Explainable ML for Fall Risk & Post-Fall Mortality Prediction in Nursing Home Residents

José Alberto Benítez-Andrades[1], Irene Aguado-Caballero[1], Arturo Crespo-Álvaro[1], Sergio Rubio-Martín[1], Alicia Merayo-Corcoba[1] and María Teresa García-Ordás[1]

[1] ALBA Research Group, Department of Electric, Systems and Automatics Engineering, Universidad de León, Spain.

(June, 2026)

srubm@unileon.es

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN

PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in HER (SUICIDETECT)”
MICIU/AEI/10.13039/501100011033

Table of Contents

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

About us:

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 "NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)" MICIU/AEI/10.13039/501100011033

universidad de león

ALBA

About us: ALBA members

Mª Teresa García Ordás

José Alberto Benítez Andrades

Isaías García Rodríguez

Alicia Merayo Corcoba

Sergio Rubio Martín

Arturo Crespo Álvaro

LinkedIn

@ALBALAB_EU

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

AGENCIA ESTATAL DE INVESTIGACIÓN

PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

universidad de león

ALBA

Introduction

Falls are a major cause of **morbidity, mortality, and functional decline** in older adults — with a particularly high burden in nursing home residents, who frequently present frailty, mobility limitations, and multiple chronic conditions.

Clinical Impact

Falls lead to injuries, hospital admissions, loss of autonomy, and increased care needs.

Limitations of Current Tools

Conventional scales and clinical judgement may be limited when used in isolation, especially for complex or severe outcomes.

The Gap

Most published models focus on fall occurrence; far less attention has been paid to post-fall mortality or fall burden.

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)”
MICIU/AEI/10.13039/501100011033

State of the art

How can AI improve fall risk assessment and outcome prediction in institutionalized older adults?

- **Traditional approaches** relied on **wearable sensors** and **motion-capture systems** to detect gait instability and abnormal movement patterns.
- **Recent research** increasingly uses **structured clinical data** from EHRs and institutional records, enabling **scalable** and **low-cost** fall risk prediction.
- **Explainable AI (XAI)** techniques such as **SHAP** have become widely adopted, providing transparency and supporting clinical decision-making.
- **Class imbalance** remains a major challenge. **Synthetic data generation** (e.g., VAEs and generative models) has been proposed to improve model robustness.
- Most studies focus on **fall occurrence prediction**, while clinically relevant outcomes such as **recurrent falls**, **post-fall mortality**, and **injury severity** remain underexplored.
- **Research gap:** Few works simultaneously address **multiple fall-related outcomes** while combining **imbalance-aware learning** and **explainable machine learning**.

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 "NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)" MICIU/AEI/10.13039/501100011033

Datasets & Main Goal of the Project

Three Task-Specific Datasets

Resident-Year Dataset

N = 1,445 — fall occurrence; ~42% fall events

Falls-Only Cohort

N = 1,368 — post-fall mortality; only ~6% positive cases

Functional Subset

N = 200 — includes Tinetti & Berg scores

Predictor Variables

All predictors are **routinely collected structured variables**:

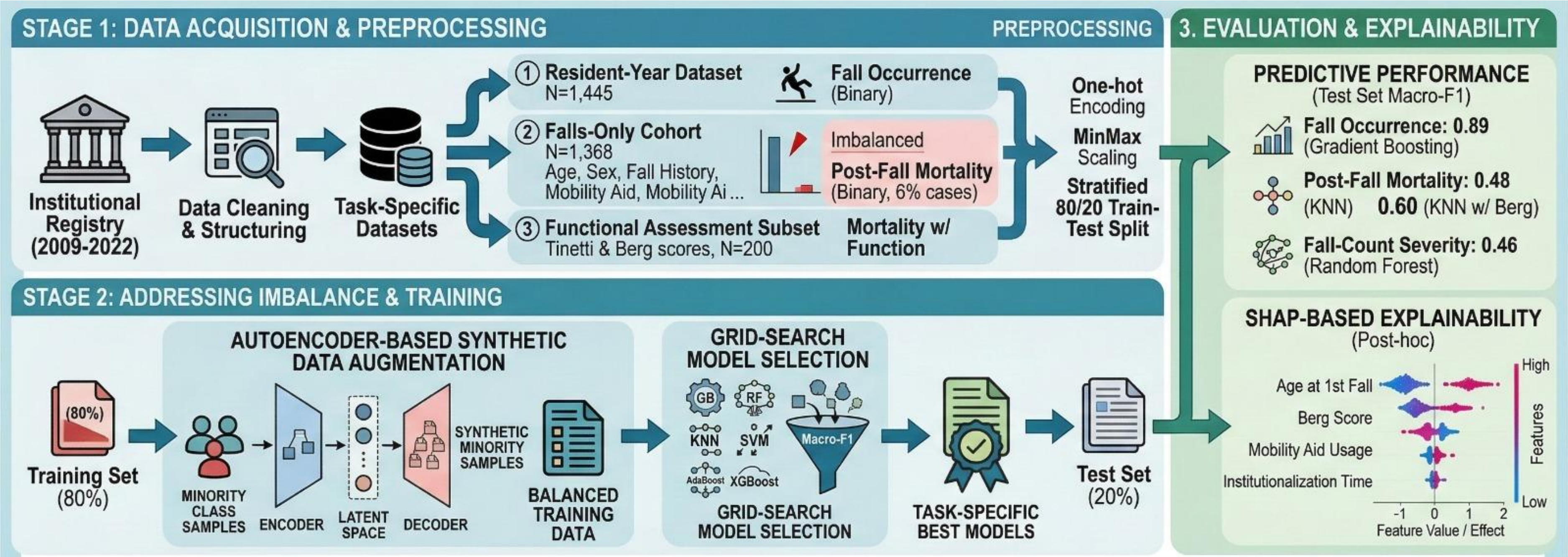
- Age & sex
- Institutionalisation time
- Mobility aid usage
- Fall history indicators
- Tinetti & Berg Balance Scale scores (subset only)

