Prediction and Prevention in Primary Schools

Using Scorecards to Identify Students at Risk of Repetition & Dropout

Rwanda has made drastic improvements in primary education, achieving near universal enrollment. However, high rates of repetition and dropout continue to hold back the education system and the economy. Predictive analysis offers the ability to better preempt repetition and dropout.

Using national survey data, we develop a 10-question scorecard that can be used by schools to predict which students are at highest risk of dropout or repetition. **Identifying high risk students will allow MINEDUC to shift to "prevention" rather than "response" and better target resources in the process.**

This policy brief details the depth of current dropout and repetition problems, our predictive approach, the results of our analysis, and lays out a road map to implement our finding.

Key Recommendations:

1. Use predictive scorecards to identify pupils at risk of dropout or repetition.
2. Combine scorecards with Imihigo contracts & new prevention programs.
3. Pilot and evaluate different approaches before scaling nationally.

Why Repetition and Dropout Matters

High levels of repetition in early primary grades lead to crowded classrooms filled with students of all different ages. The graph below shows the large degree of "bunching" in P1 to P3 and the low enrollment numbers from P6 onward.

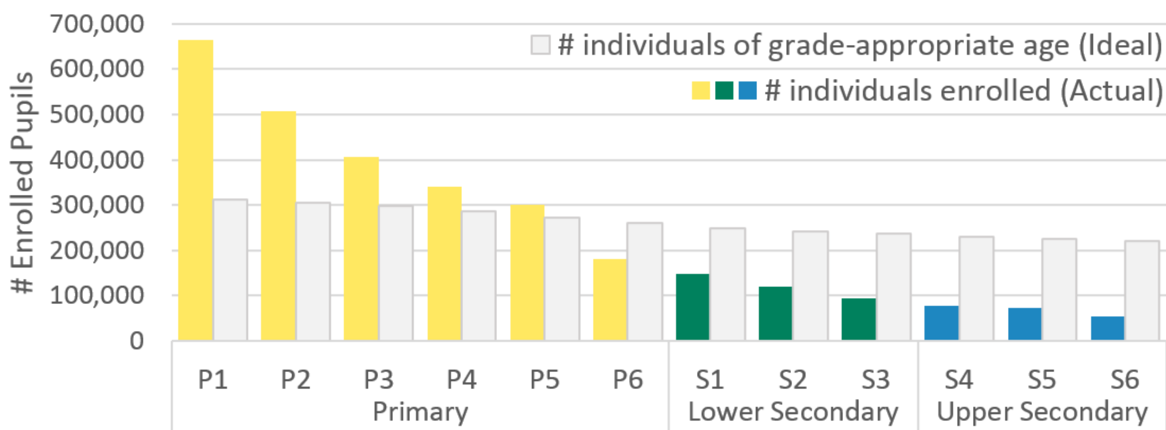
On average, students take 8 years to complete 6 grades. Aggregated across the education system, the cost of repetition in budget terms is equal to schooling 133% of the existing primary population.

Dropout:
Less than 70% of students who start primary reach P6

Repetition:
The students who reach P6 repeat 2 grades on average to get there

Additionally, the small number of student who reach secondary school limit the skills of the labor force overall, restricting the government's ability to achieve ambitious goals like Vision 2020.

A Visual Story of Repetition and Dropout: Actual enrolment versus Ideal on-track enrolment



A Way Forward: Using Prediction to Identify At-Risk Students

"Machine learning algorithms" are a set of data analysis techniques that can be used to predict outcomes in the future based on past data. They are used to solved many different policy problems in developed countries from predicting who will default on a loan to predicting who will drop out of school. We adapt these methodologies to the resource constraints of schools in Rwanda by:

1) **Collect Data:** start with 200 variables that may be used to predict drop out or repetition
2) **Select the 10 Best Questions and Create a Scorecard:** use machine learning to narrow down the 200 variables to create a 10-question scorecard that fits on a sheet of paper
3) **Build the Scorecard into Existing and New Programs:** use scorecards to help teachers meet new Imihigo goals or better target programs focused on dropout and repetition

*"Accurate prediction allows schools to **identify and focus** on the highest risk students **before** they dropout or repeat"*

