Cattle Breeding Value Modeling

The background:

A South African animal registration and data information management organization, which is used by various breeders all over South Africa, have a significant size and diverse database of Bonsmara cattle features. Bonsmara farmers commonly utilize this data base to guide them in deciding which animals to pair such that the best offspring is produced for a specific category. For each these specific categories an animal is given a score scores (such as the animal milk value, fertility value, and carcass value), which ultimately determines which animals are most frequently bought at auctions.

The problem:

Determine which cow and bull to pair to maximise a scoring value (sought after by the farmer) given the cow and bulls genetic building blocks and phenotypic data.

The solution:

COGO Data trained several machine learning models that take three generations worth of genetic and phenotypic data for a selected cow and bull and predicts the scoring values of the offspring. These models were then combined into a ensemble that makes more accurate predictions.

The result:

The ensemble model obtained slightly improved predictions over the traditional model. However, the ensemble was able to improve the accuracy for one variable by 55%. This shows that the model can be optimized to further improve predictability for other variables.