Global Partnership for Sustainable Development Data

Applying Machine Learning (ML) to Digital Health Programs Case Study

The purpose of this case study is to better understand GPSDD's contribution to and outcomes to-date of the Innovation Fund project led by Dimagi, applying machine learning to digital health programs.

To capture the dynamism of the collaborative process, the case studies are designed to explore outcomes at the various stages of the process, which can be broadly categorized as: 1. Catalyze connections between stakeholders; 2. Coordinate to understand the issue, develop, and prioritize solutions; 3. Impact decision-making by using, sharing, adapting, and re-using data.

This activity is considered to be in stage two. We hope to do a follow up case study to further explore and document outcomes from stage three.

This case study was written by Dimagi, the prime organization implementing this project and a 2016 Innovation Fund recipient. For this project, Dimagi partnered with a community health empowerment organization, which, for confidentiality reasons, will be referred to as implementing partner (IP) throughout this piece.

Data4Development Challenge

With Sub-Saharan Africa carrying a large proportion of the global burden of HIV, and survival rates attributed to antiretroviral treatment, retention in care of HIV positive patients is a priority for patient health and epidemic control. One of the biggest challenges in reducing transmission of HIV, TB, and other diseases is the treatment of defaulters, or patients who fail to return for treatment. Missing appointments or defaulting from care prevents effective treatment and has the financial consequence of draining regional budgets. Timely identification of patients at high risk of not returning is unlikely in the current system, where operational indicators are frequently collected by hand at the point-of-care. These are then aggregated over months, before being used to report the operation's status to funders and stakeholders, often with a significant time lag, such as a year later. Furthermore, data gathered is mostly for routine data collection and there is no culture of real-time data usage and application on the front line, as it is often seen as messy and incomplete.

Importantly, when patients default from care, or miss their clinical or antiretroviral pick up appointments, frontline workers[1] (FLWs) are often the direct point of contact responsible for re-engaging a patient in the return pathway to care. During patient touch-points within the care cycle, data collected on mobile devices by FLWs travel up to the reporting level to reflect relevant operational, programmatic, and patient level metrics. Often, this reporting follows a pathway separate to an FLW's day-to-day patient interactions and decision-making processes, and as a result may not be accessed by FLWs or their supervisors. This presents an opportunity to design a system to feed this information back to the front lines of care, specifically within the context of managing HIV defaulters. Furthermore, data analytical methods can be applied to the question of managing or prioritizing HIV defaulters for both programmatic and frontline worker applications.

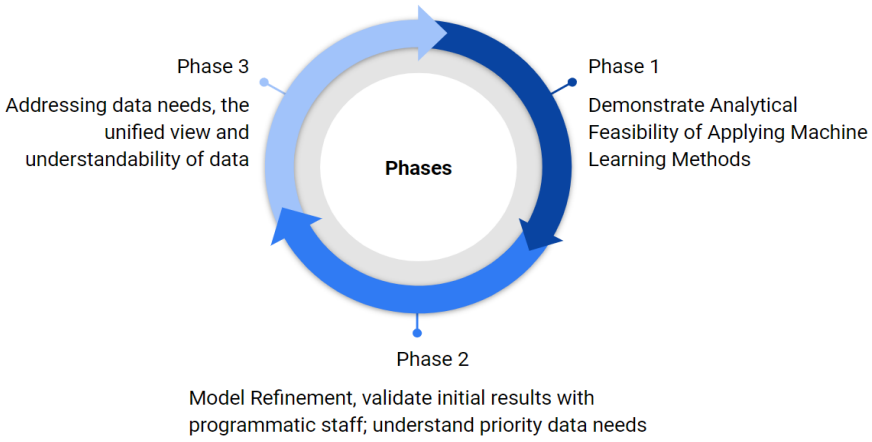
[1] Frontline workers (FLWs) connect communities and families to health systems for care delivery and services, playing a key role in preventing, treating and managing health conditions. Often, FLWs operate within low resource or hard to reach populations, therefore filling a delivery gap to provide critical services to the most vulnerable.

