training Online

JINJA, MACROS, PACKAGES

DBT

DBT FUNDAMENTALS

DBT

SNOWFLAKE FUNDAMENTALS 4-DAY

Snowflake

HANDS ON ESSENTIALS - DATA SHARING

Snowflake

HANDS ON ESSENTIALS - DATA APPLICATIONS

Snowflake

HANDS ON ESSENTIALS - DATA WAREHOUSE

Snowflake

"TO BE A LEADER"

Bilda Sydöst, Högsby

# PROJECTS

Inhouse Dataplatform För quality dataflows (DIML):

Diml is an XML-based tool that Region Skåne is developing to manage qualitative data flows and ensure interoperability between different systems. The tool uses existing and open standards as much as possible, such as DCAT-AP, OpenAPI, and BPMN, to facilitate integration and data exchange. With the help of structural metadata, it enables:

1. Minimizing manual work by automating processes

2. Ensuring vendor- independent standards

3. Building on metadata as the foundation for all data management, which facilitates future integrations and analyses.

Metod: Agile, Scrum

Datalake for healthcare data:

Store and analyze healthcare data in Scania region from two main systems () to enable ML learning models to predict firstly. The probability of chock. secondly, the probability of leaving the hospital within 3 days.

* Because of the data sensitivity the architecture was mainly on-prem with a onetime extraction to an on-prem SQL server.
* Data were extracted and encrypted using SSIS packages generated using tools developed in house.
* Then transformed using SQL and aggregate it to a Kimball model.
* Finally, validate and ensure Data quality using Jupyter Notebook.

. Train ML models using SVM on-prem.

* Create Dashboards in pyton together with a chief physician "överläkare".

Methodology: Agile

Data Pipeline developing and maintenance:

Store and analyze telecom financial data from a different team (billing, credit control, sales, marketing, …) to generate insights on different dimensions such as region, country, product. Using ELT pipelines, the main steps were:

* Set a daily, monthly, or hourly procedures using Luigi, to store the Data in a temporary AWS S3 bucket.
* Move the data to a permanent Bucket in a parquet format, using python.
* Transform the data, merge with the master files and aggregate it, using DBT.
* Validate and ensure Data quality.
* Store the data in MySQL or Snowflake DB.
* Create Dashboards in Tableau, based on the stakeholder's requirement.

Methodology: DevOps

Migration project: MySQL -> Snowflake

Have been a part of the big migration process from MySQL to Snowflake. The scope has been to move all the tables, views, and procedures to Snowflake. The step has been:

* Review the Current MySQL database and remove redundancies to start with.
* Document all Tableau data sources that need to be updated to point to the Snowflake database.
* Review, document, and create all database artifacts that need to be recreated in Snowflake.
* Document all pipelines in the data lake that needed to be updated to point to the Snowflake bucket.
* Update Pipelines so that they also point to the new Snowflake bucket.
* Test and validate the data along with the different dimensions.
* Update Pipelines to remove MySQL egress jobs after going live with Snowflake.

Methodology: DevOps

Data Pipeline creation and maintenance:

Create and maintain CI/CD pipelines to ingest data from different source (Excel, APIs, DB, …) to the data lake. Then moving the data between the different AWS S3 buckets, using Luigi as an orchestration tool. The main layers of the ETL were:

* First layer: Validate the data and the integrity of it, then ingest and store the raw data as it is.
* Second layer: Transform the data and apply business logic to it, validate it, then move a copy to a different pond, using PySpark.
* Third layer: Egress the transformed data to a Snowflake database.
* Finally, create a Tableau DS based of the data in the DB and create dashboards based of it.

Methodology: DevOps

Language-Agnostic Sentiment Classifier for Messaging

A master thesis project done at Sinch. The project aim is to limit hate speech in modern day life, by using a pre-trained transformer-based model, to identify the sentiment of a text message, and classify as Positive, Neutral or Negative. The challenge was to produce a Language agnostic model, that works for any given language, with a high classification accuracy and low inference time.

We compared a dozen of different models with different fine-tuning techniques, to satisfy the requirements. At the end, the project was a success, and our work was pulled into production. (still to our current day).

Methodology: Agile