Model for Sustainable Deployment of Climate Smart Agriculture Practices among Smallholder Farmers in Kakamega County

Simon Ndung'u
Moses Thiga, PhD
Vitalis Ogema, PhD
Philip Wandahwa, Phd

Sponsored by Artificial Intelligence Journal www.elsevier.com/locate/artint, Google TensorFlow

Background

- Kenya is dominated by 4.5 million smallholder farmers who produce over 75% of agricultural production
- CSA interventions have been developed to increase smallholder farmers' resilience to climate change, reduce GHG emissions and increase agricultural productivity.
- CSAs are production methods that combine various sustainable methods that address climate

Problem

- The current CSA interventions are supply-driven; proposing blanket recommendations for all smallholder farmers in all agro-ecological zones
- Smallholder farmers lack critical climate smart agriculture decision making tools including information on the appropriate interventions and technologies to implement in their farms.
- This makes the farmers to abandon the technologies once the projects close.

Study / Project Objectives

- To develop a suitable Machine Learning model for the deployment and adaptation of CSA practices among smallholder farmers in Kakamega county
- To prototype the Machine Learning model for the deployment and adaptation of CSA practices among smallholder farmers in Kakamega county

Background Literature

There are several Models developed for agriculture:

1) Johann et al. (2016) estimated the soil moisture content using an autoregressive error function: this model is suitable to estimate soil moisture in controlled systems that apply no no-till machinery.

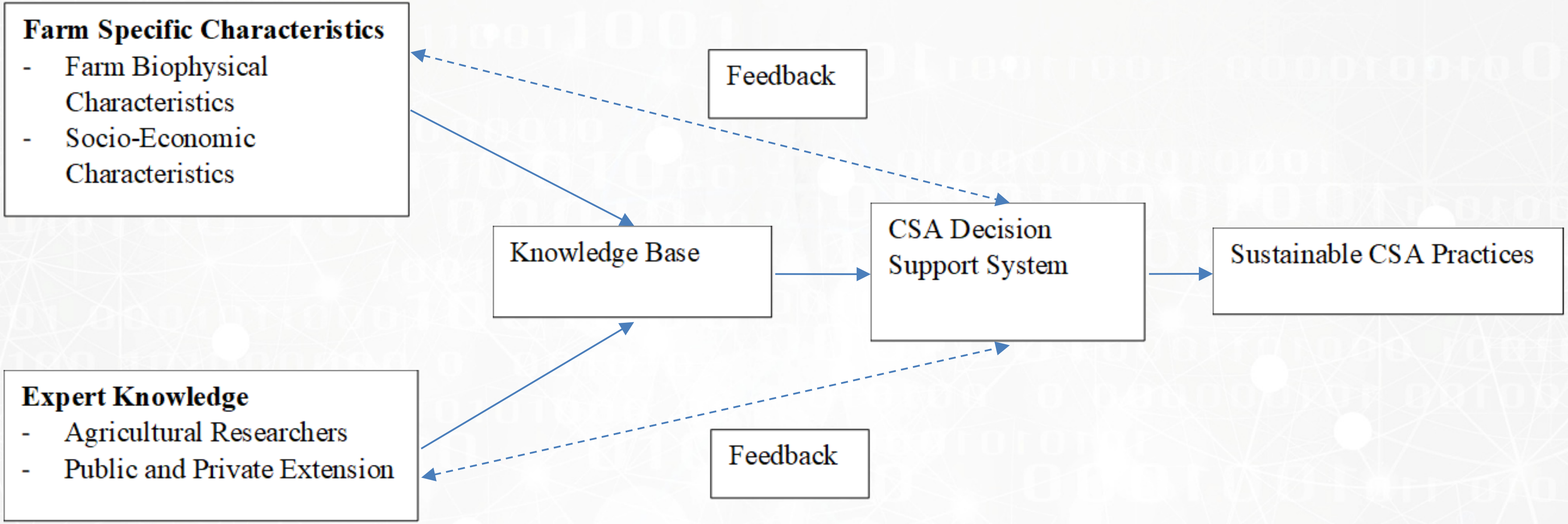
2) Chen, et al. (2014) designed a Wireless Sensor Network (WSN) to monitor multi-layer soil temperature and moisture in a farmland field to improve water utilization and to collect basic data for research on soil water infiltration variations for intelligent precision irrigation.

3) Panchard (2007) developed a DSS aimed at improving resource-poor farmers' farming strategies in the choice of crop varieties, planting and harvesting dates, pests and disease control and efficient use of irrigation water.

