GEO-AI Challenge for Cropland Mapping

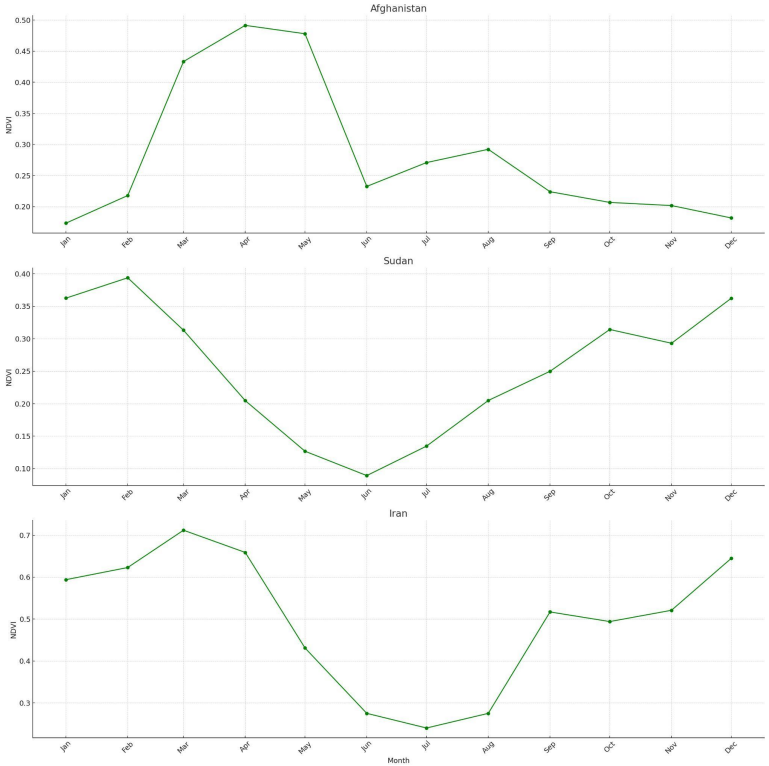
Background

- Timely, accurate crop maps are vital for numerous agricultural applications.
- Limitations of Existing Products: Lack annual updates, reducing usefulness for monitoring changes.
- New Opportunities: High-resolution open data (e.g. Sentinel 2) and machine learning models enable scalable modeling of cropland extent.

Methodology and WorkFlow

- Crop Season identification using NDVI Temporal Profile
- Sentinel 2 Data Processing - Cloud Masking/Compositing
- Vegetation Indices generation
- Model Development - HyperParameter Tuning and Classification Accuracy
- Crop Extent generation

Crop Season from NDVI Profile



Sentinel- 2 Data Processing

- Cloud Masking was done using the Cloud Probability dataset.
- All pixels with a cloud probability of 50 or above were masked out.
- The value of 50 was chosen to ensure a good balance between quality and quantity for the analysis.

