DORON BEN-CHAYIM née Snoyman

Data Scientist

Passionate Data Scientist with over 2 years of Python and SQL experience in statistical analysis, machine learning and AI solutions.

I love working in cross-dimensional teams to solve difficult and unique problems, bringing actionable insights to help drive business decisions.

CONTACT

✉ doronbenchayim@gmail.com

📞 0537914932

in https://www.linkedin.com/in/doronbenchayim

⌗ github.com/Doron-Ben-Chayim

Skills

- English (Native), Hebrew (Fluent)
- Programming Languages:
    - Python ,Sql, Matlab, Bash, Java
- Data Visualization
    - Matplotlib, seaborn, plotly
- Data Cleaning
- Data Exploration
- Data Analysis
- Signal Processing
- Feature Engineering
- Data Synthesis
- Dimensionality reduction
- Machine Learning
- Statistics
- Computer Vision
- NLP
- Data mining/Web Scraping
- Supervised Learning
- Unsupervised Learning
- Anomoly Detection
- Fluent in following modules: Pandas, Numpy, Scikit-learn, Docker, Tensorflow/keras, spark, flask, heroku, Git Linux/Windows, AWS
- SaaS
- Adaptability
- Communication
- Creative
- Time Management
- Teamwork

Relevant experience

Olive Diagnostics

Data Scientist   01/22- 12/22

- Created Python code for the entire ETL data pipeline utilized in the worlds first smart toilet.
- Utilized Python libraries such as Seaborn and Matplotlib to perform Exploratory Data Analysis (EDA) to visualize and extract vital information from the data.
- Developed signal processing algorithms used in data cleaning, enabling 50% more data extraction.
- Worked across multidisciplinary teams combining scientific and statistical techniques, feature engineering 30% of the current features.
- Designed and optimized the data processing pipeline, reducing feature creation time from 3-4 hours to 2-3 minutes.
- Built, trained and optimized powerful machine learning models increasing accuracy,recall & precision to over 90% .
- Generated realistic synthetic data utilzing AI algoirthms such as Variational-Auto Encoders (VAR), SMOTE & GANS to overcome the data scarcity.
- Deployed and maintained trained models onto the cloud using AWS.
- Designed, maintained and deployed data dashboards outlining vital data insights and statistics using Plotly, Dash, Flask and Heroku.

Cellebrite

Data Science Intern   09/21-12/21

- Developed a program that generated realistic smartphone data such as text, speech and images, for in-house product testing.
- Developed AI generative 'topic specific' chatbots, using NLP models such as GPT2 & DialoGPT.
- Generated realistic Text-to-Speech models using tacatron2 and voice cloning Neural network models.
- Designed and researched authentic Social-Network models using modules such as networkX.
- Curated and generated realistic fake faces and images using NN models such as StyleGAN.

Education Background

Masters Data Science, UNSW, Sydney (Remote)   2023- 2025
ITC - Israel Tech Challenge Data Science Program   2021
- 6 month, full-time training course in collaboration with leading tech companies, which qualifies talented STEM BSc/MSc graduates as professional Data Scientists.

Bachelors of Civil Engineering (Hons) UNSW, Sydney   2015-2019
Bachelors of Commerce (Banking and Finance) UNSW, Sydney   2015-2019