4) GPFARM, developed by Ascough Li et al. (2002), contains risk analyses that combine projected crop yield and animal production data with concurrent environmental impact data.

5) Aggarwal et al (2018), developed a CSV approach that gives guidance before and during the planting season on the most suitable CSA practices and technologies

Conceptual Framework



Methodology

1) Primary data was collected from 428 smallholder farmers in Kakamega County (182 adopters and 246 dis-adopters). This exercise yielded 610 variables

2) Pearson's Correlation coefficient was used to identify the variables that influence smallholder farmers' adoption/dis-adoption of CSA technologies. This exercise yielded 61 variables

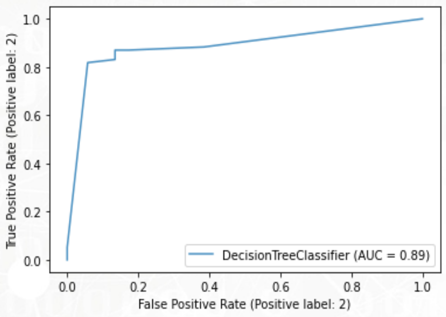
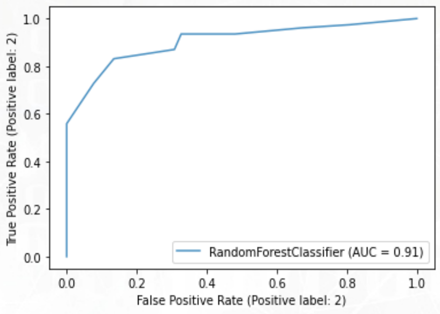
3) The Google Collaboratory notebook was used for the model fitting and testing process. Model fitting was done to measure how well the ML models generalize to similar data to that on which they were trained.

4) Decision Tree and Random Forest Models were identified as

Confusion Matrix

Decision Tree Classifier			Random Forest Classifier		
	Adopter	Dis-adopter		Adopter	Dis-adopter
Adopter	45	7	Adopter	45	7
Dis-adopter	11	66	Dis-adopter	13	64

Visualization ROC Curves

Decision Tree Classifier	Random Forest Classifier
	

The models produced AUCs of 0.89 and 0.91 under the Decision Tree Classifier and Random Forest Classifier, respectively

Model Metrics

Metric	Decision Tree	Random Forest
Training Accuracy	0.9431438127090301	0.9966555183946488
Prediction Accuracy	0.8604651162790697	0.8449612403100775
Precision / Sensitivity	0.8035714285714286	0.7758620689655172
Recall	0.8653846153846154	0.8653846153846154
Specificity	0.8653846153846154	0.8653846153846154

Model Classification Report

	Decision Tree Classifier				Random Forest Classifier			
Metric	Precision	Recall	F1-Score	Support	Precision	Recall	F1-Score	Support
Adopt	0.80	0.87	0.83	52	0.78	0.87	0.82	52
Dis-Adopt	0.90	0.86	0.88	77	0.90	0.83	0.86	77
Accuracy			0.86	129			0.84	129
macro avg	0.85	0.86	0.86	129	0.84	0.85	0.84	129
weighted avg	0.86	0.86	0.86	129	0.85	0.84	0.85	129

Model Accuracy

Metric	Decision Tree	Random Forest Classifier
MEA	0.13953488372093023	0.15503875968992248
MSE	0.13953488372093023	0.15503875968992248
RMSE	0.37354368381881425	0.3937496154790789
Accuracy	90.31%	89.53%

Identification of important Features

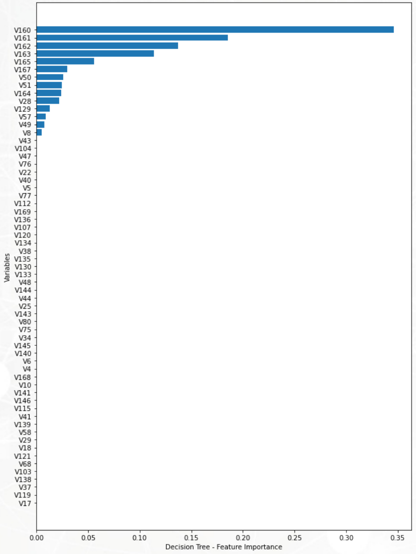
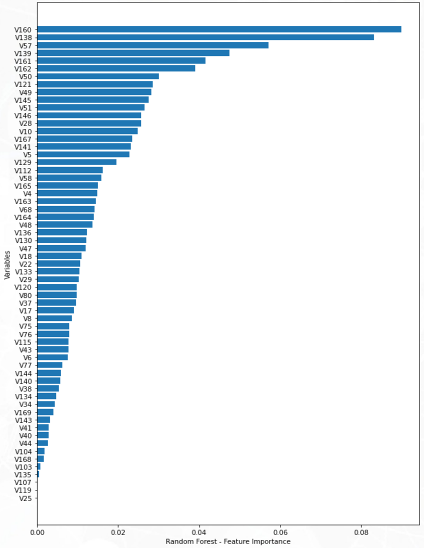
Random Forest Classifier

