



Tomi Illes

Data Scientist



portfolio: tomilles.com



linkedin.com/in/tomilles



tomilles97@gmail.com

Technikal skills

- **Data Science:**
Python, Pandas, NumPy, SQL, regex, Matlab, R, ETL, IoT, EDA, Seaborn, Matplotlib, web scraping, Selenium
- **Big Data:**
Databricks, PySpark, DWH
- **ML:**
SKlearn, PyTorch, spaCy, SciPy, Tensorflow, Keras, CNN, NLP
- **Cloud - Azure:**
IoT Hub, Logic apps, ADX, KQL (Kusto), Synapse
- **Data modelling:**
PBI, Tableau, SSIS, SSMS
- **Web:**
HTML, CSS, jQuery

ABOUT ME

An enthusiastic data engineer with great interest in applied ML to solve new, emerging problems. I often work on various, independent projects which requires me to be able to quickly learn new technologies and apply them to real world problems. In current my role I develop, maintain and improve streaming data pipelines for End-To-End IoT solutions, and further analyse the outputs to get an insight into the data collected. Furthermore, I usually participate in various ad-hoc side projects where data science and ML skills are essential to tackle the problems in hand.

EXPERIENCE

IoT Data Engineer

IFUA Horváth, October 2022 — present



- Plan, create, manage and maintain end-to-end IoT data pipelines in Azure using IoT Hub, ADX, DataLakes, Logic Apps, Synapse
- Analyse historical data from IoT devices to identify outliers, correlations and energy saving opportunities for clients in Databricks using PySpark
- Develop ETL processes and backend data models for BI visualisation purposes

Data Engineer (freelancing)

Neuron Solutions, 2022, 6 months



- Carried out image based fault detection model inferences, EDA, and performance tests and evaluation to improve the ML model
- Using Pandas, SpaCy, NER and regex extracted information from medical text data to prepare it for ML training purposes to predict lung tumor stages

IoT developer (Junior)

IFUA Horváth, December 2021 — October 2022



- Developed IoT streaming data pipelines in Azure
- Created ETL processes in SSIS and SSMS
- Regular research on novel IoT techniques and advances

Data Science Intern

IFUA Horváth, 2017, 2 months



- Identified customer behaviours through data analysis using k-means clustering and basic data science tools in Python.

PROJECTS

NLP engineer

Hackerintro, September 2021



- Built a rule-based model and trained a custom NER model in spaCy to detect individuals' names and tech skills in CVs / resumes.

Master's final project

Durham University, 2020



- The extraction of musical instruments from bachata songs using semantic segmentation for audio source separation

Stock price prediction and classification 2021

- Developing a statistical CNN model with various features (e.g. multiscale) to predict the stock price movement of the next day of a given market based on the previous X days.

Interpersonal skills

- Communication
- Collaboration
- Teamwork
- Leadership
- Public speaking
- Creative & critical thinking
- Analytical mindset
- Teaching
- Organisational & management skills
- Report writing

Extracurricular activities & positions

- European Solidarity Corps volunteer
2020 – 2021, *Asociación Las Niñas del Tul, Granada*
- Latin dance instructor
2016 – 2020, *UK*
- Hotel entertainer/ animator
2019, *Hotel Gran Castillo Tagoro (Lanzarote), Activ*
- Founder and president of Latin Social Dance Society
2017 – 2019, *Durham University*
- Active Societies representative
2018 – 2019, *Durham Students' Union*
- Treasurer and social secretary
2016 – 2017, *Collingwood College Basketball Club, DU*
- Languages: English, Hungarian, Spanish
- Interests: dancing (latin), basketball, skiing, hiking

EDUCATION, CERTIFICATIONS & ACHIEVEMENTS

Academic

MSci in Computer Science and Mathematics

within the Natural Sciences programme

2015 – 2020, Durham University · Class II Division 1 (2.1)

Milestone Institute

2014 – 2015

The institute gathers talented Hungarians from secondary schools and prepares them to study at the top universities of the world, introducing them to foreign university systems (e.g. UK), providing extracurricular classes beyond those taught in secondary schools, mentorship, student clubs, and opportunities to grow both personally and professionally before, during, and after attending university.

Online

Deep Learning Specialisation by Andrew Ng

2021, deeplearning.ai, Coursera

- Neural Networks and Deep Learning
- Improving Deep Neural Networks
- Structuring Machine Learning Projects
- Convolutional Neural Networks
- Sequence Models

Azure Fundamentals Specialisation AZ-900

2022, Coursera

- Azure Cloud Services
- Azure Management and Security
- Azure Services and Lifecycles
- Preparing for the AZ-900

Create Machine Learning Models in Microsoft Azure

2022, Coursera

Achievements & Conferences

Guest lecturer: Introduction to AI and ML + use cases

2022, Business IT club, Corvinus University, Budapest

Presentation of our IoT solutions with live ML demo

2022, IoT Live Show by Yettel, Budapest

Data+AI Summit by Databricks

2022, online

+ Data Analysis with Databricks SQL training course

Data Innovation Summit

2022, online

Data Analysis with Databricks SQL training course

Reinforcement Conference

2022, Budapest

