Michal Malyska

Machine Learning Team Lead, Semantic Health

https://michalmalyska.com

@ malyskamichal@gmail.com

Links -



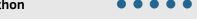
My Github

Academic Website

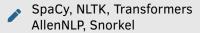
LinkedIn

Programming

Python



PyTorch, TensorFlow



Pandas, Scikit-learn, LGBM

♀ R



📃 tidyverse, RSTAN, brms, INLA

SOL



>_ VBA, SAS

Machine Learning -

Natural Language Processing Semi-Supervised Learning

Information Extraction

Statistical Learning

Explainable ML

Knowledge Graphs

Other -

Git Git

Docker, MLFlow, WandB

> Bash, ZSH

UMLS, SNOMED-CT

Contributions

MedSpaCy

AllenNLP

SciSpaCy

MS-BERT

Work Experience

January 2021

Machine Learning Team Lead

- Present

· Lead research and implementation of the clinical NLP product for a team of 5 ML Scientists from the idea stage to multiple successful client deployments.

· Work with the CEO to create and implement the organization's Machine Learning short, and long-term vision.

· Supervise a number of research projects with outside organizations resulting in product improvement and scientific publications.

June 2019 -

Machine Learning Scientist

December 2020

· Assessed, implemented, and improved deep learning algorithms from research and open source repositories to achieve and surpass current state-of-the-art performance on several clinical NLP tasks.

· Improved the data pipeline and model evaluation procedures for complicated NLP datasets in a highly multi-label setting.

January 2018 Analyst

- August 2018 · Worked client-site on business analytics, loss forecasting, and predictive modelling for a large insurance client.

> · Created an actionable overview of internal and external data, and analytical tools for a large public sector client

Education

2019 – 2022 **Master of Science - Statistics**

University of Toronto

University of Toronto

4.0 GPA

2015 – 2019 **Honours Bachelor of Science - Statistics**

4.0 GPA in Statistics, 3.78 overall

Teaching and Extracurricular

Summer 2020 Course Instructor

University of Toronto

Prepared materials and taught STA220 - Practice of Statistics I, to a class of 250 students in an online setting right at the start of the

pandemic

2018 – 2020 **Teaching Assistant** University of Toronto

Prepared and conducted tutorials for a number of courses, including:

Statistical Methods for Machine Learning, Data Analysis II

2020 Research Visitor St Michael's Hospital, Toronto

Publications

Active learning for medical code assignment 2021

arXiv:2104.05741

ACM CHIL 2021 workshop track

2020 **Multiple Sclerosis Severity Classification From Clinical Text**

2020.clinicalnlp-1.2

EMNLP 2020 Clinical NLP workshop

Competitions

2017 - 20191st. Place, Mentor **ASA Datafest**

Created a business case from click data for Expedia, aimed at improving the suggestion engine and customer retention, served as a

mentor in the subsequent years

October 2017 1st. Place

McKinsey Open Data Challenge

Created a business case and an MVP aimed at re-routing low urgency patients to hospitals with lowest estimated combined travel and wait time