Examining the association between elementary students lexcio-syntactic writing features and cognitive-motivational profiles using Natural Language Processing

Melissa Hunte
University of Toronto (OISE), Toronto, Ontario, Canada

Christine Barron
University of Toronto, Toronto, Ontario, Canada

Jeanne Sinclair
University of Toronto (OISE), Toronto, Ontario, Canada

Hyunah Kim
OISE, University of Toronto, Toronto, Ontario, Canada

Samantha McCormick McCormick
University of Toronto (OISE), Toronto, Ontario, Canada

Megan Vincett Vincett
University of Toronto (OISE), Toronto, Ontario, Canada

Eunhee Eunice Jang
University of Toronto, Toronto, Ontario, Canada

Abstract

Natural language processing (NLP) provides an innovative avenue to understand and explore human language content, yet minimal research has utilized it to classify students literacy, cognition, or motivation. This study investigated the association between grade 4-6 students (n = 143) writing and their cognitive-motivational profiles (CMPs) based on their self-regulated learning, locus of control, writing self-efficacy, and goal-orientation. LPA (Mplus 7.4) results indicated a two-class CMP solution with predominantly positive or negative CMPs. Using NLP, 404 lexico-syntactic writing features were extracted from students writing. Random forest with 10-fold cross-validation was implemented in Weka 3.8 (with SMOTE to equate class instances) to accurately (93%) classify students CMPs (class 1 True Positive Rate (TPR) = .942; class 2 TPR = .925) based on the NLP-processed lexico-syntactic writing features. These results highlight the potential for machine learning to analyze students writing and accurately classify learner profiles to provide formative feedback and customized interventions.