UNIVERSITY of GUELPH

COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

MSc Defence

Thursday September 2, 2021 at 10AM via Zoom

Comparing Machine Learning and Deep Learning Techniques for Product Matching

Jeremy Foxcroft

Chair: Dr. Mark Wineberg
Advisor: Dr. Luiza Antonie
Advisory Member: Dr. Gary Grewal
Non-Advisory Member: Dr. Kris Inwood [Economics]

ABSTRACT:

Product matching is a specific application of record linkage where different digital records that refer to the same product are identified. In this thesis we design a framework to compare state-of-the-art product matching systems including traditional machine learning models and more recent deep learning approaches. We then employ this system to perform comparisons on both open source product matching benchmarks and real-world modern day industrial product data, measuring performance with both F1 measure and precision-recall curves.

We find that traditional machine learning techniques remain superior for clean, structured data and that this superior performance translates seamlessly from the open source product matching benchmarks to the real-world data. We also propose a new application for product matching: forecasting demand for products that are new to market.