Response

Dimagi[2] is a digital health provider that partners with HIV programs to support clinical care pathways to frontline workers (FLWs) to manage HIV patients and embed service provision in communities. Dimagi received a $249,474.00 grant through the joint Global Partnership for Sustainable Development Data (GPSDD) and World Bank Innovation Fund to ask the following question: *"Within a digital health system, can we use Machine Learning to prioritize patients most at risk to default, and how would this be designed into the workflow of a frontline worker?"*

Dimagi's approach consisted of three key phases: 1. Demonstrating the feasibility of applying machine learning; 2. Refining the model through user testing; and 3. Curating a unified view data set developed from machine learning insights. The first phase required diving into the data to assess feasibility of machine learning as a method for defaulter tracing. The data came from CommCare HQ, a Dimagi-developed open source digital health platform that aggregates patient data from health systems. For this project, the data came from the application of the platform in 6 Sub-Saharan African countries. The data set contains a wide array of input and output variables relevant to defaulters and regional health system analysis. In our first project phase, we were able to pull historical data to look at defaulter trends over a multi-year time period.

GPSDD's contribution to the project beyond the funding involved staff time in monitoring and reviewing project deliverables and facilitating connections with other innovation fund projects largely through the lessons learned workshop on the sidelines of the Data for Development Festival in March 2018.



[2] Dimagi is an award-winning enterprise, created in 2002 out of Harvard and MIT. Dimagi has a team of 120+ engineers, scientists, agriculture and public health experts, and project implementation staff in offices in the United States (HQ), Senegal, India, Guatemala and South Africa. With its open source mobile platform CommCare, used in over 60 countries, the company helps address gaps in service delivery by frontline health workers in low and middle income countries.

Global Partnership
for Sustainable Development Data

Phase 1: Machine Learning Model Feasibility

Essential to a machine learning predictive model is defining a target, which is the outcome of interest that we would like to predict. In our case this is the event of **defaulting from care**. There is programmatic variation in the definition of defaulters, and Dimagi considered a range of definitions throughout the project cycle. For this initial phase, based on the implementing partner's (IPs) programmatic structure and CommCare data collection, the initial target we focused on considered defaulters as clients who had missed a scheduled appointment and were actively followed up by an FLW based at a facility. Here, we assigned a target value of 1 if a client is actively followed up and their outcome is documented, otherwise assigning the target value was 0.

During the feasibility testing phase, two types of analysis were conducted:

1. (Supervised) predictive modelling – essentially this means taking a portion of the data and asking a machine learning algorithm to *learn* what the underlying pattern to defaulting is in the program. This is traditionally known as a "classification" effort, such that when the algorithm is shown a **new** patient it can answer the question: ***"What is the probability this patient will default in the next defined time window, e.g. 90 days?"***

2. (Unsupervised) cluster analysis - Cluster analysis or clustering is the task of finding similar groupings of clients in the data set. Ideally this can help the program understand the ways that clients in the same group are more similar in outcomes and behavior to each other than to those in other groups. This hopefully would allow for program insights and understanding that normally could only be achieved with lengthy and expensive clinical trials and baseline studies. Whilst difficult to produce and having no guarantee of their being underlying groupings, the hope would be this approach would help answer the programmatic question: ***"Are there any particular groups in the data that are more vulnerable to poor outcomes?"***

The model performed well at the task of recognizing potential defaulters, demonstrating feasibility of machine learning methods on the data set. When tested with unknown patients, the algorithm correctly classified them as followed up by an FLW or not 3 times out of 4. Furthermore, when all patients were ranked in order of risk, 20% of the cases used in the model were responsible for 60% of those who were actively followed up on. The results for patients based on the model illustrated the potential to help a FLW focus efforts on the cases that are most at risk of falling out of the program. When the model was applied to a more restricted dataset, for just South Africa, the prediction was more accurate at 78%.

Phase 2: Model Refinement and Field Observations

In depth field work allowed us to bring the machine learning-developed risk score to the frontline health worker. This informed how a relevant user interface could be designed and be of utility to a frontline health worker. The outputs were the result of meetings with programmatic stakeholders, open ended discussions

with monitoring and evaluation staff, and prototyping and observation sessions with end users. As Dimagi, we aimed to understand our entry points in the data value chain, and how machine learning and risk prioritization fit into these levels. The feedback from the users was incorporated into the design of a second prototype.

