Supervisor Name: Daniel Felsky

Project Title: Integrative modeling of mental health and substance use in Ontario youth using machine learning

Hospital/Research Institution: The Centre for Addiction and Mental Health

Email: daniel.felsky@camh.ca

Field of Research (2 keywords): mental health, machine learning

Department: Krembil Centre for Neuroinformatics

School of Graduate Studies Appointment (IMS, LMP, IHPME etc)? Yes/No:

If YES, please name: Yes (IMS Associate Member)
**Brief Project Description (<300 words):**

The complex combinations of socio-demographic, geographic, lifestyle, and experiential factors that give rise to pervasive mental illness in Ontario Youth are not well understood. In collaboration with scientists at the CAMH Institute for Mental Health Policy Research, we have compiled high-dimensional data from over two decades of cross-sectional surveys distributed biannually to thousands of youth at schools across Ontario.

The successful candidate will be responsible for describing these data and developing models of primary mental health and substance use outcomes using statistical software (e.g. R, Python). The approach has three steps: 1) dimension reduction of questionnaire items to identify principal study outcomes, 2) statistical modeling of key outcomes, and 3) model interpretation.

Dimensionality reduction will require familiarity with the central concepts of principal components analysis. Modeling of depression, anxiety, social function, and substance use will be performed using gradient-boosted regression trees to allow for interactive and non-linear effects of all factors not used to define model outcomes. Interpretation will be performed using several “explainable AI” techniques, such as SHAP, to identify the constellation of predictive factors which are most associated with our key outcomes, and which of those factors may moderate the effects of others. One aim is to identify which major determinants of mental illness have changed over time or differ between geographic regions.

For success in this project, the student will require experience with statistical programming, as well as (at minimum) foundational exposure to regression modeling, cross-validation, and data visualization. This work has the potential to A) guide future research toward the most important factors determining mental illness in youth and B) inform provincial policy on the development of intervention strategies targeting the most vulnerable groups.