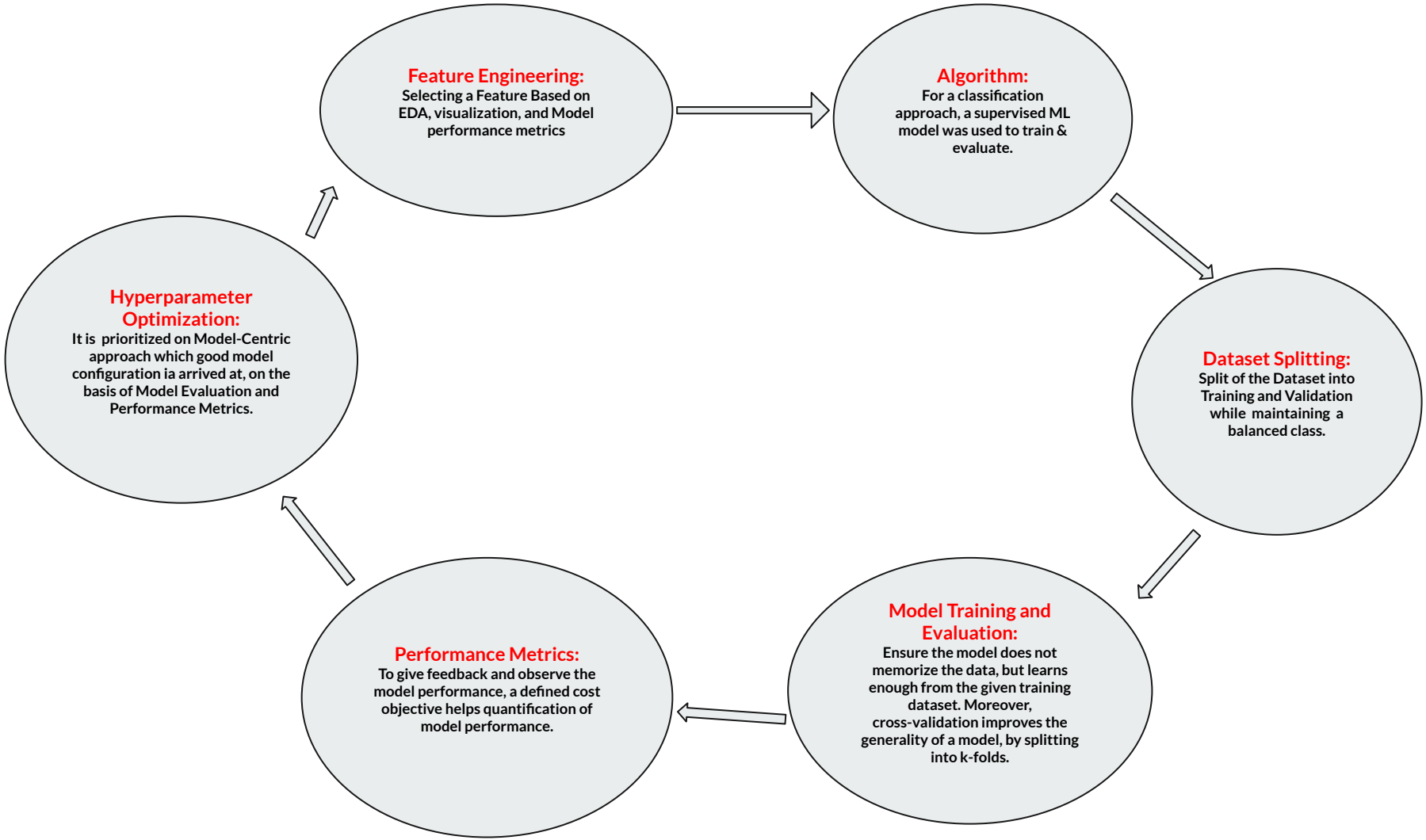
Vegetation Indices

- A wide variety of vegetation Indices were generated for each composite of the crop season to establish contrast between crop and non crop pixels
- Following Indices shortlisted
- NDVI, NDRE, ARVI, RVI, SCI, NDMI, GNDVI, BSI

Model Development

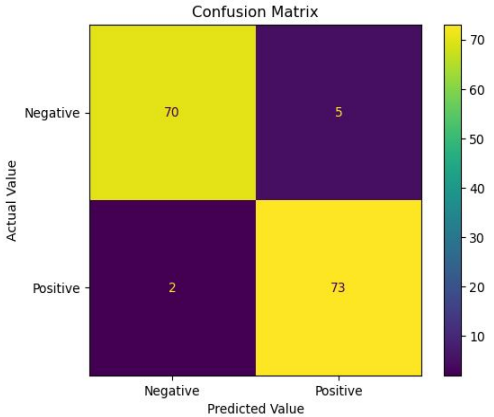
- It is an iterative process in which many models were derived, evaluated, and tested until achieving the best performance in terms of accuracy results.

- Typically focus on model creation, experimentation, and evaluation.

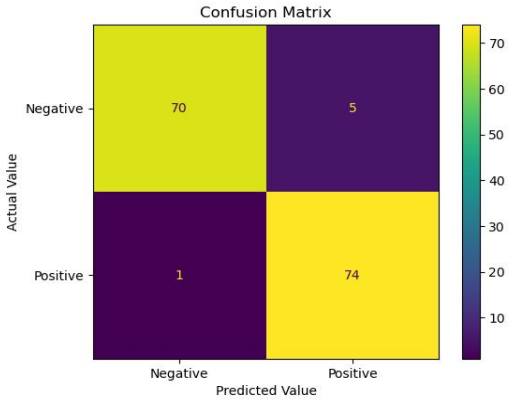

Feature Engineering:
Selecting a Feature Based on EDA, visualization, and Model performance metrics
Algorithm:
For a classification approach, a supervised ML model was used to train & evaluate.
Dataset Splitting:
Split of the Dataset into Training and Validation while maintaining a balanced class.
Model Training and Evaluation:
Ensure the model does not memorize the data, but learns enough from the given training dataset. Moreover, cross-validation improves the generality of a model, by splitting into k-folds.
Performance Metrics:
To give feedback and observe the model performance, a defined cost objective helps quantification of model performance.
Hyperparameter Optimization:
It is prioritized on Model-Centric approach which good model configuration ia arrived at, on the basis of Model Evaluation and Performance Metrics.

