# NAÏVE BAYES

# TITLE & ABSTRACT

**What question were the investigators asking?**

Sử dụng giải thuật Naive Bayes để phân tích hiệu suất học tập của sinh viên

## What techniques did they use?

Họ lấy dữ liệu học tập từ các sinh viên khoa Điện toán sau đó thực hiện tiền xử lý dữ liệu để tìm ra phân lớp và nhóm dữ liệu. Trong nghiên cứu này, kỹ thuật phân lớp được mô tả là dựa trên giải thuật Naive Bayes và được dùng cho khai thác các dữ liệu về học thuật.

## What were their major results?

Nó hỗ trợ sinh viên và giảng viên đánh giá hiệu suất học tập để đưa ra kế hoạch điều chỉnh giúp sinh viên không trượt kỳ thi.

## What conclusions did they draw?

Đây là nghiên cứu phương pháp đề phòng cho sinh viên trong tiến độ học tập.

# INTRODUCTION

## Why did they start?

Nhiều tổ chức giáo dục tiên tiến được xây dựng trải dài Pakistan. Ngược lại, chất lượng giáo dục được đánh giá bởi tỉ lệ tốt nghiệp thành công của sinh viên và bằng chứng nhận tốt nghiệp có đủ để giữ chân sinh viên. Dự doán hiệu suất học tập của sinh viên là một phao cứu sinh giúp sinh viên nhận ra dựa trên khả năng trượt bài kiểm tra, vì vậy sự quản lý sẽ cung cấp hỗ trợ đúng lúc và đem đến những bước quan trọng để chỉ dẫn cho sinh viên hỗ trợ tăng lại hiệu suất học tập. Khả năng dự đoán hiệu suất của học sinh là bắt buộc trong ngành giáo dục. Sử dụng phương pháp khai thác dữ liệu với dữ liệu có số lượng lớn và phát hiện mẫu thông tin được che giấu có tính hợp tác trong việc ra quyết định. Đó là xác định các khía cạnh khác nhau đã ảnh hưởng đến các hoạt động học tập và hiệu suất của sinh viên trong toàn ngành giáo dục

## Why is this research important?

Có ý kiến ​​cho rằng thông tin tương quan nên được truyền đến giảng viên trước khi bắt đầu các bài kiểm tra. Nó đã giúp các giảng viên giảm tỷ lệ thất bại xuống mức đáng kể và nâng cao hiệu suất của sinh viên

## Is it justified (review of literature?)

Trong nghiên cứu này, dựa trên kỹ thuật khai thác dữ liệu và phân cụm dữ liệu của naïve bayes đã giúp cho giới học thuật để dự đoán SGPA (GPA) của sinh viên và dựa trên đó giảng viên là giai đoạn thiết yếu để cải thiện kết quả học tập cử nhân của sinh viên. SGPA (Điểm trung bình học kỳ) thường được sử dụng để đánh giá kết quả học tập sinh viên, điểm trung bình thấp nhất được các khoa khác nhau đặt ra và được duy trì. Do đó, điểm trung bình chung vẫn là yếu tố được chia sẻ nhiều nhất được các nhà tổ chức giáo dục sử dụng để đánh giá sự tiến hóa trong lĩnh vực học thuật sinh viên.

## What is the research question/s of interest (purpose/objectives) or hypothesis?

Việc thiếu hụt phương pháp giáo dục là điều hiển nhiên trong ngành công nghiệp chiến thuật thường phàn nàn rằng chương trình mới tốt nghiệp ngành khoa học máy tính chưa sẵn sàng để giải quyết các dự án kỹ thuật phần mềm trong thế giới thực

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Author | Name | Year | Idea | Terms defined? |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