Predicting the Future

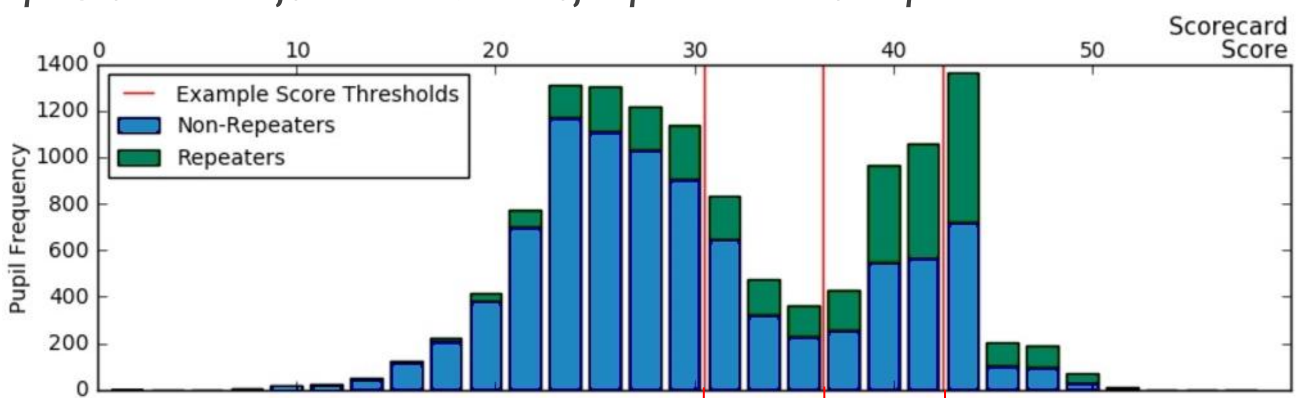
Applying machine learning techniques to data allows us to create scorecards that accurately predicts a pupil's dropout or repetition risk. We can identify 2 out of 3 at-risk pupils, while only misclassifying 1 out of 3 non-at risk pupils.

The Scorecard:

Behind the Scenes

A **10-question** scorecard gives a **score from 0 to 64,** with higher scoring kids having higher risk. The graph below shows that on average the green repeating pupils have higher scores than the blue non-repeating ones.

Example Scorecard Performance: Number of Repeaters and Non-repeaters at Predicted Risk Levels



		Threshold of 31		Threshold of 37		Threshold of 43	
		Prediction		Prediction		Prediction	
100 Students ▼		at-risk	**not** at-risk	at-risk	**not** at-risk	at-risk	**not** at-risk
26 Repeaters	▶	19	7	15	11	7	19
74 **Non**-repeaters	▶	28	46	19	55	8	66
100 Students	▶	*47*	*53*	*34*	*66*	*15*	*85*

Unique Rwandan Solutions: Imihigo

Imihigo contracts can provide teachers with incentives to act to prevent repetition and dropout of at-risk pupils. Scorecard results allow teachers to focus on those pupils most at-risk.

Targeted programs

Channeling at-risk pupils into preventive and remedial programs can lower the dropout or repetition rate. Scorecard predictions allow better program targeting, assuring that resources go to the most at-risk students.

An important **trade-off** lies in setting the threshold for at-risk classification: a low threshold identifies most repeaters as at-risk, but also miss-identifies several non-repeating pupils as at-risk. The opposite is true for a higher threshold.

Example **scorecard predictors**, are age and grade, questions about learning outcomes like "*Can the student perform a written calculation?*", questions on non-school activities like "*Has the student fetched water in the last 7 days?*", and other questions, for example "*Does the student own a mobile phone?*".

Roadmap to Implementation:
From Scorecard to Improved Promotion and Retention

1) Finalize Algorithm and Scorecard with Partners

Due to data limitations, our current algorithm is designed only to predict repetition. With new data, available in April 2017 from MINEDUC's partner Laterite, this analysis can be extended to (1) improve the existing scorecard and (2) design a scorecard for dropout.

This research can be continually supported as a part of Laterite's Dropout Study in partnership with UNICEF and DFID. This process will give MINEDUC opportunity to provide input into a final scorecard and processes to pilot its use. We propose two different uses for scorecards below

2) Implementing Scorecards with:

A) New Imihigo Contracts

- Targets for improving repetition and dropout should be added to Head Teacher level Imihigo contracts, setting goals at the school level.
- Scorecards can be introduced alongside this change as a means to help teachers target their efforts and reach their goals.

B) New, Targeted Programs

- Laterite, UNICEF, and other partner organizations are prepared to support MINEDUC in the design of innovative programs that accompany scorecards.
- Such programs should address root causes identified by Laterite's dropout and repetition study.

3) Pilot Programs, Evaluate, and Scale Successes

This novel use of scorecards to address dropout and repetition in public schools is new in Sub-Saharan Africa. As such, pilots of these programs, comparing the effectiveness of (A) *scorecards + Imihigo* and (B) *scorecards + programs*, should be rigorously evaluated before scaling up. Digital data collection and scorecards can eventually replace manual methods.

Future Correspondence for Authors: Ben - BenBrockman314@gmail.com & Vincent - Vincent.Vanderputten@gmail.com
In partial fulfilment of the requirements of the MPA/ID program at the Harvard Kennedy School. The authors extend a special thank you to MINEDUC Rwanda, Laterite Rwanda, and the Harvard Kennedy School for supporting this research. Full report accompanying this brief available upon request. Cover photo: courtesy of Laterite.