Evaluating Machine Learning Methods for Modeling the Distribution of the Two Form Bumble Bee (Bombus bifarius) in Wyoming

Ramesh, Padmalakshmi (School: Laramie High School)

Machine Learning (ML) algorithms, part of Artificial Intelligence, allow scientists to analyze large datasets and identify patterns. Processing large datasets prior to the introduction of ML methods was tedious. Ecologists have been collecting data on plants and animals, but without efficient methods, many patterns have not been detected. Bees are a group of insects that are essential to the environment. Bumblebees (Bombus bifarius), a species of bees, can tolerate cooler temperatures making them important in cold environments. Bumblebees along with the rest of the bee populations are declining due to factors such as climate change and increased use of pesticides. Knowledge of their habitats and monitoring changes in their population are essential to prevent further decline. ML methods can be used to analyze and generate habitat models to determine areas for conservation. Bombus bifarius observational data were acquired from Wyoming Natural Diversity Database, along with elevation, temperature, precipitation, and landcover data. Presence (n=221) and pseudo-absence (n=221) data were analyzed with five ML methods in R software. These models were evaluated based on their ability to correctly predict the presence and absence of Bombus bifarius. The Naïve Bayes model had the highest accuracy (80.3%) and was used to generate the first statewide habitat map of Bombus bifarius in Wyoming. The ML models give an insight into the preferred habitats of Bombus bifarius and the distribution map can help create focused efforts to protect this species. These ML models could detect hidden patterns in the data about Bombus bifarius habitats in Wyoming.

Awards Won: