Daniel T. Soukup

Data Scientist & Educator

☎ *+43 665 65115723*
✉ *daniel.t.soukup@gmail.com*
🖱 *danieltsoukup.github.io*
⊕ *danieltsoukup*
in *danieltsoukup*

Strengths

- 4+ years of R&D experience in math, machine and deep learning.
- Strong knowledge of data privacy, NLP and spatial analysis techniques.
- Engaging public speaker and content creator; dedicated project lead.
- High proficiency with `Python` and the software engineering work cycle.

Skills

- `Python`
- `pandas, numpy`
- `scikit-learn`
- `Keras, PyTorch`
- `NLTK, gensim, spaCy`
- `GloVe, w2v, BPEmb`
- `geopandas, skmob`
- `seaborn, folium`
- `PySpark`

- `Some Java & SQL`
- `Docker`
- `Linux & git workflow`
- `AWS & Google Cloud`

- `graph theory`
- `logic`
- `teaching & research`
- `science communication`

Work experience

Educator, Data Science @ BrainStation (2020 - present)

- Deliver and develop content for the full-time Data Science program.

Data Scientist @ Mostly AI (2019-2020)

- Synthetic data generation using deep neural networks; customer, transaction and dynamic mobility data modeling.
- Privacy analysis of data sharing; sequence embedding and clustering methods; bias mitigation in ML.

Data Scientist @ Uniqa Insurance Group (2019)

- Long-form text classification for medical claims processing; sub-word embeddings; combined CNN/LSTM models; transformers.

Course Instructor @ U of Toronto/Calgary/Vienna (2014 - 2019)

- Coordinating multi-sectional lectures, 200+ students and 5+ teaching assistant in calculus, linear algebra and combinatorics.
- Daniel B. deLury Teaching Award; mentoring; excellent student reviews.

Research Mathematician @ U of Calgary/Vienna (2016 - 2019)

- Local/global analysis of large graphs and digraphs; random combinatorial objects; decomposition problems for algebraic and topological structures.
- 20 refereed publications in top international journals; invited & plenary talks at conferences.
- Major EU funding for independent research projects (MSCA, €224,933).

Education

University of Toronto (2015)

- PhD in Mathematics
- Ontario Trillium Scholarship

Eötvös Loránd University (2011)

- MSc in Mathematics
- BSc in Mathematics

MOOCs (2018 - present)

- Applied Text Mining; Databases and SQL for Data Science [`IBM Db2 Warehouse`].
- Object-Oriented Software Design and Architecture Specialization.

Extras

- Synthetic Data Meetup (2020): protecting individual privacy with synthetic location data (article).
- data4good hackathon (2019): text analysis for the geo-distribution of urban greening projects.
- Cryptography outreach for high schools (2018).