Additionally, discussion with program leads allowed refinement of the machine learning model itself, which resulted in target creation that represented our population of patients more closely. During this time, we gained insight into how prioritization can be independently classified as both a machine learning risk score, and/or a prioritization cascade based on clinical and appointment patient data. FLWs can potentially be assisted by a combination of both these prioritization classifications. This phase demonstrated the value of combining both the analytical, data-driven approach with on-the-ground context and frontline health worker needs for a "best fit" solution. An important insight during this phase was understanding the data impact of application misuse and how that translates to model bias or error in interpretation. For example, if users are not consistently using the application to follow up on patients, the model will not accurately reflect the real world follow up of defaulters, which may be happening outside of the digital data collection system. To help address these biases, convergence of data collection tools is key to creating the optimal data structure for a machine learning model to run on, which leads us to Phase 3.

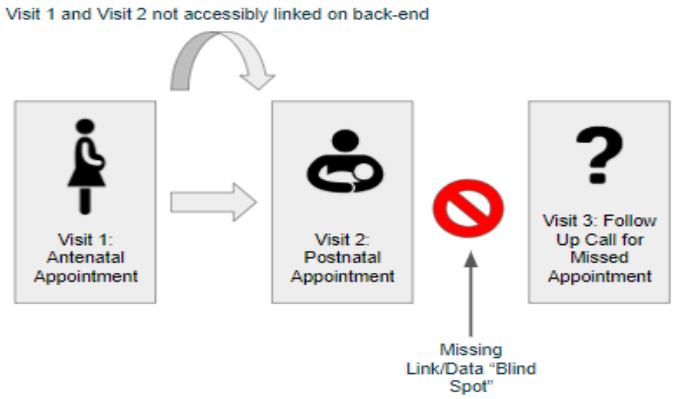
Phase 3: Understandability of data, the unified view as an outcome.

In creating machine learning models specifically applied to the FLW system, we identified an important challenge of being able to make sense of the raw data collected by CommCare. Even the Dimagi team and machine learning experts struggled to get a clear view of what was happening given the large amount of transactional data. We observed that the partner organization would benefit in their standard M&E activities by seeing a transformed view of the data over time, or a "unified view." The status quo data challenges for standard reporting included the following:

- Patient data about defaulting came from multiple appointment types
- These forms were not linked, therefore it was a challenge to observe patient history
- Aggregate statistics were used to inform defaulter metrics were based on monthly appointments from a single form type
- Insight into follow up of patients was limited, as well as patient outcomes after follow up

Consider an HIV positive pregnant mother who has 3 touchpoints with a facility system. She has an antenatal visit (AN), then gives birth and attends her postnatal visit (PN). She is expected to come back to the facility on a given day to do an ART drug pick up. She misses this appointment, and the FLW fills out a follow up form for her. In the current system, there is a data challenge around linking the AN-PN-missed appointment history of the patient.

To address these gaps, Dimagi developed an initial data prototype of a "unified view" that aimed to unpack nested objects to a more usable one-row-per-visit longitudinal view for multiple appointment types, linked to an individual patient. This longitudinal view was previously not visible to organization and was designed to provide a more efficient way to access patient visits over time. As a result, it enabled a retrospective view



of patient history in the system, previously unobserved by the IP. This design also focused on identifying key variables relevant to programming and reporting, which is an opportunity to develop a precise set of high-impact key analysis and reporting variables for an organization. By combining these data needs, the unified view was a mechanism to allow for pattern identification for missed appointments and system touch points.

Patient Name	Event	Appointment Type	Date	Outcome	Key Variable A
X	Visit 1: Antenatal		01-03-2018	Attended Appt	1
X	Visit 2: Postnatal		15-05-2018	Attended Appt	1
X	Visit 3: Missed ART Refill, follow up form filled out		01-04-2018	Missed Appt	0

The envisioned data users of this output included M&E users in regional country offices, statistical officers at facilities who would like to generate defaulter statistics, and supervisors who are interested in monthly or weekly appointment information of their patient base. Providing information in this format can improve programmatic insight into patient history that was previously unobserved.

