Author : Tracey Li

Machine Learning for Personalized Care

How can we personalize health services so that clients receive the care that *they*, as individuals, need?

At D-tree, this question has become increasingly important as technologies improve and programs expand and scale up.

We are continually striving, alongside our government partners, to improve the quality and efficiency of health programs in order to have a positive impact on the lives of as many people as possible. Since governments have limited resources, we think carefully about how to allocate resources in the best way, so that health service delivery will result in the best outcomes for as many clients as possible.

Personalized care leads to better outcomes for each client, as they receive the services that will most benefit their health. It also leads to improved allocation of resources, because health workers only deliver the services that each individual needs, rather than delivering a generic package of services that may include content which is not relevant to every individual.

To help us with this task, we have started to explore the use of machine learning.

What is machine learning?

Machine learning is a technique that enables a computer to automatically identify patterns in a dataset, without being given explicit instructions about what to look for. The amount of human input that is required is therefore low, compared to ‘manual’ methods of data analysis. Machine learning is valuable when the dataset that needs to be analyzed is too large or complex to be able to efficiently perform the analysis using traditional tools.

There are several different types of machine learning. We are currently focusing on one type, called *predictive modeling*. This involves developing a model that can predict a chosen ‘target outcome’ from the input data that we provide. For example, we could choose our target outcome to be whether a woman will have a successful delivery (either ‘yes’ or ‘no’). The result would tell us whether she would benefit from receiving special services, in order to improve her chances of a successful delivery. The prediction would be based on input information about her health and socioeconomic circumstances. We could also include readily available inputs about environmental factors, such as the population of her village, which can indirectly provide information about the level of infrastructure that exists around her home.

How do you do it?

To create a predictive model, we first need to ‘train’ the model with a historic dataset, that includes both the input data as well as the target outcome that actually occurred. The model ‘learns’ from this dataset, meaning that it finds patterns that link the input data with the target outcome. We then use these patterns to predict the target outcomes for our current clients, pregnant women, for whom we have the input data but who have not yet given birth. If the model predicts that a woman will have an unsuccessful delivery, this means that she has similar characteristics to women in the historical dataset whose deliveries were unsuccessful. We can then provide her with additional services to improve her chances of a successful delivery.

Why is machine learning (predictive modeling) a good tool for us?

An essential requirement for building a good predictive model is having a high-quality dataset to train the model. Due to D-tree's digital platform, and the success of multi-year programs such as our Safer Deliveries program which includes longitudinal data for tens of thousands of previous clients in the program, we already have such datasets. Improving the quality of our health service delivery is one of the reasons why we collect the data in the first place, and we are now using machine learning techniques to extract insights from the data that cannot be done using traditional analysis techniques.

How we use machine learning effectively

D-tree's strength lies in our technical expertise combined with our ability to implement. We deploy technologies in a manner that is appropriate to the context, culture, and constraints of the location in which we are working, in close collaboration with governments and public health experts. In the case of a machine learning model, we take special care to address the following considerations:

- The prediction step needs to be able to be **conducted offline**. This is to enable community health volunteers (CHVs), even in the most remote settings, to immediately assess a client's risk level during a visit, whilst the CHV is inputting data about the client into the app.
- Machine learning models are not perfect and will not make predictions that are 100 percent accurate. Since there will always be some errors, it is important to understand the consequences of these errors, and to **mitigate any risks** that may arise as a result of them. In our case, it is more harmful for the model to incorrectly predict that a woman is 'low-risk' when in reality she is high-risk (this is called a 'false negative') than it is for the model to incorrectly predict that a woman is 'high-risk' (a 'false positive'). We take this into account by training the model to minimize the false negative rate.
- Machine learning models can tell us about who is likely to be at high risk of a negative outcome, but they often do not provide much information about *why* this is the case. Therefore, if we use the output of the machine learning model as our only source of information, the intervention that we provide in response to that information can only be fairly generic, and address the reasons that are most likely for everyone in general. In order to be most effective at an individual level, machine learning needs to be **coupled with public health expertise** in order to understand what is causing the increased risk, and not just whether there is an increased risk.
- Machine learning is an unfamiliar tool to many of our stakeholders. We therefore need to make efforts to **communicate what we are doing and why**. We also need to provide information that enables non-technical experts to understand how machine learning works, and what its benefits

and constraints are. By doing so, we will build confidence in the use of machine learning and its outputs.

The future of machine learning at D-tree

We are still in the early stages of using machine learning at D-tree. We have recently deployed a machine learning model produced by our collaborators at N-Lab at the University of Nottingham that predicts whether a woman is at risk of her baby dying during delivery. Women that are identified to be high-risk are provided with additional services, beyond our standard package of pregnancy services.

After the initial pilot phase, we will evaluate the effectiveness of the initiative and reflect on what we have learnt during the process, before deciding how to proceed. In the meantime, we are also thinking about how machine learning could potentially help us to improve other aspects of our program, such as improving training for CHVs, or providing automated communications to clients, in between CHV visits.

We see a huge opportunity to leverage the power of the datasets that have been collected through our digital platform, to ensure that all our clients receive tailored care that meets their needs.

—

The work to deploy a machine learning model to create tailored care pathways for clients at higher risk of perinatal mortality in Zanzibar is made possible by support from the Wehubit program implemented by Enabel.