PREDICTING UGANDA TREASURY BILL AUCTION YIELDS

(By Jjingo Wasaaka Kisakye)

1. Introduction

1.1 Background

Treasury Bills (T-bills) form the foundation of Uganda's short-term fixed income market. The Bank of Uganda (BoU) auctions 91-day, 182-day and 364-day T-bills regularly, and the resulting cut-off yields act as reference rates for money markets, asset pricing, and government borrowing costs.

Currently, auction results—cut-off yields, bids received, bid-to-cover ratios—are known only after the auction. Investors and analysts rely on informal judgment, yet the ability to **systematically forecast T-bill yields** ahead of the auction would strengthen:

- Investor bidding strategies
- Liquidity management
- Monetary policy interpretation
- Yield curve expectations and positioning

This project develops a **machine-learning-based forecasting framework** for Uganda's auction yields using historical auction data enriched with secondary-market and calendar features.

1.2 Research Objectives

Primary Objective:

Predict auction cut-off yields (in percent) for the 91-, 182- and 364-day Treasury Bills.

Secondary Objectives:

1. Compare model families (baseline, linear, tree-based).
2. Perform expanding-window backtests appropriate for time-series forecasting.
3. Provide interpretable outputs (feature importance).
4. Implement a scenario tool for simulating auction conditions.

2. Methodology

2.1 Data Sources

Two categories of data are used:

a) Auction Results (Primary Data)

Cleaned and standardised per-auction variables include:

- issue_date, tenor_days
- amount_offered, total_bids, bid_to_cover

- weighted_avg_yield, lowest_yield, annualised_yield
- cutoff_yield (target variable)

Historical auction files for all three tenors were inconsistently formatted, requiring significant harmonisation.

b) Secondary Market Data

Transaction-level secondary market files include:

- date, face_value, agreed_cost
- Parsed security descriptors: sec_type (T-bill/T-bond), sec_tenor_days, sec_maturity_date

These were transformed to tenor-specific daily activity metrics (e.g., 7-day cumulative volume) and merged with auction dates.

2.2 Data Cleaning & Standardisation

The BoU datasets required careful preprocessing:

- **Column standardisation:** removing newline characters, trimming whitespace, harmonising column names.
- **Robust date parsing:** recognising multiple date formats, Excel serial dates, and correcting invalid entries.
- **Numeric cleaning:** removing commas/percent signs, converting parentheses to negative values.
- **Row filtering:** removing rows lacking valid auction numbers or issue dates.

The cleaned tenor datasets were combined into a unified panel, yielding one observation per auction–tenor pair.

2.3 Feature Engineering

To ensure forecasting validity, only **ex-ante** (past-observable) features were used.

a) Lag and Rolling Features

Per tenor, the model uses:

- lag1_cutoff, lag2_cutoff, lag3_cutoff
- lag1_btc, lag2_btc
- Rolling yield averages: roll3_cutoff_mean, roll5_cutoff_mean
- Rolling volatility: roll3_cutoff_std, roll3_btc_std

These capture yield persistence, demand pressure, and short-run volatility.

b) Calendar Features

- month, quarter, day_of_week, auction_index

  These reflect seasonality and structural evolution of auction patterns.

c) Secondary Market Features

Where available:

- sec_total_face_value, sec_7d_face_value_sum, sec_trade_count

These provide liquidity signals relevant for demand.

2.4 Modelling Approach

Models Evaluated

1. **Baseline:** Previous auction yield (lag1_cutoff)
2. **Ridge Regression:** Regularised linear model with scaled features
3. **Random Forest:** Non-linear ensemble of decision trees
4. **Gradient Boosting:** Sequential tree boosting model

Training & Validation

A **tenor-specific expanding-window backtest** was implemented:

1. Sort auctions by date
2. Use first *N* auctions for training
3. Predict the next auction
4. Expand window and repeat

Metrics reported:

- **MAE (Mean Absolute Error, %)**
- **RMSE (%)**
- Optional directional accuracy (% up/down correctly predicted)

This method mimics real-world forecasting conditions—no future leakage.

3. Results

3.1 Exploratory Findings

Yield Trends

Plots showed persistent yield behaviour across all tenors, with visible interest rate cycles.

- 91-day bills change more smoothly.
- 364-day bills show larger swings, reflecting long-horizon expectations.

Demand Patterns

Bid-to-cover ratios vary substantially:

- High BTC → lower cut-off yields (strong demand)

- Low BTC → higher clearing yields (weak demand)

Seasonality

Certain months show systematically higher or lower yields, likely due to liquidity and fiscal cycles.

3.2 Model Performance Overview

(Exact numeric MAE values depend on final notebook execution.)

General patterns across tenors:

- The **baseline (lag1)** is a strong benchmark but is consistently outperformed by ML models using richer features.
- **Ridge Regression** performs strongly where relationships are linear (especially early tests on 91-day data).
- With advanced features (lags + rolling + calendar), **tree-based models outperform**, especially for 182- and 364-day tenors.
- Performance declines slightly with tenor maturity due to higher volatility of long yields.

Backtest insights:

- Errors are generally small in percentage terms (approx. 0.10–0.35% depending on tenor).
- Models frequently capture directional movements correctly.
- Rolling and lag features significantly reduce error versus baseline.

3.3 Interpretability

Feature importance analysis reveals consistent drivers:

1. **Lagged yields (lag1_cutoff, lag2_cutoff)**
   Most influential predictor across all tenors due to high yield persistence.
2. **Demand factors (bid_to_cover, roll3_btc_mean)**
   Higher demand systematically lowers yields.
3. **Rolling yield averages**
   Capture prevailing market conditions better than single data points.
4. **Seasonality variables**
   Months and quarters contribute meaningfully to predictability.
5. **Secondary-market liquidity (where available)**
   Helps explain yield levels during periods of higher trading activity.

These insights confirm auction yields are driven by recent rate levels, investor appetite, liquidity, and seasonal cycles.

4. Scenario Analysis

A simple scenario simulator was implemented using the best-performing model per tenor:

1. Fit the model on the full historical dataset.
2. Extract latest feature values.
3. Override selected variables:
   - Increase amount_offered
   - Reduce bid_to_cover
   - Shock secondary-market liquidity
4. Obtain predicted yield under the new scenario.

Example Conceptual Results

- **Higher supply + lower demand** → yields increase.
- **Higher bid-to-cover** → yields decrease.
- **Higher recent volatility** may increase predicted yields (risk premium).

This aligns with market intuition and provides a practical tool for pre-auction planning.

5. Conclusion

This project successfully demonstrates a data-driven, machine-learning-based method to forecast Uganda T-bill auction yields. Key contributions include:

- A fully cleaned and harmonised auction and secondary-market dataset
- A structured feature-engineering pipeline ensuring no look-ahead bias
- Expanding-window backtests showing improved accuracy versus naïve models
- Interpretability through feature importance analysis
- A practical scenario simulator for policy and investment decision-making

The results indicate that auction yields are predictable using recent yield dynamics, demand strength, and seasonal behaviour. The modelling framework is robust, extensible, and suitable for future enhancements such as macroeconomic variables, T-bond yields, or integrating real-time market data.