18. *Predicting School Performance Using a Combination of Traditional and Non-Traditional Education Data from South Africa* (Henry Wandera, Vukosi Marivate, and David Moinina Sengeh)

The application of Big Data Analytics (BDA) in education is transforming learning, teaching and administration in schools. Current Education Data Mining (EDM) research focuses on teaching and personalised learning in higher institutions mostly in western countries with limited research conducted in African countries. Most research has been conducted using small datasets, simple learning analytics techniques and machine learning black box models to predict students' performance. Black box modelling approaches use complex structures which are difficult to be easily interpreted by stakeholders. EDM approaches and tree based machine learning techniques were synthesised to identify important features that can predict school performance across African countries such as South Africa. LightGBM a gradient boosting framework and interpretable tree based algorithms was applied on combined data sources from community surveys, school master lists and examination results to perform feature importance. The challenge faced in EDM research is limited education data sources; different existing datasets from government reports and archives were merged. Community survey data was used to determine the standards of living in secondary schools within those communities. Cell phone internet, toilets, security, usable water sources, number of teachers and students, school location, and family head were identified as control variables impacting the attainment of schools. LightGBM, underlies the developed prediction model. It empowered the model with high accuracy, stability and easy interpretation hence outperforming XGBoost, decision tree and random forest algorithms. Index Terms—performance prediction, school performance, national statistics.

The full paper is available on the following link:

Authors:

- Mr Henry Wandera is an Assistant Lecturer at University of Pretoria (UP) (u17253129@tuks.co.za).
- Dr Vukosi Marivate is a Senior Lecturer at UP (vukosi.marivate@cs.up.ac.za).
- Dr David Moinina Sengeh is the Adviser to the President of Sierra Leone; Chief Innovation Officer in the Directorate of Science, Technology and Innovation in the Government of Sierra Leone (dsengeh@statehouse.gov.sl).