Uganda Sanitation for Health Activity (USHA)

Artificial Intelligence and Machine Learning: An Alternative to Surveys to Determine Sanitation Service Levels

Context & Challenge

In 2021, USHA and Tetra Tech's Data, Analytics, and Technology (DAT) team explored the use of artificial intelligence and machine learning (AI/ML) applications to categorize images of newly constructed or upgraded toilets to determine if toilets met the WHO/UNICEF/Joint Monitoring Program (JMP) minimum standards for household sanitation services. Over 270,000 images of latrines (e.g., interface and superstructure) were collected during project implementation. To process and analyze these images, USHA applied a machine learning model to classify and analyze image-related data. This reduced the time required for manual analysis and enabled USHA to test if machine learning models could replace or complement the Activity's related survey work completed by field enumerators during baseline and endline surveys. Specifically, USHA sought to compare the accuracy of these models to the outputs from enumerator classification to determine if machine learning models can validate survey data from the field. Several measures of accuracy can help to determine if artificial intelligence can classify whether a latrine is "improved" or "unimproved" with a certain level of confidence.

Analysis: Latrine Classification with Machine Learning

This analysis utilized Lobe.ai, TensorFlow, and Python to complete the classification and evaluation of the images collected via USHA's MBSIA and CLTS+ baseline and endline surveys. Lobe.ai is an open-source artificial intelligence image analysis software developed by Microsoft and can be used which can generate image classification models. TensorFlow is a neural network platform that develops and deploys largescale machine learning models using classification models. Python is a programming language that instructs TensorFlow how to execute the model.

USHA MBSIA and CLTS+ Approaches

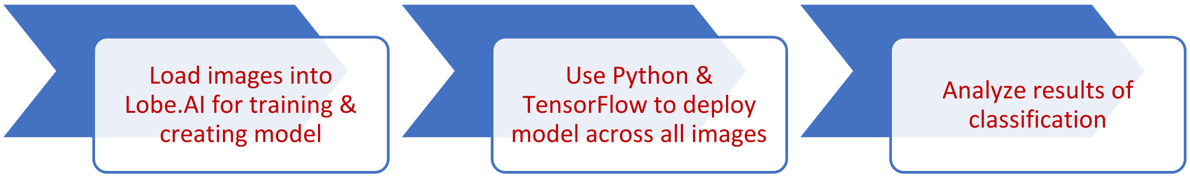
Market Based Sanitation Implementation Approach (MBSIA) interventions were aimed at creating a well-functioning sanitation enterprise – this included developing desirable and affordable products; marketing these products using tailored messages; facilitating a network delivery model to provide the required information, materials, and services to customers.

Community Led Total Sanitation Approach (CLTS+) interventions aimed to primarily end open defecation practices. This was done by delivering a range of affordable new and upgraded latrine products that respond to local latrine preferences and common construction practices using locally available materials.

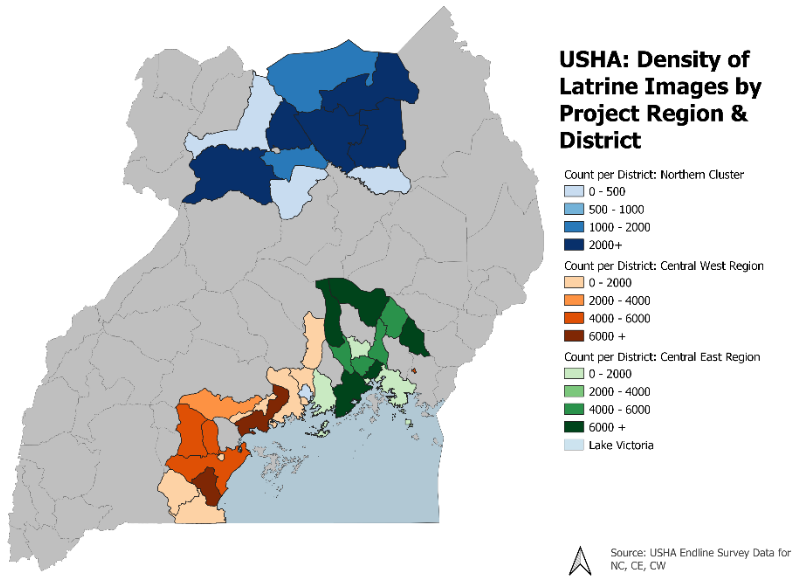
Machine learning uses training data to create predictive models which be applied data sets. In this analysis, the training data included several hundred images taken during intervention surveys which were loaded into Lobe.ai to create 3 distinct models: one model to classify the superstructure, a second model to assess if the latrine had a door, and a third model to classify the latrine's floor material.

Each model was based on unique labels i.e. "washable" or "unwashable" in the case of the floor classification. These models were then applied to the entire set of images to classify each latrine image based on labels.

