Junior Data Scientist /
Junior Data Analyst in Amsterdam or Berlin.

+31650522592 Jurriaanengels@hotmail.nl
https://pyprint.github.io/JurriaanEngels.github.io/
https://www.linkedin.com/in/jurriaan-engels-9ab6b3113/
https://github.com/PyPrint

Jurriaan Engels

Skills

- Python (Pandas, Tensorflow, Keras, Numpy, Matplotlib, Plotly, Dash, Seaborn, Folium, SciKit-learn, SciPy, Beautifulsoup, NetworkX)
- Jupyter Notebook, JupyterLab
- Excel
- Spark (PySpark, Spark SQL, SystemML)
- VBA, Microsoft Office
- Rstudio
- SQL (Microsoft SQL Server, MySQL, SQLite)
- HTML, CSS
- Microsoft Power BI
- Microsoft Azure (Databricks)
- Git

Projects

INTERACTIVE SPOTIFY MUSIC KNOWLEDGE DASHBOARD – Personal Project Mar 2023

- Used **Python (Plotly, Dash)** to make a highly interactive dashboard on Spotify data taken from Kaggle to gain insights in popular music from the past two decades.
- Performed ETL on the spotify dataset using **Pandas** and **SQL (SQLite)** to create three new dataframes, visualized the dataset using a number of charts with **Matplotlib, Seaborn** and **Plotly.express**, and found interesting music facts with EDA.

DATA SCIENCE SUPERMARKET PRICE APPLICATION – Personal Project Apr 2023

- By **webscraping (BeautifulSoup, requests, pandas, SQL** in **Python)** unstructured Dutch supermarket data, I obtain datasets on all supermarkets. I will use these datasets to make an application in which supermarket customers can compare prices of certain products, ingredient lists and full-on recipes.
- I will also implement ChatGPT (GPT-3.5) by OpenAI in order to assist the process of obtaining ingredient lists.
- To highlight some interesting findings, I will use **Tableau Public**. This project is still ongoing.

AMSTERDAM POWERBI AIRBNB INSIGHT PROJECT – Personal Project Mar 2023

- To gain insight into the current Airbnb market in Amsterdam, I developed an interactive **Power BI** dashboard. Customers of Airbnb can use this dashboard to make decisions on booking certain Airbnbs. To preprocess the data, I used both **Python** and **Power Query Editor.**

MACHINE LEARNING SEOUL BIKE PROJECT – Personal Project Mar 2023

- I used the publicly available Seoul Bike Sharing Demand dataset to predict the amount of bikes needed on a certain day in **Python**. After loading, exploring and preprocessing the data using **Pandas**, I use the machine learning libraries **Scikit-learn** and **Keras** to construct two machine learning models.

Work Experience

JUNIOR CONSULTANT (DATA SCIENCE TRAINEESHIP) - A.S.R. (House of Bèta) - Utrecht, The Netherlands Dec 2022 – Feb 2023

- Building sustainable relationships and trust with customer accounts through open and interactive communication while following communication procedures, guidelines and policies.
- Providing both specific content-related and general information about products, policies and services of life insurances at A.S.R.

DATA SCIENCE GRADUATE INTERN – TNO – The Hague, The Netherlands Sep 2021 – May 2022

- Researched privacy enhancing techniques in the field of graphical transaction data to generate privacy preserving data.
- Communicated with three professional data scientist supervisors from TNO, supervisor at the UvA and two professional data scientists at two big financial institutions on my findings and setting quality metrics for a possible model.
- Created a generative deep learning model using **Tensorflow** and **Keras** on a big data transactional dataset in **Microsoft Azure Databricks** using the **Apache Spark API.**

Education

MASTER OF SCIENCE IN MATHEMATICS – University of Amsterdam – Amsterdam, The Netherlands Sep 2019 – May 2022
Courses: Statistical Data Analysis, Advanced Machine Learning, Dynamic Programming and Reinforcement Learning.
BACHELOR OF SCIENCE IN MATHEMATICS – University of Amsterdam – Amsterdam, The Netherlands Sep 2016 – June 2019

Certificates

AGILE SCRUM FOUNDATION – Basic fundamentals of agile and scrum project management techniques.
MACHINE LEARNING A TO Z – Course on data preprocessing, regression, classification, clustering, RL, NLP, Deep learning, Dimensionality Reduction and Model Selection in Python and Rstudio.
IBM DATA SCIENCE PROFESSIONAL CERTIFICATE (9.4)– IBM course in which I developed hands-on skills using the tools, languages and libraries used by professional data scientists, ETL techniques, visualization tools, building and evaluating machine learning models and pipelines using real-world data and publishing a report for stakeholders.
IBM ADVANCED DATA SCIENCE (ongoing) **-** IBM course in which I am learning how to build fully scalable end-to-end data integration tools, deep learning pipelines, learning how to use Apache Spark, SparkML, Tensorflow, Keras, Pytorch, Deeplearning4j, Apache CouchCB, MQTT.