Artificial Intelligence Prediction of Virtual Reference Values for use in Sensor-Based Nitrogen Management

*Nishan Khatri[1,2] ; Jackson Stansell[1]

[1]Sentinel Fertigation, Lincoln, USA; [2]University Of Nebraska, Omaha, USA

Motivation

N status indices, such as the sufficiency index (SI), may be derived from vegetation index (VI) values to inform in-season N management. Because N status index computation requires a comparison of VI values in non-N-limited crop regions with VI values in potentially N-limited crop regions, the establishment of a N-rich or N-poor strip has traditionally been required during a fertilizer application prior to crop canopy reflectance data use. The establishment of N-rich and N-poor strips, however, is cumbersome since it is time-consuming for operators and requires significant data preparation and processing to be useful in image processing routines. Previous research has been undertaken to investigate N status index derivation methods that do not require an N-rich or N-poor strip. One of the more promising approaches is the virtual reference (VR) concept which uses the cumulative 95th percentile of the VI distribution to provide the reference VI value traditionally derived from an N-rich strip. Application of this technique for in-season N management decisions, however, has shown that the VR tends to overapply N and that the VI reference value determined using this method may not align with the VI value measured in an established high-N strip. Realizing the importance of Artificial Intelligence (AI) for the agricultural revolution, we apply AI techniques to determine appropriate VI values to use for computing N status indices without the need for a N-rich or N-poor strip. We train different machine learning and deep learning models to minimize the error rate between the target value and the value predicted by our AI model. The metrics used to evaluate the effectiveness of our approach are Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). With both RMSE and MAE values near 0.01 in most AI models, as well as NDRE differences and SI values that closely mirror those computed from the original dataset, our approach shows significant early promise for enabling the computation of N status indices without an N-rich or N-poor strip.

Methods

1. **Input**: Raw data in the form of GeoTIFF images or shapefiles containing NDRE values and excel file with mean NDRE values for each pair of AOIs.
2. **Preprocessing**: Extract NDRE values from each AOI along with their corresponding target values of N-rich and N-poor regions from excel file.
3. **Feature Engineering**:
   - Machine Learning(ML) algorithms requires feature engineering.
   - Make shape of each AOIs same: Calculate 1st to 99th percentile values and use them as feature.
4. **AI algorithms**:
   - Regression like problem with image or shapefile containing NDRE values as input.
   - Two separate models for N-poor and N-rich strips.
   - Divide the dataset into training(85%) and validation(15%):
   - Deep Learning(DL): Tune a custom Convolution Neural Network(CNN) architecture that could handle multiple feature size as size of each pair of AOIs is different.
   - Machine Learning(ML): Implement different ML algorithms: *Random Forest*, *XGBoost* and *Support Vector Regressor(SVR)*
5. **Evaluate**: Use Root Mean Square Error(RMSE) and Mean Absolute Error(MAE)
6. **Predict**:
   - Save the trained model and make prediction.
   - Analyze the predicted values.

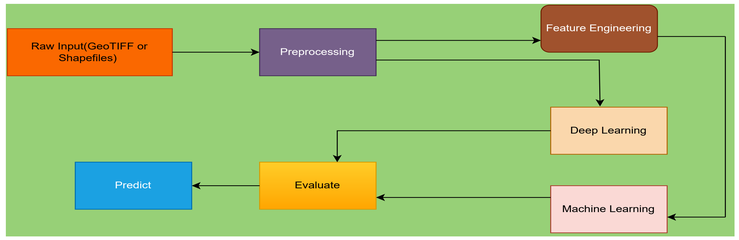

Fig 1 – Artificial intelligence learning methodology diagram comparing deep learning and machine learning approach processes.

Result

Evaluation Result on Validation Dataset:

Algorithm	RMSE		MAE	
	N-poor	N-rich	N-poor	N-rich
Deep Learning	**0.0187**	**0.0170**	**0.0131**	**0.0120**
Random Forest	0.0186	0.0175	0.0133	0.0124
XGBoost	0.0192	0.0207	0.0303	0.0155
SVR	0.0376	0.0378	0.0138	0.0326

Training and validation loss:

