Name(s)

Adham Elarabawy

Project Number

S0812

Project Title

Predicting Optimal Farming Regions via Machine Learning Trained on Novel Vegetation Index

Abstract

Objectives

As the world population grows and the usable farming area shrinks, the demand for nutrition increases, leading to an escalating pressure on farmers to increase their yield. Due to this, many farmers resort to using fertilizer and utilizing biotechnology to make the individual products larger and more nutritious. However, the uncertainty of overall crop yield is yet to be fully addressed, which this project attempts to undertake. Precisely and reliably predicting plant growth and relative crop yield proves to be very helpful for developing countries as well as existing farms, as it allows farmers to invest an optimum amount of resources due to their knowledge of its respective potential, as well as helps optimize their farmland placement for yield and reliability.

Methods

Firstly, in order to quantify the potential for vegetation, a novel composite vegetation index computed through calculations on image bands of different wavelengths from multispectral images is derived. To compute the novel composite vegetation index (NCVI), multispectral images are obtained from the Landsat 6, 7, & 8 satellites, from which the infrared and red bands are retrieved. In order to prevent cloud reflectances from influencing the NCVI calculation and consequently impacting the prediction algorithm, machine learning is utilized to disregard pixels determined as clouds. Once the cloud pixels are filtered out of the satellite images, NCVI is computed and passed into the output dataset utilized by the prediction algorithm.

Results

The algorithm reliably predicts relative crop yield to within 10% for 3 subsequent seasons when trained on only 11 seasons of prior data. Furthermore, the predicted relative crop yield output from the algorithm is compared to the true crop yield of two industrial farms in Fresno, the results of which lead the refinement and further development.

Conclusions

As opposed to the current leading vegetation indices, not only does the novel inclusion of moisture into the vegetation index lead to more immediate observations in short-term environmental changes, but it also counteracts the typical saturation of vegetation indices at higher densities of vegetation through the increased granularity.

Summary Statement

A novel composite vegetation index computed from multispectral images is passed to a machine learning algorithm to identify and predict optimal farming regions.

Help Received

I developed and tested the software presented independently. I received some preliminary help and exposure to the topic from my previous software engineering internship with GroGuru Inc.