



# COLLEGE of ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

## PhD Seminar 2

**Wednesday May 3, 2023 at 1pm via Zoom [Remote]**

**Akshay Chadha**

*Application of Machine Learning Techniques for  
Genotypic-Phenotypic Linkage of H5 Avian Influenza and PRRS Viruses*

**Advisor:** Dr. Rozita Dara

**Co-Advisor:** Dr. Zvonimir Poljak [Population Medicine]

**Advisory:** Dr. Dan Gillis

**Advisory:** Dr. David Pearl [Population Medicine]

### **Abstract:**

Over the past decades, numerous viral outbreaks resulted in widespread economic and livestock losses, e.g., avian influenza virus (AIV), porcine reproductive and respiratory syndrome virus (PRRSV), and several other viruses, including the recent COVID-19 outbreak. Early classification is crucial to timely intervention, especially in healthcare monitoring, intrusion detection, and viral outbreaks. The genotypic-phenotypic relationship establishment studies may include sequence classification, a widely studied task using machine learning (ML) techniques. Despite the availability of high-dimensional viral genetic sequences, for several viruses, the genotypic-phenotypic relationship establishment has yet to be understood entirely.

Establishing genetic-phenotypic relations is an interesting but challenging task due to an incomplete understanding of how virulence-encoding information is constrained inside limited vocabulary-based high-dimensional genomic sequences. The success of ML techniques poses an important solution for identifying the pathogenicity of different viruses in a data-driven manner conditional on sufficient input data used for training and if such a relationship exists.

This research is aimed to (1) investigate the classification of viral genomic sequence datasets based on phenotypic outcomes observed in the form of clinical impact for pathogenicity classification of haemagglutinin (HA) gene sequences of H5Nx AIV and PRRSV using open reading frame-5 region. (2) assess the plausible contribution of sequence positions in establishing genotypic-phenotypic relationships. The study signifies the challenges in ML applications when sufficient data may be unavailable. The results of this research may provide a perspective to virologists, veterinarians, and scholars specifically working in the field of AI and PRRS and also lays a pathway for future genotypic-phenotypic studies.