V4, V5, V6, V10, V38, V39, V40, V41, V42, V43, V44, V45, V46, V49, V50, V51, V58, V59, V103, V107, V112, V114, V115, V116, V117, V119, V120, V129, V136, V138, V139, V140, V141, V143, V144, V145, V146, V160, V161, V162, V163, V164, V165, V166, V167, V168, V169

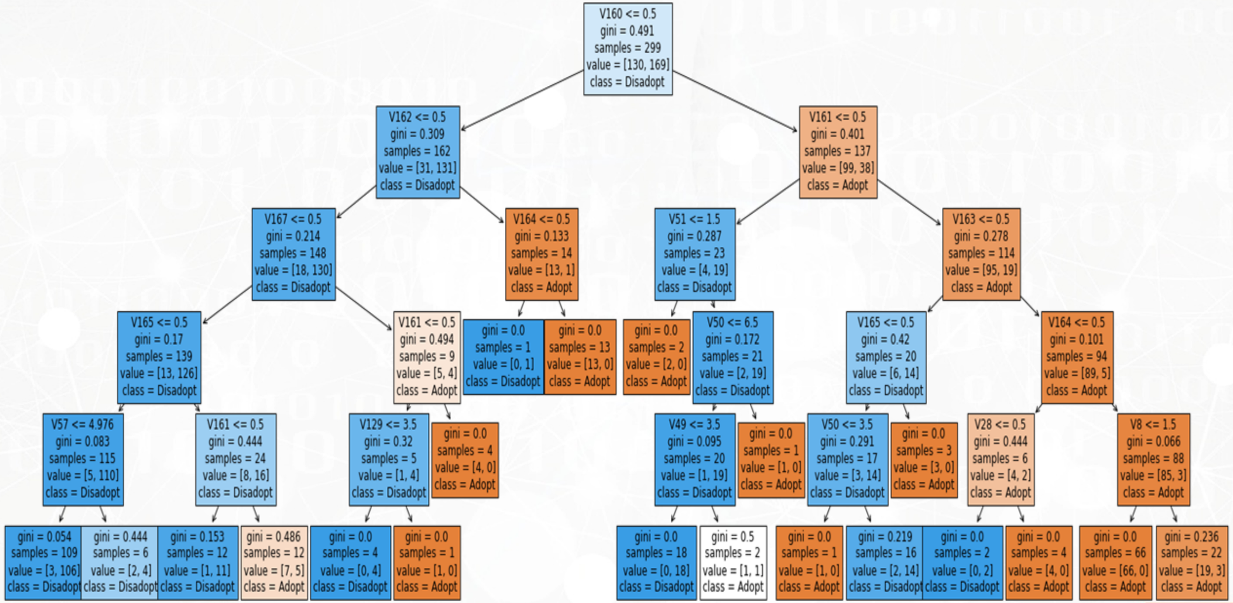
Decision Tree Classifier

V3, V28, V49, V50, V51, V57, V129, V160, V161, V162, V163, V164, V165, V167

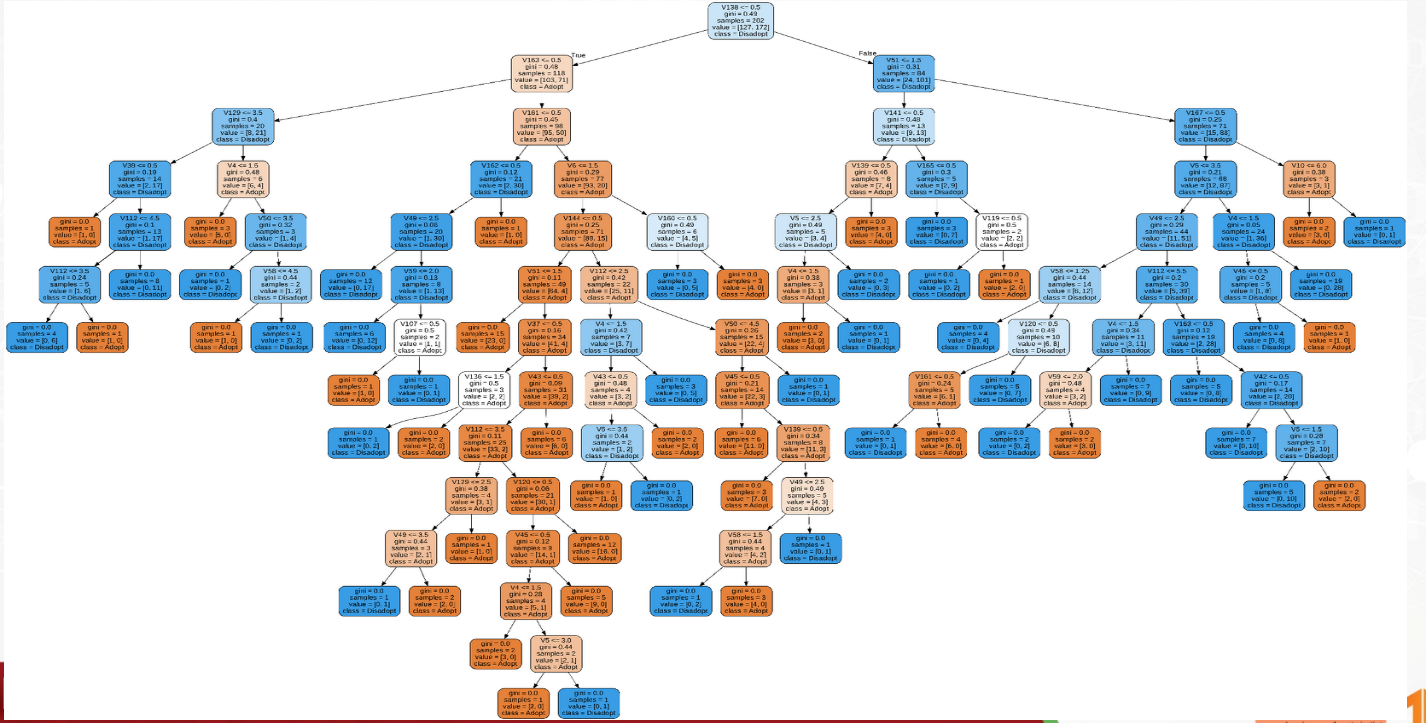
Visualizing important Features

Decision Tree Classifier	Random Forest Classifier
	

Decision Tree Visualization



Random Forest Visualization



Discussion / Implications

a) The Decision Tree and Random Forest Classifier Models could predict the Smallholder Adoption at 86.05 and 84.50 respectively

b) The Decision Tree Classifier Model predicted Smallholder CSA adoption using 14 variables while Random Forest Classifier Model used 29 Variables

c) The important Variables for Decision Tree Classifier Model are V160, V161, V162, V163, V164, V165, V167, V28, V49, V50, V51, V57, V129 and V8

Discussion / Implications

1) The study yielded 610 Variables. Decision Tree predicted Smallholder CSA adoption using 14 Variables while Random Forest used 47 Variables
2) Implication: If data is collected on the 14 variables, it is therefore possible to predict CSA adoption
3) Using ML Algorithms, it is now possible to identify suitable smallholder farmers for CSA adoption
4) ML should be mainstreamed in the deployment of CSA practices among smallholder farmers

Conclusions

- The challenge of high CSA dis-adoption rates among smallholder farmers in Kakamega County informed this study
- Using the random forest classifier and decision tree, this study identified the most important variables that influence smallholder farmers adoption and dis-adoption of CSA practices.
- With data on the following: CA practice (V160), SWC practice (V161), PPT practice (V162), Composting Practiced (V163), ISLM/ISFM Practiced (V165), Water Harvesting practice (V167), and the Farmer Category (V51) Precision (V57), Agroforestry practiced (V164), Household Monthly income (V129), Farming Experience (V49), Year of CSA Training (V50), Wheelbarrow owned (V28) and Farm Decision Making ability (V8), it is possible to predict smallholder farmers CSA adoption

Future Work / Directions

- This study tested 2 ML Models: Decision Tree and Random Forest Classifiers; it will be necessary for future studies to test additional models that require less data

- This Study was conducted in Kakamega County, future studies should test the deployment of CSA adoption through larger samples that cover bigger regions

- Future studies should model the adoption of CSA practices among livestock farmers, fish farmers, and other livelihoods.

THANK YOU!