Mathematical & Statistical Sciences Department

Announcing:

Ph.D. Defense of Masabho Peter Milali

Machine Learning and Data Mining Methods to Estimate Party Status and Age of Wild Mosquito Vectors of Infectious Diseases from Near-infrared Spectra

By Masabho Peter Milali

Friday, February 21, 2020

9:00 - 11:00 A.M.

Olin Engineering 540

Committee Director: Dr. George Corliss
Committee Member: Dr. Naveen Bansal
Dr. Samson Kiware
Dr. Serdar Bozdag
Dr. Richard J. Povinelli

Abstract

This dissertation first demonstrates that training models on near-infrared spectra (NIRS) using artificial neural network (ANN) as an architecture yields models with higher accuracies than training models using partial least squares (PLS) as an architecture. In addition, irrespective of the model architecture used, direct training of a binary classifier scores higher accuracy than training a regressor and interpreting it as a binary classifier. Furthermore, for the first time, this dissertation shows that training ANN models on autoencoded near-infrared spectra yields models that estimate parity status of wild mosquitoes with an accuracy of \( \approx 93\% \), which is strong enough to support NIRS models as an alternative to ovary dissections. Results from this dissertation also show that there is no significant difference between spectra collected from semi-field raised and wild mosquitoes of the same species supporting the on-going practice of training models on semi-field raised mosquitoes to estimate the age class in days of wild mosquitoes. Finally, the study shows that an ANN model trained on semi-field mosquitoes classifies wild mosquitoes into either less than or greater or equal to seven days old with an average accuracy of 76%. In conclusion, the results in this report strongly suggest the use of artificial neural networks as a suitable architecture to train models that estimate parity status and age in days of wild mosquito vectors of infectious diseases. The results further suggest near-infrared spectroscopy as an appropriate alternative tool to estimate different parameters of mosquito vectors of infectious diseases.

Open to the public!