A Machine-Learning Approach to Nowcast the GDP in Sub-Saharan Africa

AFRICAN ECONOMIC CONFERENCE
DECEMBER 10TH, 2020

Presenters: Tara Iyer and Franck Ouattara

Views expressed herein are those of the authors and do not necessarily reflect the views of the IMF, its management, or its Executive Board.
Why nowcasting matters?

- Nowcasting is the prediction of the present.
- Monetary policy decisions in real time are based on assessments of current and future economic conditions using incomplete data.
- Macroeconomic analysis and estimating economic activity present distinct challenges in Sub-Saharan Africa, in particular during the COVID-19 pandemic.
Three pillars of our nowcasting project

1. Nowcasting SSA GDP

- SSA analyzed as an aggregate economy (SSA represented by about 10 largest countries)
- Updated regularly for information of the IMF staff
- Also reported in IMF Sub-Saharan Africa Regional Economic Outlook.
Phases of machine learning modeling

**Input** (predictors) and **Output** (GDP) data fed into the model

- **Tuning**
  - Set parameters of each model to maximize that model’s *likely* out-of-sample performance using cross validation in the training set.

- **Evaluation**
  - Gauge *actual* out-of-sample performance to choose the top models in the test set.

- **Nowcast**
  - Use the chosen models to nowcast GDP growth.
Machine-learning models considered in this project

- **Models used for “horse racing”:** Support Vector Machine (SVM), Random Forest, Elastic Net, Ensemble, …

- **SVM:** Algorithm that locates hyperplanes that differentiate among data points plotted on an n-dimensional space.

- **Random Forest:** Combines a large number of decision trees to approximate a non-linear regression.

- **Elastic Net:** Penalized regression model with coefficient shrinkage, trading some bias to improve out-of-sample performance.

- **Ensemble:** Meta-algorithm that combines decisions from different machine learning techniques.
Three pillars of our nowcasting project

1. Nowcasting SSA GDP

2. Nowcasting GDPs in individual countries
   - Provided for IMF country teams’ references for surveillance
   - Plan to complement with factor model approach

2. Nowcasting GDPs in individual countries (whose quarterly GDPs are available)
   - Discuss the result with the IMF country teams and reflect feedback to improve
Three pillars of our nowcasting project

1. Nowcasting SSA GDP
2. Nowcasting GDPs in individual countries
3. Working on countries with limited data (work in progress)

- Several SSA countries do not have quarterly GDPs
- Panel approach pursued
- Availability of predictors limited, but nontraditional data could be used
Sub-Saharan Africa’s GDP appears to have bottomed out

Sub-Saharan Africa: Year-on-Year Rolling Quarterly Real GDP Growth, Data and Nowcasts

Sub-Saharan Africa: Annualized Quarter-on-Quarter Rolling Real GDP Growth, Data and Nowcasts

Sources: Haver; IMF internal databases; and IMF staff calculations.
This reflects the evolution of high-frequency indicators including...

South Africa: **New Vehicles Sold** and Real GDP Growth

Nigeria: **Purchasing Managers’ Index** and Real GDP Growth

WTI Crude Oil Spot Market Price and Sub-Saharan Africa’s Real GDP Growth

Note: Purchasing managers’ index implies an expansion if >50; and a contraction if <50.
Sources: Haver; and IMF staff calculations.
While not an input to our nowcast, COVID-19-related data provide insights

Selected Economies: Daily New COVID-19 Cases

Note: New COVID-19 cases show the total across countries (max=100).

Sources: Center for Systems Science and Engineering, Johns Hopkins University; Google LLC, “Google COVID-19 Community Mobility Reports”; and IMF staff calculations.
Thank You
Extra Slides
Frequently Asked Questions

1. How do we know machine learning will provide better projections than country team’s own? Nowcasts are not to replace country team’s projections. A model-based counterpart is to be provided to the more routine forecasts produced by staff, which have traditionally been based on expert knowledge.

2. Will nowcasting have value added for smaller economies which do not have much data? We are starting from larger economies as they tend to have more data and we could accumulate know-how. Simultaneously, we will explore how to use non-traditional data. By incorporating the know-how and the available non-traditional data, we plan to work to provide some useful results. Growth projections in other countries could also be a valuable input. The literature is expanding rapidly.

3. Can we put all data into machine learning rather than selecting? In principle, we can do that, but not all indicators are available before GDP data are available. We should select variables given the availability. The model can then drop the predictors less useful.

4. Why is a nowcast for Country X on quarter Y is so much different from the actual observation? Nowcasting is not a magic box in terms of projections. While the numbers may different between a nowcast and the actual observation, the nowcasting exercise can provide valuable information for surveillance. Direction of growth (i.e., quarter Y’s nowcast relative to one quarter before), the contribution of predictors to the result, etc.
Nowcasting is the prediction of the present

- Nowcasting models employ data that are available at the same or higher frequencies than the variables of interest.
- If quarterly GDP data are available with a lag, how to monitor the economy in the absence of data?
  → Solution: Nowcast.
Machine learning is an extension of non-parametric statistics

- Focuses primarily on out-of-sample performance
- Able to handle a large amount of predictors
- Well suited to nowcast GDP:
  - Program the “machine” to exploit historical (possibly non-linear) statistical patterns, and predict current GDP growth
Introduction to Machine Learning

Permutation testing set?

Permutation 2.
Average
Error = 1.1
Introduction to Machine Learning
“Shapley” Decomposition: Kenya, Random Forest

- Accounting for nonlinearities, fairly attribute contributions to an individual nowcast.
- Which features pushed our nowcast away from the sample average?
- For example, lower-than-average CPI pushed our nowcast up by 0.16.
“LIME” Decomposition: Kenya, Random Forest

- Linearize model around individual nowcast; conduct sensitivity analysis
- Which features impact the prediction most?
- For example, if CPI were to drop to zero, growth nowcast would be 0.8 higher.

**Local Linear Model: Key Features**

- CPI.average (5.0)
- Manufacturing.Softdrinks (0.0)
- Price.of.rent.Crude.Oil (0.0)
- Agriculture.Horticulture (0.0)
- Exchange.Rate.Average (0.57)
- Manufacturing.Sugar (4.67)
- Agriculture.Horticulture (8.11)
- Price.of.Brent.Crude.Oil (-12.66)
- Manufacturing.Softdrinks (17.02)
- CPI.average (5.93)