# METHOD

## What did they do

Data preparation and Data Pre-processing

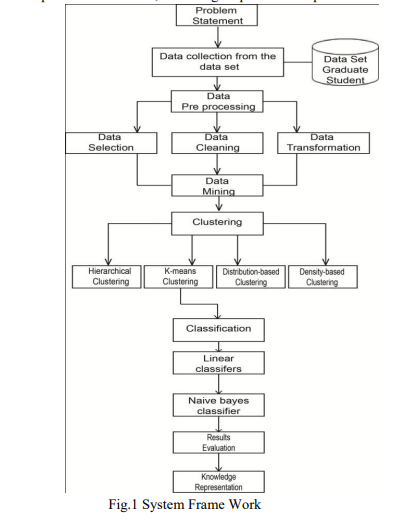
Data Selection and Transformation

Data clustering

Classification

Using Naïve Bayes Algorithm

## Flow execution



# RESULT

## What did they find?

* Table

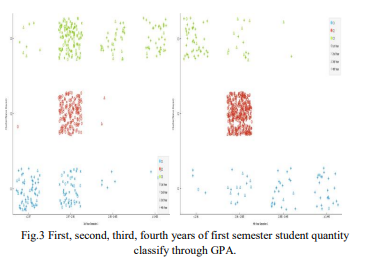
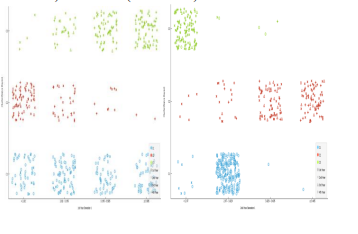
A picture containing table

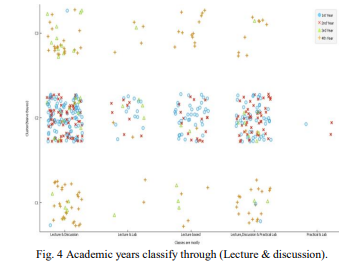
Description automatically generated

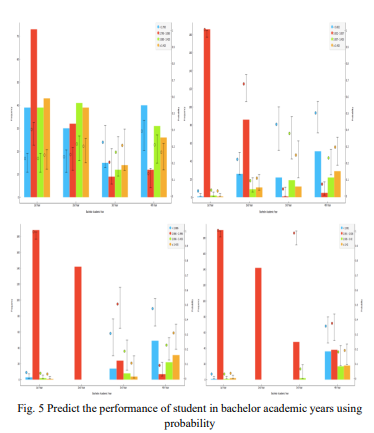
* Figure

Chart, scatter chart

Description automatically generated





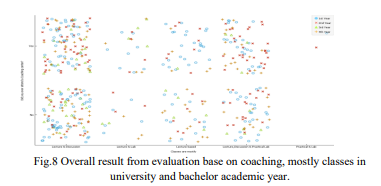


Chart, bar chart

Description automatically generated

Diagram

Description automatically generated



Chart, bar chart

Description automatically generated

## Were the number of subjects in each group or subgroup used in the analysis specified? Parameter

Vâng, số lượng môn học của mỗi nhóm hoặc mỗi nhóm nhỏ hơn được phân tích cụ thể với thước đo là người.

## Were the subject characteristics summarized?

Vâng, các tính chất được tóm gọn và sử dụng phương thức so sánh để phân tích

## Did the results relate to the specified objective/hypotheses?

Vâng, Kết quả liên quan đến mục tiêu hỗ trợ sinh viên cải thiện hiệu suất học tập

## Do the tables and figures “speak for themselves” ?

Vâng, các diagram và bảng đã thể hiện rõ các tính chất của dữ liệu.

## Are the tables adequately titled, labeled?

Vâng, các bảng có tiêu đề đầy đủ và dán nhãn.

# DISCUSSION

## What do the results mean?

Nghiên cứu này đã giúp sinh viên và giảng viên cải thiện kết quả của sinh viên có nguy cơ trượt bài kiểm tra. Nghiên cứu này cũng là nỗ lực để xác định những sinh viên cần được xem xét đặc biệt để giảm thiểu tỉ lệ trượt bài kiểm tra và thực hiện hành động phù hợp cho kỳ thi học kỳ sắp tới.

## Did authors discuss the results in relations to the objectives/hypotheses?

Vâng, tác giả có bàn luận về kết quả trong mối liên hệ với mục đích nghiên cứu

## Were the results discussed in relation to those from similar studies?

Vâng, Nghiên cứu có kham khảo từ 16 nghiên cứu khác về chủ đề học thuật

## Are the authors justified in the strength of the statements they make in the study?

Tác giả chứng minh được bằng sức mạnh của các tuyên bố trong nghiên cứu

## Did they offer an alternative explanation for the result?

Người ta đã sử dụng cách tiếp cận phân lớp là phân lớp Naïve Bayesian để dự đoán điểm trung bình chung của sinh viên sau đại học. Ngoài ra, nó phân cụm sinh viên thành các tập bằng thuật toán phân cụm K-Means. Dữ liệu như điểm chuyên cần, thảo luận và bài tập được thu thập từ bộ dữ liệu trước đó của học sinh, để dự đoán kết quả học tập vào cuối học kỳ.