Year 5: Developing a Novel Multiple Linear Regression Model To Optimize Honey Bee Gut Immunity Using a Lactic Acid Bacteria Probiotic Mixture

Madan, Varun (School: Lake Highland Preparatory School)

Honeybees are essential to society, performing over 80 percent of worldwide pollination. However, commercial hive populations have been decreasing at an alarming rate in the United States. Research throughout the first three years of the project revealed that treating the hives with a specialized lactic acid bacteria probiotic mixture containing B. infantis, B. bifidium, and L. kunkeei significantly reduced the counts of the harmful gut parasite Nosema ceranae. An existing Scikit-learn multiple linear regression model was used last year to discern the optimal therapeutic index of the previously proven bacterial treatment for any given initial concentration. However, a major problem with this (and other existing machine learning alternatives) was that the model would not minimize the sum of the residuals unless the explicit training command was given. To this end, a novel multiple linear regression model was developed with continual-learning algorithms, and its accuracy was compared with experimental values & existing ML alternatives. The results showed an extremely significant improvement in output accuracy, as the margin of error in predicted Nosema concentration significantly decreased at every treatment dosage value throughout the experiment. Additionally, when compared to existing machine learning libraries, the model’s continual-learning algorithms predicted output values with significantly more accuracy (peak of 92.8%) without sacrificing significant processing time. After continuing to integrate this multiple linear regression model into a user-friendly iOS application, farmers will potentially have the ability to determine the ideal dosage of the bacterial treatment just moments after diagnosing their hives' Nosema concentrations.

Awards Won:

First Award of $5,000

EU Contest for Young Scientists Award