Scaling *The Great Wall of Canada*: Technological solutions for more accessible and equitable language proficiency testing

Dr. M. Gregory Tweedie, Saina Kiani, Anika Tabassum
Werklund School of Education, The University of Calgary
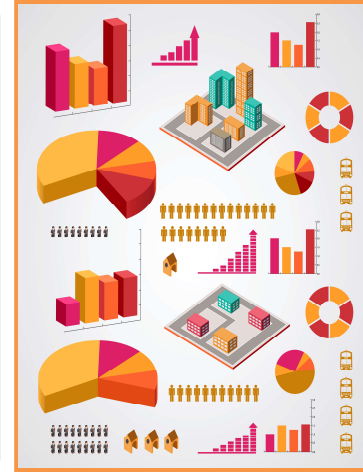
UNIVERSITY OF CALGARY

Background

- Meaningful employment: a critical milestone in newcomers' settlement process.
- Newcomers' un-/underemployment: costs to Canadian society include **economic**, **educational**, and **social cohesion**.
- Current situation of newcomers' language testing: **The Great Wall of Canada**: inequitable, inefficient, inaccessible.
  - Multiple test attempts (average of 3; as many as 14: see Hu & Trenkic, 2019; Hamid, 2016), USD$250 / attempt
  - "Black box" results (no feedback on performance)
  - Predatory marketplace of test tutors
- Strong research support: **locally developed tests** of **general language proficiency** + **test-taker demographic data** (e.g., age, nationality, test location), **when taken together**, can **predict language test results (**Daller et al, 2020; Müller et al, 2019; Daller et al, 2017; Daller et al, 2013; Eckes et al, 2006; Barkaoui, 2019; 2017; Davis, 2016; Orr, 2002; Wilson, 1987).
- **TestPredikt** innovation - test-takers receive:
  - An **accurate score prediction** before taking the test.
  - A data-driven **prediction** of **which** test and **where** they are likely to receive the **highest score**.
  - A recommended **tutor** customized to their **specific** test-taking **needs**.

Research Aims and Questions

Research questions:

- To what extent does performance on a general language proficiency test accurately predict scores on international standardized language tests?

- To what extent do demographic factors accurately predict scores on international standardized language tests?

Methodology

- Through self-reported test taker data on a variety of English language assessments, an array of features were reprocessed from quality to quantities.

- The process of converting sentences (literally list of words) to numbers is a complex process requiring knowledge of categorical hash assignment.

- Upon input readiness, a neural network was designed to digest inputs and de-form them to the desired scales.

- The training component is undertaken with a training dataset able to format weights of nodes and develop a predictive model.

- The model can then be deployed to the cloud for a prediction in less than a second.

Input features → Prediction

Results and Implications

Using machine learning neural networks, test takers:
- provide basic demographic data;
- complete an assessment-specific measure of general language proficiency; and,
- complete a vocabulary level test.

Our trials have shown increasingly accurate (>80%) predictions on:
- **Which test?** A recommendation on which language test they are most likely to receive a score required for their goals.
- **Where?** A prediction on the test centre (if applicable) where they are likely to receive the highest score.
- **What score?** A prediction of their anticipated score, indicated by a mathematical "confidence meter".
- **Who?** A recommendation for qualified tutors, given their individual language learning needs as identified in the pre-assessments.
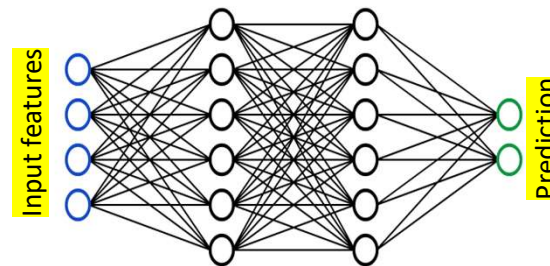
Newcomers will benefit in several ways:

1. **Reduced costs**.

2. **Reduced test attempts**.

3. **Reduced time**.

Anticipated software launch: December 31, 2021
- Contact: gregory.tweedie@ucalgary.ca
- Register for more information: https://testpredikt.ca/