SAAT - 1: Academic Analytics and Early Warning System

Towards a Dynamic Vectorial Mathematical Model for Academic Risk Prediction and Management

Author

Johnny Villalobos Murillo

jvillalobos@una.ac.cr

Abstract

Higher education institutions face the ongoing challenge of identifying students at risk of academic delay, poor performance, or dropout at an early stage. Traditional approaches are typically based on isolated indicators and retrospective analyses, which often hinder timely intervention.

This paper proposes the Academic Analytics and Early Warning System (SAAT), a dynamic vectorial mathematical model that transforms academic observations into analytical indicators and subsequently into risk metrics, enabling the construction of temporal trajectories for each student and the generation of explainable alerts.

The model defines three interrelated vector spaces:

• Observed Variables Space.

• Analytical Indices Space.

• Academic Risk Space.

The proposed architecture allows the complete reconstruction of a student's academic state at any point in time and measures its evolution through change functions specifically designed for educational analytics.

Keywords—
Academic Analytics, Early Warning Systems, Educational Artificial Intelligence, Predictive Systems, Temporal Trajectories, Characteristic Vectors.

I. INTRODUCTION

Student dropout represents one of the major challenges faced by higher education institutions. Traditional monitoring mechanisms are generally based on isolated indicators, static reports, and analyses performed only after the problem has already materialized.

The increasing availability of academic data enables the development of analytical systems capable of detecting behavioral patterns and anticipating risk situations.

In this context, the Academic Analytics and Early Warning System (SAAT) is proposed as a formal model that transforms academic observations into vector representations capable of describing the temporal evolution of each student.

The proposed research seeks to answer the following question:

Is it possible to construct an explainable mathematical model capable of representing, analyzing, and predicting student academic behavior through hierarchical vector transformations?

II. THEORETICAL FOUNDATIONS

The proposed framework is based on three major areas.

A. Academic Analytics

Academic analytics uses institutional data to understand and improve teaching and learning processes.

Its primary objective is to support decision-making through quantitative evidence.

B. Early Warning Systems

Early Warning Systems aim to identify students at risk of academic failure before such risk materializes.

These systems enable preventive and targeted interventions.

C. Vector Spaces

From a mathematical perspective, each student can be represented as a feature vector.

Vector spaces enable:

• Student comparison.

• Change measurement.

• Pattern detection.

• Application of Artificial Intelligence techniques.

III. FORMAL MODEL DEFINITION

Let

$V(t)$

be the vector of observed variables for a student at time t.

Formally,

$V(t) \in \mathbb{R}^{16}$

where

$V(t)=$

$[V_1, V_2, \ldots, V_{16}]^T$

Each component represents an observed academic variable.

IV. VARIABLE SPACE

We define

$\mathbb{R}^{16}$

as the student observation space.

Each student occupies a point within this space.

Graphically:

```
  V1
   |
   |
   |
   •
  /|\
 / | \
```

V2 ... V16

The set of all observations forms the variable space V.

V. INDEX SPACE

Indices represent analytical dimensions derived from observed variables.

Let

A

be a weighting matrix obtained from the relationship:

index_variable

Then

$I(t)=A \cdot V(t)$

where

$I(t) \in \mathbb{R}^8$

Definition

$I(t)=$

$[I_1, I_2, \ldots, I_8]^T$

Each index synthesizes multiple variables.

For example,

$I_1=$

$0.25V_1 +$
$0.25V_2 +$
$0.25V_3 +$
$0.25V_4$

VI. RISK SPACE

Let

B

be a weighting matrix obtained from the relationship:

risk_index

Then

$R(t)=B \cdot I(t)$

where

$R(t) \in \mathbb{R}^4$

Definition

$R(t)=$

$[R_1, R_2, R_3, R_4]^T$

For example,

$R_1=$

$0.40I_1 +$
$0.30I_2 +$
$0.30I_3$

VII. FUNDAMENTAL THEOREM OF SAAT

Theorem 1

Every academic risk calculated by SAAT can be expressed directly as a linear transformation of the observed variables vector.

Proof

We know that

$I(t)=A\cdot V(t)$

and

$R(t)=B\cdot I(t)$

Substituting,

$R(t)=B(A\cdot V(t))$

By associativity,

$R(t)=(BA)V(t)$

Defining

$M=BA$

we obtain

$R(t)=M\cdot V(t)$

Therefore, the theorem is proven.

□

VIII. DIMENSIONALITY REDUCTION PRINCIPLE

SAAT implements a hierarchical dimensionality reduction process:

$\mathbb{R}^{16}$

$\downarrow$

$\mathbb{R}^{8}$

$\downarrow$

$\mathbb{R}^{4}$

That is:

16 Variables

↓

8 Indices

↓

4 Risks

This process transforms detailed information into strategic knowledge.

IX. STUDENT MATHEMATICAL STATE

The academic state of a student is defined as

$S_e(t)=$

$(V_e(t), I_e(t), R_e(t))$

This state simultaneously contains:

• Observed information.

• Analytical information.

• Risk information.

X. TEMPORAL ACADEMIC TRAJECTORY

Let

$t_1<t_2<...<t_n$

The student's trajectory is defined as

$T_e=$

$\{$

$S_e(t_1)$,

$S_e(t_2)$,

...

$S_e(t_n)$

$\}$

This structure allows the study of the student's historical evolution.

XI. FICE FUNCTION

Student Change Impact Factor

Let

$\Delta I =$

$I(t_2) - I(t_1)$

We define

FICE=

$\|\Delta I\|$

Using the Euclidean norm,

FICE=

$\sqrt{\Sigma(I_i(t_2) - I_i(t_1))^2}$

Interpretation

It measures the magnitude of analytical change experienced by a student.

XII. FRCE FUNCTION

Student Risk Change Factor

Let

$\Delta R =$

$R(t_2) - R(t_1)$

We define

FRCE=

$\|\Delta R\|$

Using the Euclidean norm,

FRCE=

$\sqrt{\Sigma(R_k(t_2) - R_k(t_1))^2}$

Interpretation

It measures the magnitude of risk variation experienced by a student.

XIII. RESEARCH HYPOTHESIS

General Hypothesis

The use of temporal characteristic vectors enables the modeling and prediction of student academic behavior with greater explanatory power than traditional isolated indicators.

XIV. PROPOSED SCIENTIFIC CONTRIBUTION

The primary innovation of SAAT lies in explicitly separating three levels of representation:

VCE → Observational State

ICE → Analytical State

RCE → Risk State

This separation allows the complete reconstruction of the process leading to an academic alert.

Unlike black-box models, SAAT preserves full traceability of the factors responsible for generating a prediction.

XV. FORMAL DEFINITION OF SAAT

SAAT=

(V,A,I,B,R,T)

where:

V = Variable Space

A = Index Construction Matrix

I = Index Space

B = Risk Construction Matrix

R = Risk Space

T = Temporal Dimension

XVI. FUNDAMENTAL EQUATION

$R(t)=B \cdot A \cdot V(t)$

XVII. CONCLUSIONS

SAAT constitutes a formal mathematical architecture for academic analytics and early warning systems.

The model enables:

• Representation of students through dynamic vectors.

• Construction of temporal academic trajectories.

• Measurement of analytical changes.

• Measurement of risk variations.

• Generation of explainable alerts.

• Integration of future Artificial Intelligence techniques upon a solid mathematical foundation.

The proposed formulation transforms SAAT into a conceptual and computational framework capable of evolving into advanced predictive systems for academic management in higher education institutions.