Classification Accuracy

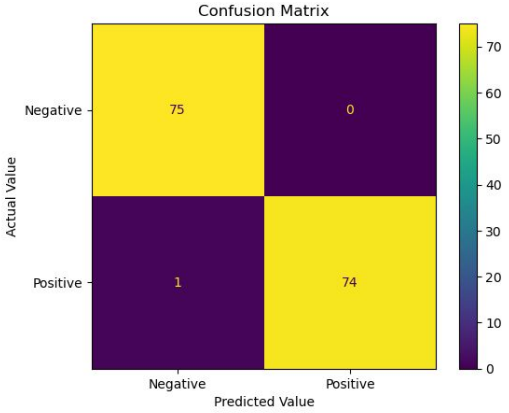
Afghanistan



Iran

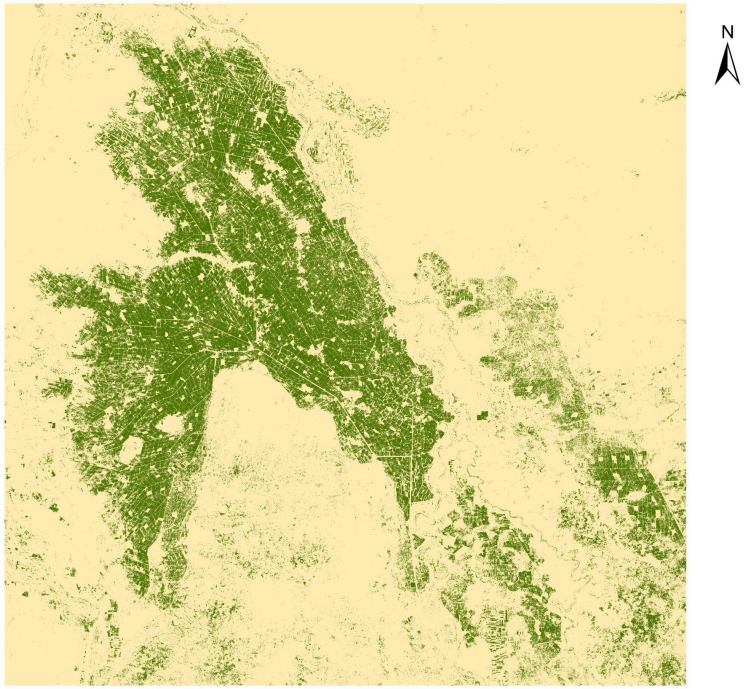


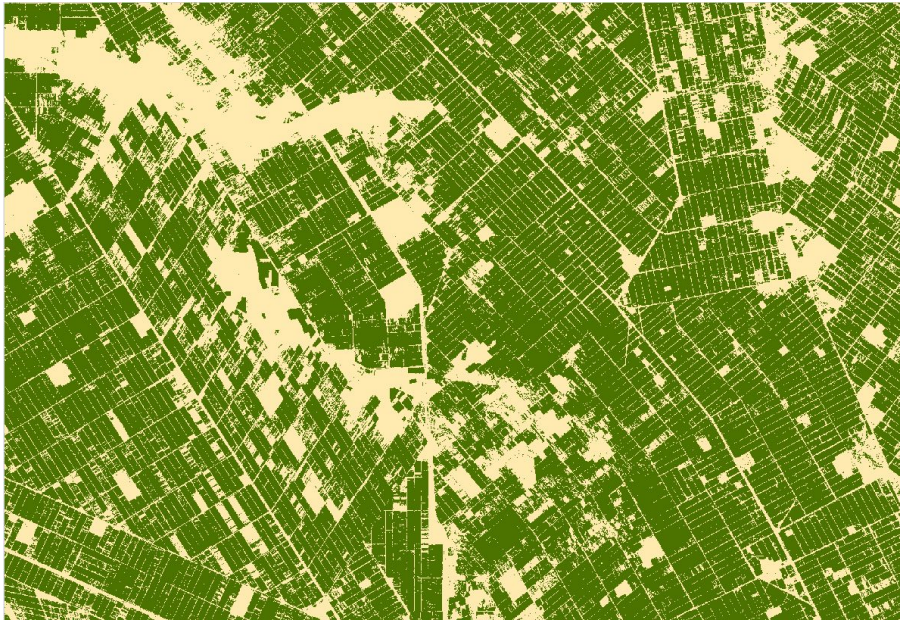
Sudan



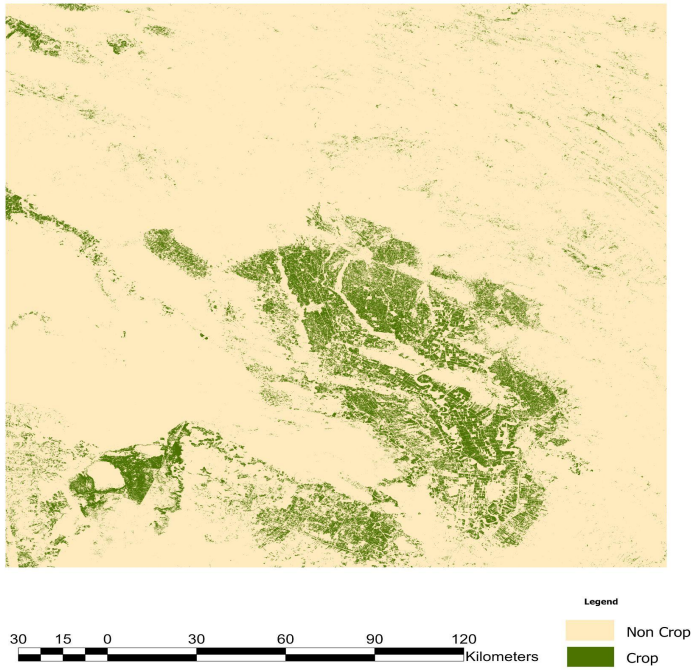
- **Negative: Non-Crop**
- **Positive: Crop**

Spatial Extent - Sudan





Spatial Extent - Iran



Spatial Extent - Afghanistan