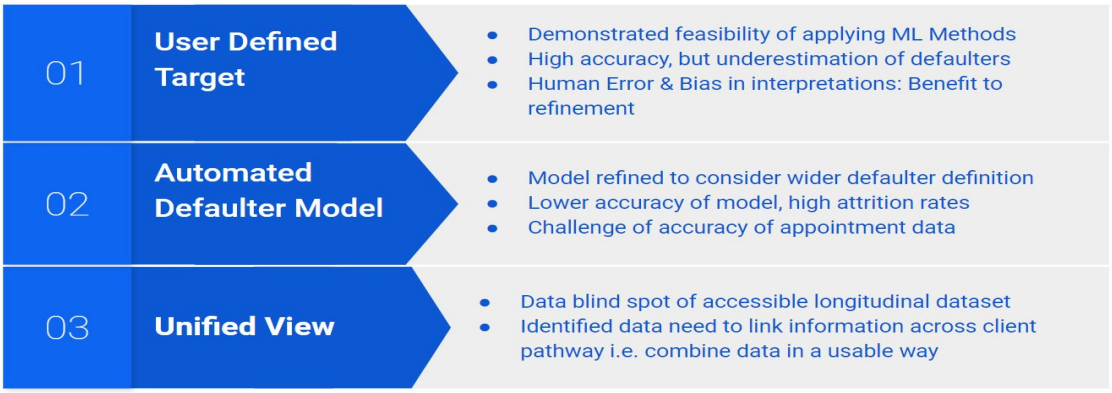
Challenge Identified: Data from touchpoints with mobile application not synthesised into a single coherent view

Solution: One-row-per-patient "Unified View" that gives a longitudinal view

Potential Impact: Improved programmatic insight into patient history previously unobserved

Global Partnership for Sustainable Development Data

Outcomes

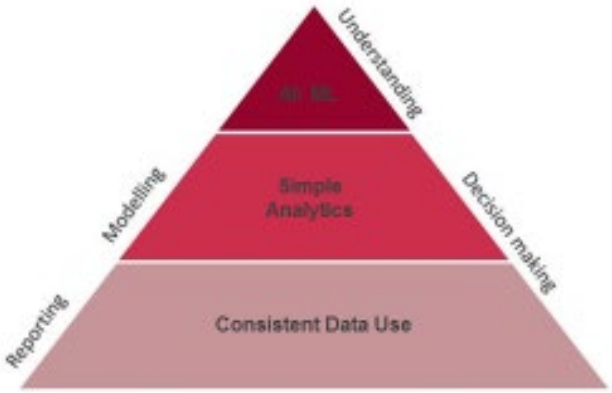


Data and Skills

This project demonstrates the feasibility of applying machine learning to predicting defaulter risk. The predictive machine learning models developed through this project have the potential to help front line health programs identify patients who are likely to fail to return for treatment (for diseases such as HIV and TB) by quantifying the underlying risk factors of defaulting from care. In addition, the unified view interface builds data-driven decision-making into an information system for care providers and equips the health workers with a decision-assisting job aid. This in turn will help health programs to take action on patients at high risk of not returning, and thus cut down on costs related to more resource demanding mitigation strategies, such as physically locating non-returning patients.

Knowledge and Resources

Partnering with GPSDD enabled us to have an experimental space to pursue the analytical route of applying machine learning to a pre-defined problem. This highlighted the need for Dimagi to dedicate time to clear and consistent use of data within an organization as a key process to laying the groundwork to doing more sophisticated analysis. Thanks to this collaboration, and the platform to innovate within the data value chain of our partner organizations and end users, it has assisted us to develop the following hierarchy of data use goals to further pursue the long-term goal of applying machine learning to support our partners' work.



This pyramid, often depicted in the journey to AI/ML for organizations, helps see the relationship between stages of improved data capacity and the expected outcomes. Even if data is being collected in real-time, there is foundational work required before the program can benefit from advanced analytics. Programs often must first invest in understanding the data and using it consistently, and then introduce more simple analytics, before really being able to harness artificial intelligence or machine learning algorithms. In addition to benefits that can be derived from this case study to predict and reduce defaulting for the implementing partner, Dimagi will additionally use the above framework in future projects with other partners to help set expectations and guide our collaboration with partners wanting to leverage machine learning to improve their programs.