School of Mathematical and Statistical Sciences
Arizona State University

M E M O R A N D U M

DATE: 03/27/2023

TO: Faculty and Students

FROM: Professor(s) Chair/Co-Chairs of Defense for the Committee Members

Paul R. Hahn
Chelsea Krantsevich

PhD in Applied Mathematics

Hedibert Lopes
Rosemary Renaut
Visar Berisha
Yi Zheng

DEFENSE ANNOUNCEMENT

Candidate: Chelsea Krantsevich

Defense Date: Monday, April 10, 2023

Defense Time: 9:00 AM

Virtual Meeting Link:
https://asu.zoom.us/j/89719219037?pwd=Y21sTkhKbmg2cGowWnI0di9pMUpzQT09

Title: Machine learning for the design of screening tests: general principles and applications in criminology and digital medicine

Please share this information with colleagues and other students, especially those studying in similar fields. Faculty and students are encouraged to attend. The defending candidate will give a 40 minute talk, after which the committee members will ask questions. There may be time for questions from those in attendance. But, guests are primarily invited to attend as observers and will be excused when the committee begins its deliberations or if the committee wishes to question the candidate privately.

ABSTRACT
-See next page-

PO Box 871804 Tempe, AZ 85287-1804
(480) 965-3951 Fax: (480) 965-8119
http://math.asu.edu

ABSTRACT

This thesis explores applications of machine learning methods in service of the design of screening tests, which are ubiquitous in applications from social work, to crim- inology, to healthcare. In the first part, a Bayesian decision theory framework is presented for designing tree-based adaptive tests. On an application to youth delin- quency in Honduras, the method produces a 15-item instrument that is almost as accurate as a full-length 150+ item test. The framework includes specific considera- tions for the context in which the test will be administered, and provides uncertainty quantification around the trade-o↵s of shortening lengthy tests.

In the second part, classification complexity is explored via theoretical and empirical results from statistical learning theory, information theory, and empirical data complexity measures. These formalisms are used to understand the relative merit of speech elicitation task engineering and speech feature engineering when designing a speech-based screening test to detect cognitive impairment. Through an extensive classification analysis on a clinical speech dataset from patients with normal cogni- tion and Alzheimer's disease, we show that the speech data collection protocol has a fundamental impact on test performance. Carefully designed task and feature engineering are required for best results, and methods are proposed for automatically extracting insights to guide speech-based screening test design.