Main Goal: Three Prediction Tasks

- **Fall occurrence** — binary classification
- **Post-fall mortality** — binary, falls-only cohort
- **Fall-count severity** — multiclass (0, 1, 2–3, ≥4 falls)

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 "NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)" MICIU/AEI/10.13039/501100011033

Methodological pipeline



PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

Models & Dataset split

10 Algorithms Evaluated

Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, AdaBoost, KNN, SVM, Gaussian Naive Bayes, LDA, XGBoost

Grid-Search Optimization

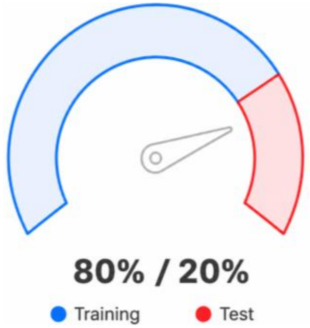
Hyperparameters tuned on training data only. Model selection guided by **Macro-F1**, which gives equal weight to all classes — critical in imbalanced settings.

Tinetty and Berg tests

Tinetti and Berg scores provide objective measures of **mobility** and **balance**.

Dataset distribution

Stratified split to maintain the proportion of samples from each class among the process.



Data imbalance

Task	Positive Class	Positive (%)	Negative (%)	Samples
Fall Occurrence	Fall	42%	58%	1,445
Post-Fall Mortality	Mortality	6%	94%	1,368
Mortality + Berg	Mortality	20%	80%	200

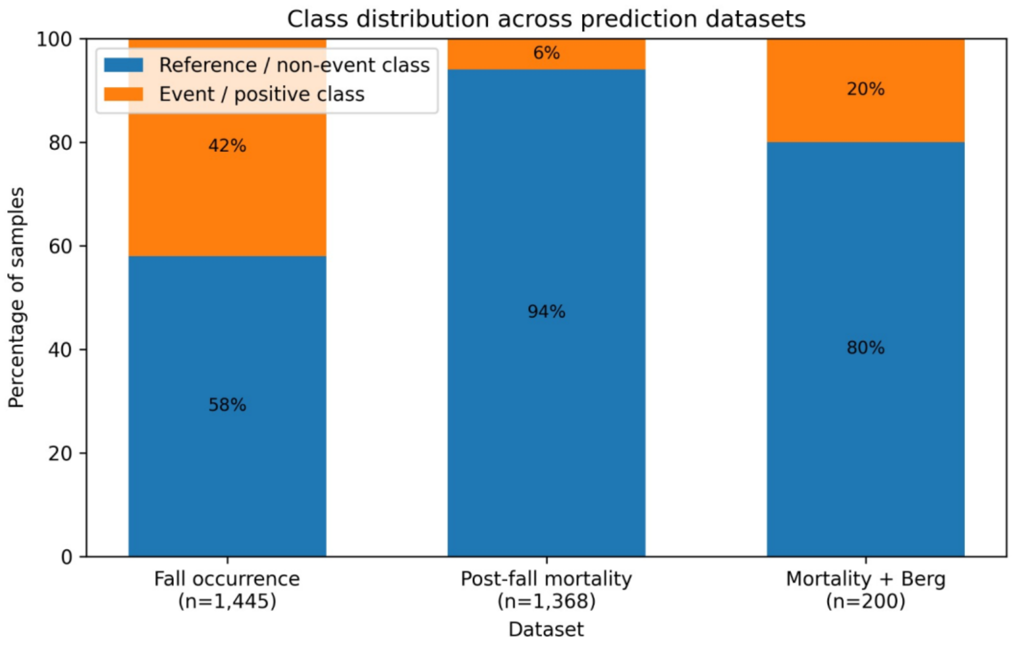
Severe class imbalance in mortality prediction tasks, motivating the use of synthetic data augmentation techniques.

PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

universidad de león

ALBA

Data Imbalance: Autoencoder Augmentation



The Problem

Post-fall mortality has only ~6% positive cases. Standard models fail to learn stable decision boundaries for rare events.

The Solution

An **autoencoder-based synthetic data augmentation** strategy generates minority-class samples in latent space. Applied **only to the training partition** — the test set remains untouched to ensure fair evaluation. This improves model exposure to rare-event patterns without contaminating evaluation.

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

universidad de león

ALBA

Predictive Performance Results

0.89: Fall Ocurrence

Gradient Boosting — strongest result; routine variables carry substantial predictive signal

0.60: Mortality + Berg

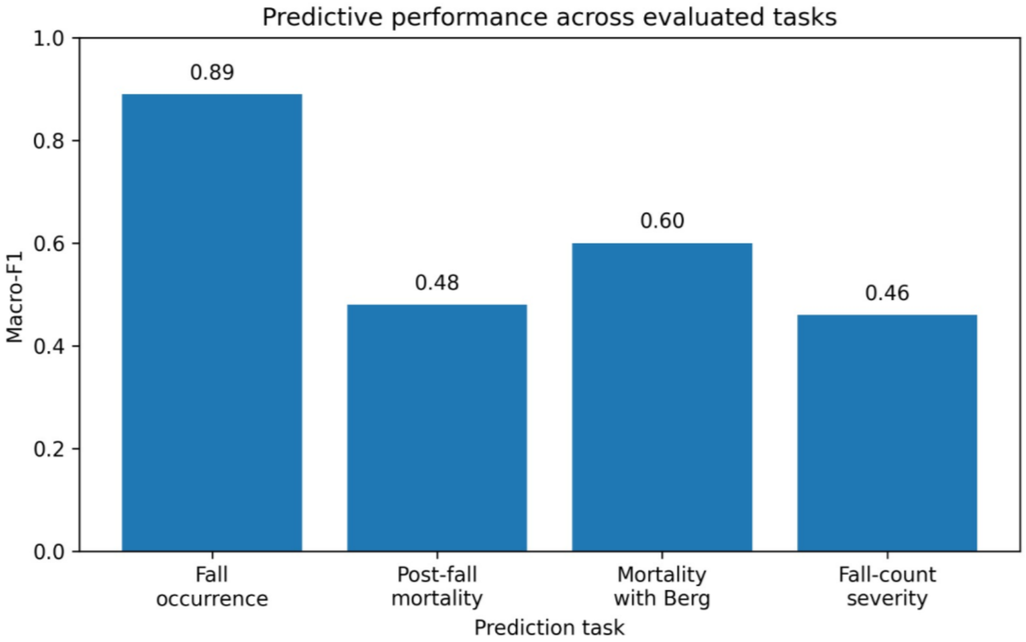
KNN — Berg Balance Scale improves performance, reflecting frailty signal

