Predicting Fat and Protein content in donor milk using machine learning

Rachel Wong, Michael Pitino, Rafid Mahmood, Ian Zhu, Debbie Stone, Deborah O'Connor, Sharon Unger, Timothy Chan

Background

Donor milk is the standard of care for hospitalized very-low-birthweight infants when mother's milk is unavailable.

Testing every milk donation for nutrient composition is costly and labour-intensive.

Nutrient variability in donated milk complicates the production of a uniform pooled product by milk banks and thus the provision of adequate nutrition to promote optimal growth of infants.

Methodology

Machine Learning (ML) models for predicting donor milk macronutrient content, focused on fat and protein.

Samples of donor milk were from the Rogers Hixon Ontario Human Milk Bank. A baseline model was established using lactation stage and infant gestational status.





Train different ML models on input data

1. Ordinary least squares regression 2. Lasso-Least-Angle regression 3. Random forest regression 4. Gradient boosting regression

Make predictions for donations and their associated pools

• 4 different models on 7 different combinations of variables

Results

ML models were much more accurate than baseline at both the individual-donation and pool-level. This allows for optimizing which donations should be placed together in donor milk pools.



A trial of a two-step predict-optimize model is underway at the Rogers Hixon Ontario Human Milk Bank.

Centre Affiliates



Rachel Wong MASc Student



Ian Zhu PhD Student



FACULTY OF APPLIED SCIENCE & ENGINEERING



Timothy Chan Professor