

# Michal Malyska

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

## Work Experience

---

### Machine Learning Team Lead

**Toronto, Canada**

*Semantic Health*

*January 2021 - December 2022*

- Responsible for R&D and delivery of company's Machine Learning offerings from idea stage to multiple successful client deployments, across various business lines, which drive a \$60M valuation.
- Lead a company-wide effort to re-engineer the development pipelines resulting in a 40% increased code velocity.
- Managed a team of 5 ML Scientists/Engineers, including growth goal setting, creating and tracking OKRs, leading 1-1s, and conducting performance reviews.
- Supervise research projects with outside organizations resulting in product improvement and scientific publications

### Machine Learning Scientist

**Toronto, Canada**

*Semantic Health*

*June 2019 - December 2020*

- Assessed, implemented, and improved deep learning algorithms from research and open source repositories to 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.

### Course Instructor

**Toronto, Canada**

*University of Toronto*

*2020-Present*

- Prepared materials and taught a mix of undergraduate and graduate courses in Statistics and Machine Learning.

### Actuarial Analyst

**Toronto, Canada**

*Deloitte*

*2018*

- Deployed 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 & Publications

---

### University of Toronto

**2015–2022**

*BSc. Statistics, MSc. Statistics*

- Research Visitor at St Michael's Hospital, President of Statistical Sciences Union
- 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

**University of Toronto**

*1st. Place, Mentor*

*May 2017, 2018, 2019*

- Created a business case from click data for Expedia, aimed at improving the suggestion engine and customer retention
- Served as a mentor during the 2018 and 2019 competitions

### McKinsey Open Data Challenge

**McKinsey & Company**

*1st. Place*

*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