0.48: Post-Fall Mortality

KNN — limited by severe class imbalance and scarce positive training cases

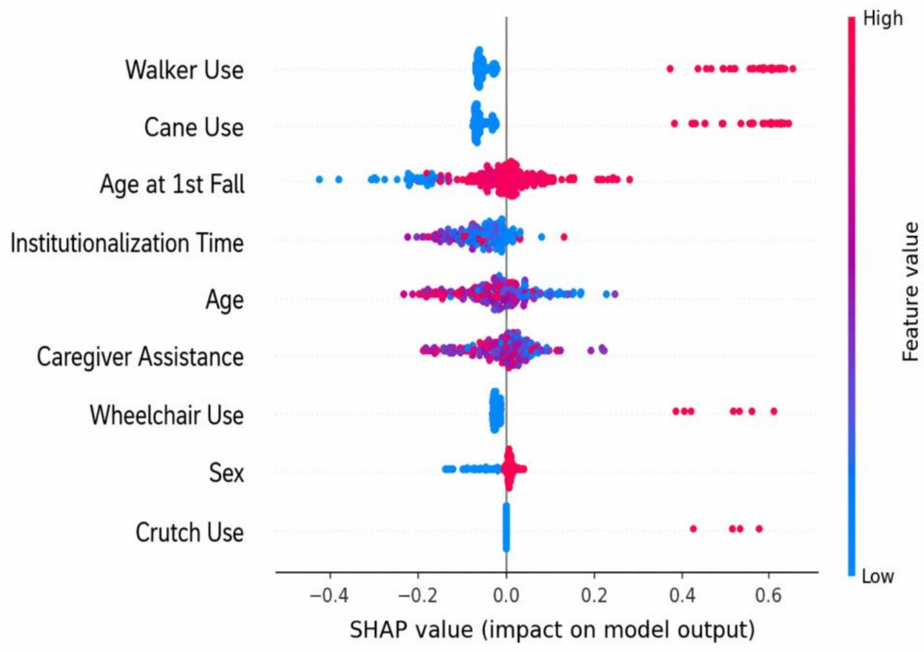
0.46: Fall-Count Severity

Random Forest — above chance in a 4-class task; structured data captures partial gradient



MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 "NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)" MICIU/AEI/10.13039/501100011033

SHAP-Based Explainability



Top Predictors — Fall Occurrence

- Age at first fall
- Institutionalization time
- Chronological age
- Mobility aid usage (walker, cane, wheelchair)

Top Predictors — Mortality (with Berg)

The Berg Balance Scale was the strongest contributor to mortality risk prediction, followed by age and accumulated fall history. SHAP analysis confirms models rely on clinically coherent predictors rather than spurious patterns — supporting trust in decision-support settings.

PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

Discussion: A Gradient of Predictive Difficulty

1

Fall-Count Severity — Macro-F1: 0.46

Most complex; requires medication, environmental, and comorbidity data not available in routine records

2

Post-Fall Mortality — Macro-F1: 0.48 / 0.60

Rare event; augmentation helps but cannot replace real minority-class observations

3

Fall Occurrence — Macro-F1: 0.89

Strongest signal; routine demographic, mobility, and fall-history variables are highly informative

Performance decreases as the prediction target becomes more specific and clinically severe — a pattern consistent with the increasing rarity and complexity of the modelled outcomes.

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN
PID2023-146168OA-I00 "NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)" MICIU/AEI/10.13039/501100011033

universidad de león

ALBA

Conclusions & Future Work

Key Findings

Strong predictive performance
Macro-F1 = 0.89

Relies on routinely available structured variables
No sensors required

Explainable predictions
SHAP improves transparency and clinical interpretability

Multiple clinically relevant endpoints in a unified framework

Limitations

Severe class imbalance
Especially for post-fall mortality (6%)

Single institutional source
Generalizability uncertain

Synthetic augmentation
Cannot replace real minority-class data

No unstructured clinical notes or medication details included

Future Work

Validate in larger, multi-institutional datasets

Explore temporal modelling for dynamic risk prediction

Incorporate richer clinical variables

Conduct prospective evaluation of real-world impact on fall prevention

Practical Impact

The framework can serve as an **initial screening layer** within routine nursing home workflows — not replacing clinical judgement, but **supporting early identification** of residents who may benefit from closer monitoring or functional evaluation.

Routinely Collected Structured Data → ML Framework (Explainable) → Risk Prediction (Falls, Mortality, Severity) → Early Intervention & Clinical Assessment

Better outcomes for nursing home residents

Fall occurrence can be predicted accurately using routinely collected structured data (Macro-F1 = 0.89), enabling explainable and scalable risk screening in nursing homes.

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES | AGENCIA ESTATAL DE INVESTIGACIÓN

PID2023-146168OA-I00 “NLP-Driven Insight Engine for Suicide Attempt Detection in EHR (SUICIDETECT)” MICIU/AEI/10.13039/501100011033

Explainable ML for Fall Risk & Post-Fall Mortality Prediction in Nursing Home Residents

José Alberto Benítez-Andrades[1], Irene Aguado-Caballero[1], Arturo Crespo-Álvaro[1], Sergio Rubio-Martín[1], Alicia Merayo-Corcoba[1] and María Teresa García-Ordás[1]

[1] *ALBA Research Group, Department of Electric, Systems and Automatics Engineering, Universidad de León, Spain.*

(June, 2026)

srubm@unileon.es

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
AGENCIA ESTATAL DE INVESTIGACIÓN

PID2023-146168OA-I00 "NLP-Driven Insight Engine for Suicide Attempt Detection in HER (SUICIDETECT)"
MICIU/AEI/10.13039/501100011033