PROJECT SUMMARY

Explainable AI for Health Workforce Planning

Prepared by Temidayo Sunday Alade, Miva Open University

Python · scikit-learn · XGBoost · SHAP · LIME · FastAPI · Node.js

52,906
de-identified registry records scored

103
Local Government Areas

43
automated tests

11
modules in the analytics package

Master of Information Technology project, Miva Open University. Registry records are used as a scoring cohort only. No individual records are shown.

WHAT IS INSIDE

1 The story behind the project
2 Words used in this document
3 The proposed system
4 The model
5 Data
6 What was tested
7 What the real records showed
8 How the key calculations are worked out
9 How a planner reads a ranked list
10 Why a modest model is still useful
11 Governance and fairness
12 Status and limitations

01 THE STORY BEHIND THE PROJECT

> *Imagine a state is told that a model has ranked its health workers by how likely they are to leave. A planner would reasonably ask: why those workers? What would change the ranking?*
>
> *A score with no reason cannot be acted on, challenged or trusted. This project asks whether a model can give the reason as well as the score.*

The project asks whether a model can warn a state which health workers are likely to leave in the next twelve months, and explain why in terms a planner can check.

02 WORDS USED IN THIS DOCUMENT

Word	What it means
Explainable AI	A model that shows why it gave a score, not only the score.
SHAP	A method that shares a score out among the features that produced it, giving each feature a contribution.
LIME	A method that explains one prediction by fitting a simple model around it.
Calibration	Whether a predicted probability matches what happens. If it says 20%, about 20 in 100 should leave.
Brier score	A measure of how far predicted probabilities are from outcomes. Lower is better.

Word	What it means
AUC	Area under the curve: how well a model ranks leavers above stayers. 0.5 is chance, 1 is perfect.
Development panel	A generated dataset with known rules, used to build and test the model when real exit data does not exist.

03 THE PROPOSED SYSTEM

The National Health Workforce Information System: a national register of health workers, facilities and licences, built on Node.js and Express over a relational store, with a separate Python analytics service on FastAPI.



04 THE MODEL

Part	Choice
Target	Non-retirement exit within twelve months: resignation, termination or dismissal.
Retirement	Handled separately as a deterministic projection from age and years of service.
Models	scikit-learn baselines and XGBoost.
Explanations	SHAP for global and individual effects; LIME for local explanations, shown beside every score.
Fairness	Religion, marital status and raw state of origin are withheld from the features.

05 DATA

Five State registry extracts (Delta, Ogun, Ondo, Plateau and Rivers): 52,906 de-identified records across 103 Local Government Areas. The extracts are enrolment snapshots with no recorded exits, so the model is trained and tested on a generated development panel and the real records are scored. This division is stated in every caption.

> *Why a generated panel? A model that predicts exits needs examples of exits. The registers record who is employed today, not who left. So the thesis builds the model on a panel with known rules, checks that the method recovers them, and applies the finished model to the real records as a scoring cohort. Results on the panel describe the generator and the method. They are not findings about Nigeria's workforce.*

06 WHAT WAS TESTED

- Design Science Research Methodology, with an eleven-module Python analytics package behind a Node.js register.
- A suite of 43 automated tests that defend the design commitments. The suite found two defects before it passed, including a failure of reproducibility in the explanation layer, and both were fixed.

Result on the development panel	Value
Regularised logistic model against an age and service rule: area under the curve	0.688 against 0.595
Brier score before and after calibration	0.1976 to 0.0868
Departures captured by reviewing the top 5 per cent of ranked workers	19.7 per cent, 3.89 times random review

The regularised logistic model was selected over the ensembles.

07 WHAT THE REAL RECORDS SHOWED

The framework projected 5,744 retirements over five years, measured workforce density across the 103 Local Government Areas, and located substantial data quality defects in the registers. Disclosure controls suppress small facility cells instead of nominally protecting them.

08 HOW THE KEY CALCULATIONS ARE WORKED OUT

THREE CALCULATIONS USED IN THE WORK

```
retirement date  =  earlier of ( age 60 ,  35 years of service )
density          =  workers ÷ population × 10,000
Brier score      =  average of ( predicted probability − outcome )²
```

An everyday example. Three workers are given exit probabilities of 0.8, 0.2 and 0.1, and the first leaves while the other two stay. Squared errors are 0.04, 0.04 and 0.01, so the Brier score is 0.09 ÷ 3 = 0.03.

09 HOW A PLANNER READS A RANKED LIST

FROM SCORE TO DECISION

1. The model gives each worker a probability of leaving in the next twelve months, and the workers are ranked from highest to lowest.
2. The planner does not act on one person. They look at the groups that rise to the top, such as a cadre in a district.
3. Beside every score, SHAP shows which features pushed it up or down, so the planner can see the reasons: for example short service, a remote facility, or a cadre in short supply.
4. LIME gives a second, local explanation of the same score, so the planner can check that two methods agree.
5. The planner then chooses a response, such as retention allowances, posting or training, and records the decision.

10 WHY A MODEST MODEL IS STILL USEFUL

Absolute performance on the development panel is modest, and the thesis says so. What matters for planning is how much better the ranking is than reviewing workers at random.

THE IDEA OF LIFT

```
lift  =  share of departures captured  ÷  share of workers reviewed
```

An everyday example. A state has 1,000 workers and about 100 will leave in the year. Reviewing 50 workers (5 per cent) at random would find about 5 leavers. The model's top 5 per cent captures 19.7 per cent of departures, which is about 20 leavers. That is roughly 3.9 times as many, which is the lift the thesis reports. The model is a tool for deciding where to look first and not a verdict on any worker.

11 GOVERNANCE AND FAIRNESS

Concern	What the design does
Unfair features	Religion, marital status and raw state of origin are withheld from the model.

Concern	What the design does
Privacy	Records are de-identified, and small facility cells are suppressed instead of nominally protected.
Trust	Every score is shown with its explanation, so a planner can challenge it.
Overclaiming	Results on the generated panel are described as results about the method, not about Nigeria's workforce.
Misuse	The ranking is for groups. It supports prioritising places and cadres and is not meant to single out individuals.

12 STATUS AND LIMITATIONS

The report was written up in six chapters for Miva Open University, dated September 2026, with Dr. Anayo Ikegwu as supervisor.

Performance is modest in absolute terms and supports the prioritisation of groups rather than the identification of individuals. It is weakest for nursing.

Choosing the logistic model on generated data is a claim about the generator, not a finding about Nigeria's workforce. Whether it behaves that way can only be answered once registers record separations.

Emigration cannot yet be separated from resignation in the available records. Strengthening register data collection is named as the precondition for field deployment.

Temidayo Sunday Alade · aladetemidayo7@gmail.com · dayo@startechone.com.ng · linkedin.com/in/aladetemidayo-sunday