Daniel T. Soukup

Data Scientist

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

Strengths

- 4 years of R&D experience in math, machine and deep learning.
- Strong knowledge of NLP and spatial analysis techniques and packages.
- High proficiency with `Python` and the software engineering work cycle.
- Dedicated project lead and team player; engaging public speaker.

Skills

- `Python`
- `pandas, numpy`
- `scikit-learn`
- `Keras, PyTorch`
- `NLTK, gensim, spaCy`
- `GloVe, w2v, BPEmb`
- `geopandas, skmob`
- `seaborn, folium`
- `PySpark`
- `Java`
- `SQL`
- `Linux & git workflow`
- `AWS & Google Cloud`
- `graph theory`
- `logic`
- `teaching & research`
- `science communication`

Work experience

Data Scientist @ Mostly AI (2019 - present)

- Developing and productionizing generative deep neural networks for synthetic data engines; technical lead for multiple customers.
- Focus: transaction and mobility data; clustering; privacy analysis of data sharing; sequence embedding methods.

Data Scientist @ Uniqa Insurance Group (2019)

- Medical claims classification using novel word-embedding techniques and combined CNN/LSTM architectures; mixed data sources; cloud computing and large-scale model training with GPUs.
- 9% improvement over the deployed model's KPI.

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

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

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

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

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, Design Patterns and Service-Oriented Architectures.

Extras

data4good hackathon (2019)

- Text analysis for the geo-distribution of urban greening projects; project website and github repo.
- Cleaning and processing unstructured data; topic modelling and NER; visualization.

Cryptography outreach for high schools (2018)