

# INVESTIGATING THE CAUSES OF JAMB FAILURE RATES AND PREDICTING FUTURE PERFORMANCE TRENDS (2020–2030)

## 1. RESEARCH BACKGROUND

The Joint Admissions and Matriculation Board (JAMB) Unified Tertiary Matriculation Examination (UTME) is the primary standardized test for university admissions in Nigeria. Over recent years, stakeholders have raised concerns over declining performance and increasing failure rates among candidates. This trend threatens access to higher education, especially among candidates from under-resourced regions or socioeconomic backgrounds.

This research project was initiated to investigate the root causes behind the rising failure rates from 2020 to 2025 and to forecast future performance trends through 2030. The goal was to derive actionable insights using real data collected from students, apply machine learning models to predict outcomes, and suggest evidence-based interventions for stakeholders including students, educators, and policymakers.

## 2. SURVEY METHODOLOGY AND DESIGN

The project employed a structured survey approach to collect primary data from candidates who sat for the JAMB UTME between 2020 and 2025. The survey was administered online via Google Forms and shared across student networks, schools, and exam preparation forums.

### Survey Design Highlights:

- **Demographic Information:** Age, gender, school type (public/private), state.
- **Exam Attempts:** Year(s) the candidate sat for JAMB, and the highest score attained.
- **Preparation Methods:** Study hours per day, types of materials used (books, online resources, private tutoring).
- **Digital Readiness:** Frequency of CBT practice, internet access, and access to devices.
- **Self-assessment:** Confidence level before the exam, perception of how well the score reflects their ability, and reasons for perceived underperformance.

### Sampling Frame:

A total of 1,000+ valid responses were collected and processed. Responses were cleaned to ensure consistency, remove nulls, and standardize data entries for analysis.

## 3. EXPLORATORY DATA ANALYSIS (EDA)

EDA was performed using Python (Pandas, Matplotlib, Seaborn). Key findings are summarized below:

**Score Distribution:**

- Majority of students scored between 180 and 250.
- High performance (260+) was rare; long tail at the low end (<150).

**Yearly Pass Rates:**

- Year-on-year pass rates (defined as scores  $\geq 200$ ) showed no consistent improvement.
- External factors such as exam administration changes and pandemic-related disruptions may have influenced performance trends.

**Performance by Demographics:**

- Age group 18–21 showed slightly higher pass rates.
- Gender had negligible influence on overall success rates.
- Students from private schools had slightly higher median scores compared to public school counterparts.

**Study Habits**

- Students studying 4–6 hours/day had the highest average scores.
- Those who studied 1–2 hours/day had wider score dispersion, including many in the failure range.

**Resources and CBT Practice:**

- Frequent CBT practice correlated positively with passing.
- Candidates with reliable internet access and digital device access performed better on average.

**Correlation Matrix:**

- Study hours, CBT practice frequency, confidence level, and school type had positive correlations with exam success.

**4. PREDICTIVE MODELING**

A supervised machine learning model was built using scikit-learn to predict whether a student would pass the JAMB exam based on their background and behavior.

**Features Used:**

- Age (converted from ranges to numeric estimates)
- Study Hours per Day (converted to numeric scale)
- School Type (one-hot encoded)

- Internet Access (binary/categorical)
- CBT Practice Frequency (ordinal)
- Confidence Level (ordinal)

**Model Used:**

- **Random Forest Classifier** was selected for its robustness and interpretability.

**Model Performance:**

- **Accuracy:** 62%
- **Precision & Recall:** Balanced performance for both pass and fail classes.
- **Top Predictors:** Study hours, CBT practice, confidence level, and internet access.

**Interpretation:**

- The model accurately learned patterns that distinguish likely passers from likely failers.
- Overfitting was controlled using train-test split and validation metrics.

## 5. FORECASTING FUTURE PASS RATES (2026–2030)

Using historical pass rate data from 2020 to 2025, a linear regression model was applied to forecast trends through 2030.

**Forecast Results:**

- The forecast suggests a slow but positive improvement in pass rates if current support and digital learning trends continue.
- Intervention strategies (CBT access, resource support) could further improve this trend.

## 6. RECOMMENDATIONS

**For Students:**

- Engage in regular CBT-based practice.
- Target 3–4+ hours of study daily.
- Utilize a mix of learning resources, including online tools.

**For Educators and Schools:**

- Incorporate digital learning early.
- Provide guided exam preparation and mock CBT sessions.

**For Policymakers:**

- Improve access to digital infrastructure in rural/public schools.
- Provide subsidized CBT resources and training.
- Encourage early orientation on exam structure.