Basic Workflow for Image Classification



Map 1 at right shows the density of images by district and region (NC, CE, CW). This shows not only the coverage of the USHA project, but also the measurable volume amount of image data leveraged for this study. While the machine learning process and subsequent analysis covered the baseline and endline modeling and classification, this learning brief will focus on the endline survey model results for the floor material. The material of the toilet "floor" or interface is one of two key defining features used globally to assess the level of household sanitation services. Toilets with washable interface materials are considered improved facilities.



Map I: Density of Images by District and Region based on GPS points in Endline Survey

Results: Regional Endline Survey Floor Classification

The endline analysis leveraged two different TensorFlow models for the floor classification: one for the Northern Cluster (NC) and one for the Central East/West Regions (CE, CW) (see Map 1 above). These models were trained differently based on what is considered appropriate flooring unique to each region. For example, in the NC, the categorization of floor types according to the JMP/WHO/UNICEF considers "slabs covered with a smooth layer of mortar, clay, or mud as improved."[1] In the NC, few interfaces are built with concrete or other washable materials, but the washable designation can include smooth clay which is commonly found in the NC. In CE and CW, latrines with smooth clay or mud floors would not be classified as improved but in the NC, where building practices are different and poverty levels are higher, these materials are commonly used. This is an important nuance that was captured in the revision to the original AI models for endline classification. The labels used to classify the floors in the two models are show in Table 1 below.

It was determined that due to these regional differences, two floor models would be developed and those models would be trained differently based on regional specificities. For the NC, this meant there

---

[1] Guidance for monitoring safely managed on-site sanitation (SMOSS) Draft prepared for Phase 2 pilots August 2022

needed to be a distinction between washable (smooth clay) and unwashable (dirt) latrine floor materials. The results of the analysis positive with images classified with high levels of accuracy in the model output (see Table I results below).

Table 1: Model Descriptions and Results			
Region	Northern Cluster	Central East	Central West
Model	Northern Cluster Floor Classification Model	Central East Floor Classification Model	Central West Floor Classification Model
Label	Smooth Washable Unwashable	Washable Unwashable	Washable Unwashable
High-level results:	**36%** of the total images are now categorized as washable. Among the latrines classified as washable in Northern Cluster, **81% of the images** had a high* level of accuracy.	**60%** of the total images are categorized as washable. Among the latrines classified as washable in Central East, **91% of the images** had a high level of accuracy.	**66%** of the total images are now categorized at washable. Among the latrines classified as washable in Central West, **84% of the images** had a high level of accuracy.

**Note: the model output provided a confidence level (high, medium, low, very low) for each image in the dataset. High is 95% confidence, medium is 85-95%, low is 50-85%, and very low is below 50% meaning that the model is 95% confident the classification is accurate for the images labeled as "high" confidence etc.*

Enumerator vs. Machine Learning Classification of Latrine Types

Another key element of this analysis was to compare enumerator's classifications from the toilet image selection and their field data to see how closely it these matched the model's classifications. The Activity's analysis found that 70% of the classifications were a match for Central East, 76% for Central West, and only 46% for Northern Cluster. The Central East and West region's results demonstrated that a model like the one developed for this analysis could be used in place of enumerators or to complement their efforts. For example, if beneficiaries sent an image of their latrine to a central digital location, this could decrease the resources and time needed to send agents to collect data. This demonstrates the value of machine learning and artificial intelligence in related efforts, particularly in hard-to-reach areas or conflict zones. With only 46% a match between the enumerator and AI classifications in the Northern Cluster, the model needs to be further refined or the method of classification by enumerators, clarified. As mentioned above, even latrines with a "smooth or clay" surface were considered washable in this zone. This distinction may need to be clarified with enumerators to ensure consistency for classification purposes.

Lessons Learned & Conclusions

Lessons Learned
ML and Image classifications work best with clearer high-resolution images of toilet floor types.
Enumerators should ensure that drop hole area of the latrine is not covered to avoid any obstructions of the ML models while classifying images.
The use of ML & Image classification models to classify latrine types can save time and resources required to administer long surveys.
Sanitation service surveys should further classify "unwashable floor" materials by durable smooth mud/mortar to ensure compliance with the JMP/WHO/UNICEF standards, especially in low-income communities such as Northern Uganda.

This analysis aimed to show how machine learning could be leveraged to help classify sanitation service ladders for large-scale sanitation and hygiene projects with accuracy and high levels of confidence. The

AI/ML model could be a beneficial tool for Implementing Partners, the Ministry of Water and Environment, and the Ministry of Health to determine and confirm sanitation service ladders using image classification. Further, sanitation and hygiene partners could explore using this AI/ML output to make data actionable while saving time and resources commonly used to administer long survey. Additional lessons learned are detailed below.

The TensorFlow models for the Floor, Superstructure, and Door are publicly available on the USHA Google Drive and the Python code and methodology is available on GitHub.[i]

---

[i] tt-ard-t4d/USHA_ML: USHA-ML-LatrineClassification