UNIVERSITY of GUELPH

ONTARIO AGRICULTURAL COLLEGE
DEPARTMENT OF ANIMAL BIOSCIENCES

PhD. Defence
Applications of data-driven algorithms for breeding programs in the dairy industry

Lucas Alcantara

Date: June 14th 2022 at 9:30am

The PhD Defence for Lucas Alcantara has been scheduled for June 14th, 2022 at 9:30m. The defence will be held online via Teams: https://teams.microsoft.com/l/meetup-join/19%3ameeting_YmY0NDc0MDktNzQ3OS00Y2VmLWIyODMtMGUxMzJjZmExZTVk%40thread.v2/0?context=%7b%22Tid%22%3a%22be62a12b-2cad-49a1-a5fa-85f4f3156a7d%22%2c%22Oid%22%3a%22fbd28915-dda5-478f-8ecb-a3682dcf0c3a%22%7d

The exam committee will consist of:

Examining Chair: Dr. Vern Osborne

Advisor: Dr. Flavio Schenkel

Adv. Committee Member: Dr. Dan Tulpan

Additional Graduate Member: Dr. Mehdi Sargolzaei

External Examiner: Dr. Cedric Gondro

Abstract:

Data-driven modelling with intelligent algorithms for the improvement of dairy breeding programs has been receiving increasing attention. In this thesis, dimensionality-reduction methods and machine learning (ML) algorithms were applied in two sets of traits within the dairy industry: conformation and fertility. Firstly, multiple polynomial regression and principal component analyses were used to assess the association of 26 conformation traits with the Pro$ selection index, providing information to guide the dairy industry on which traits are more important from a monetary standpoint and which could likely be removed from genetic evaluation or recording programs. Secondly, a flexible and robust methodology to vectorize any text-based Breeding Protocol Descriptions (BPD) and develop ML classification models to identify Timed Artificial Insemination (TAI) protocols was developed. Timed Artificial Insemination masks an animal's true fertility performance, reducing the accuracy of genetic evaluations for fertility traits. Mitigating this potential bias is challenging since there is a lack of specificity and uniformity in the recording of BPD by dairy farmers. Correct identification of TAI protocols would open the opportunity for unbiased genetic evaluation of animals based on their natural fertility. Lastly, in a third study, decision tree-based algorithms were applied to correctly determine the outcome of historical insemination records as open, pregnant, or aborted. The Canadian Data Exchange System currently stores insemination records performed by artificial insemination technicians, but those do not include the insemination outcome (IO), which is recorded on the farmers' herd management software. Therefore, researchers and geneticists do not have access to IO on a national scale for fertility research and genetic evaluations. Therefore, classification models proposed could have a positive impact on the dairy industry by using high performance predictions of IO to broaden the opportunity for new fertility research and genetic evaluations. Altogether, this thesis leveraged data-driven algorithms to provide the dairy industry with understanding on the monetary importance of individual conformation traits, a robust methodology to identify breedings that used hormonal synchronization protocols, and models to correctly determine the outcome of historical insemination records for use in breeding programs.