

Michal Malyska

✉ malyskamichal@gmail.com • 🌐 michalmalyska.com • 🐙 Github

Work Experience

Machine Learning Lead

Semantic Health

Toronto, Canada

January 2021 - December 2022

- Spearheaded the research and development of the company's Machine Learning offerings, successfully delivering multiple client deployments and driving a \$60M valuation.
- Led a company-wide initiative to re-engineer development pipelines, resulting in a 40% increase in code and deployment velocities.
- Managed a team of 5 ML Scientists/Engineers, overseeing their growth and development through goal setting, creation and tracking of OKRs, conducting 1-1s, and performing performance reviews.
- Supervised research projects with external organizations, leading to significant product improvements and scientific publications

Machine Learning Scientist

Semantic Health

Toronto, Canada

June 2019 - December 2020

- Assessed, implemented, and improved on deep learning algorithms sourced from research and open source repositories, resulting in surpassing the existing state-of-the-art performance across multiple clinical NLP tasks.
- Revamped the data pipeline and refined model evaluation procedures for challenging NLP datasets in a complex multi-label environment.

Sessional Lecturer

University of Toronto

Toronto, Canada

2020-Present

- Developed comprehensive educational materials and delivered engaging instruction for a range of undergraduate and graduate courses in Statistics and Machine Learning.
- Successfully facilitated student learning by providing constructive feedback, guidance, and individualized support.

Education

University of Toronto

BSc. Statistics (3.89 GPA), MSc. Statistics(4.0 GPA)

2015–2022

- Research Visitor at St Michael's Hospital, President of Statistical Sciences Union, Dean's List, Teaching Assistant

Publications

- BioNLP - ICDBigBird: A Contextual Embedding Model for ICD Code Classification
- ACM CHIL - Active learning for medical code assignment
- EMNLP - Multiple Sclerosis Severity Classification From Clinical Text

Competitions and Awards

ASA Datafest

1st. Place, Mentor

University of Toronto

May 2017, 2018, 2019

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

McKinsey Open Data Challenge

1st. Place

McKinsey & Company

October 2017

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

Skills

- **Programming Languages:** Python, R, SQL, JavaScript
- **Libraries :** Pytorch, Tensorflow, SpaCy, NLTK, Transformers, Pandas, Sklearn, Git, Docker, MLFlow
- **Open Source Contributions:** MedSpaCy, AllenNLP, SciSpaCy, MS-BERT