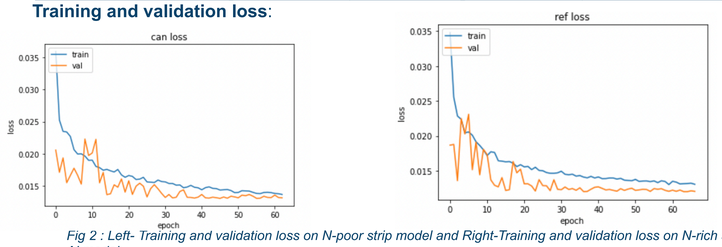

Fig 2 : Left- Training and validation loss on N-poor strip model and Right-Training and validation loss on N-rich strip AI model

Analysis of predicted values on validation data:

1. Statistical analysis of difference between N-rich and N-poor strips:

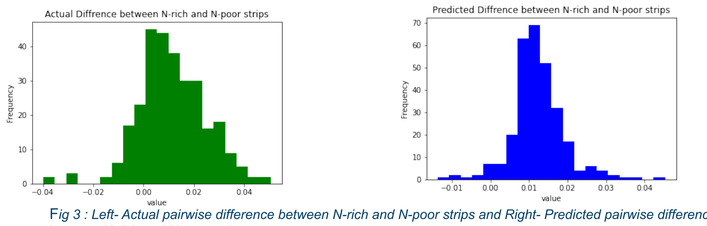

Fig 3 : Left- Actual pairwise difference between N-rich and N-poor strips and Right- Predicted pairwise difference between N-rich and N-poor strips

	Actual	Predicted
SD Difference	0.013471	0.006814
Mean Difference	0.011366	0.012220
Median Difference	0.010348	0.011690

2. Statistical analysis of Actual and Predicted SI values:

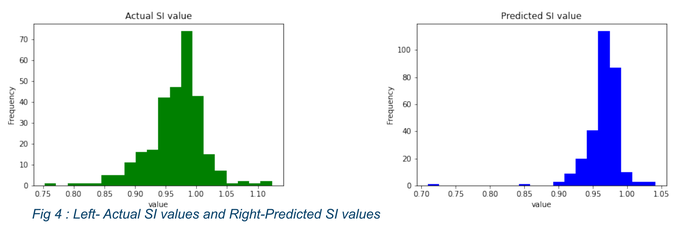

Fig 4 : Left- Actual SI values and Right-Predicted SI values

	Actual	Predicted
SD SI	0.045903	0.026009
Mean SI	0.966315	0.965150
Median SI	0.974266	0.969368

Datasets

- Unmanned Aerial Vehicle(UAV) multispectral imagery data.
- Historical data collected across multiple irrigated corn production fields during 2020 and 2021.
- Dataset has Normalized Difference Red-Edge index(NDRE) values collected in the form of:
  - GeoTIFF images
  - Shapefiles
  - Area of interest(AOI) boundaries for known N-rich and N-poor regions located in specific sub-field AOIs like sectors and management zones.
  - Excel file containing the NDRE mean values of N-rich and N-poor regions of each AOI: Act as target value for our AI model.
  - Each AOI acts as a separate data row.
- Total data rows: 1944

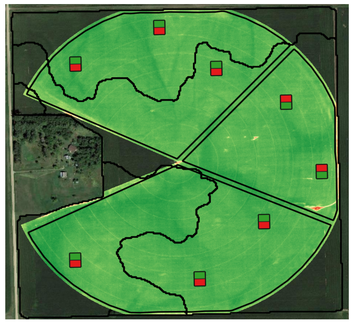

Fig 5: Sample dataset containing several Sector-Zone combination

Discussion

- Even with less data, result seems satisfactory.
- Machine learning and deep learning gives similar result.
  - Machine Learning requires feature engineering.
  - Deep learning doesn't require feature engineering but takes longer time to train.
  - Deep learning results are better.
- The current implementation takes shapefile as input.
  - Using GeoTIFF files as input is computationally expensive.

Conclusion

- RMSE and MAE values near 0.01 for AI model.
- The statistical analysis on predicted values shows decent output.
- The overall result displays convincing output for enabling computation of N status indices neglecting N-rich and N-poor strip.
- Future work:
  - Get more data which can increase the performance.
  - Work on entire GeoTIFF image file.

References

- Kelly H. Holland et al.,2012; Precision Agriculture, 14:71–85; DOI 10.1007/s11119-012-9301-6

SENTINEL FERTIGATION

2022 NITROGEN USE EFFICIENCY WORKSHOP

August 1–3

University of Nebraska, Lincoln

UNIVERSITY OF Nebraska Omaha