

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

---



**BÙI THỊ BÉ BA**

**HỆ THỐNG QUẢN LÝ VÀ DỰ BÁO TUYỂN SINH  
TẠI HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**LUẬN VĂN THẠC SỸ KỸ THUẬT**

*(Theo định hướng ứng dụng)*

TP. HỒ CHÍ MINH – NĂM 2022

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**



**BÙI THỊ BÉ BA**

**HỆ THỐNG QUẢN LÝ VÀ DỰ BÁO TUYỂN SINH  
TẠI HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

**CHUYÊN NGÀNH: HỆ THỐNG THÔNG TIN**

**MÃ SỐ: 8.48.01.04**

**LUẬN VĂN THẠC SỸ KỸ THUẬT**

*(Theo định hướng ứng dụng)*

**NGƯỜI HƯỚNG DẪN KHOA HỌC:**

**TS. HUỖNH TRỌNG THƯA**

**TP. HỒ CHÍ MINH – NĂM 2022**

## LỜI CAM ĐOAN

Tôi xin cam đoan luận văn thạc sĩ hệ thống thông tin “**Hệ thống quản lý và dự báo tuyển sinh tại Học viện Công nghệ Bru chính Viễn thông**” do tôi nghiên cứu, tổng hợp và thực hiện.

Toàn bộ nội dung luận văn, những điều được trình bày là của chính cá nhân tôi hoặc là được tham khảo, tổng hợp từ nhiều nguồn tài liệu khác nhau. Tất cả các tài liệu tham khảo, tổng hợp đều được trích xuất với nguồn gốc rõ ràng. Các số liệu, kết quả nêu trong luận văn là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

*TP.HCM, ngày 16 tháng 07 năm 2022*

**Học viên thực hiện luận văn**

**Bùi Thị Bé Ba**

## LỜI CẢM ƠN

Trước tiên tôi xin gửi lời cảm ơn sâu sắc nhất đến giảng viên hướng dẫn trực tiếp của tôi - **TS. Huỳnh Trọng Thura**. Cảm ơn Thầy đã luôn lắng nghe những quan điểm cá nhân và đưa ra những nhận xét quý báu, góp ý và dẫn dắt tôi đi đúng hướng trong suốt thời gian thực hiện đề tài luận văn thạc sĩ.

Tôi xin gửi lời cảm ơn chân thành nhất đến đến **Thầy Lê Quang Phú**. Cảm ơn Thầy đã dành thời gian thu thập dữ liệu và sớm cung cấp dữ liệu tuyển sinh thực tế của trường để tạo điều kiện cho tôi nghiên cứu và hoàn thiện luận văn.

Tôi cũng xin trân trọng cảm ơn các Thầy Cô Khoa Đào tạo Sau đại học - Học viện Công nghệ Bru chính Viễn thông Cơ sở Thành phố Hồ Chí Minh đã truyền đạt những kiến thức chuyên môn sâu rộng trong ngành cho tôi, có được nền tảng kiến thức hỗ trợ rất lớn cho tôi trong quá trình nghiên cứu, hoàn thiện luận văn.

Cuối cùng, tôi xin gửi lời cảm ơn chân thành đến gia đình, bạn bè và quý đồng nghiệp vì đã luôn đồng hành, tạo điều kiện, hỗ trợ và khuyến khích tôi cố gắng trong suốt những năm tháng học tập cũng như quá trình nghiên cứu viết luận văn này.

Xin chân thành cảm ơn!

*TP.HCM, ngày 16 tháng 07 năm 2022*

**Học viên thực hiện luận văn**

**Bùi Thị Bé Ba**

## MỤC LỤC

LỜI CAM ĐOAN .....	i
LỜI CẢM ƠN .....	ii
MỤC LỤC.....	iii
DANH MỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT .....	v
DANH SÁCH BẢNG.....	vi
MỞ ĐẦU .....	1
1. Lý do chọn đề tài.....	1
2. Tổng quan vấn đề nghiên cứu.....	2
3. Mục tiêu nghiên cứu.....	3
4. Đối tượng và phạm vi nghiên cứu.....	4
5. Phương pháp nghiên cứu .....	5
6. Cấu trúc luận văn .....	5
<b>Chương 1: NGHIÊN CỨU TỔNG QUAN .....</b>	<b>6</b>
1.1. Nghiên cứu quy trình tuyển sinh.....	6
1.2. Nghiên cứu tổng quan về dự báo .....	7
1.3. Nghiên cứu về mô hình phân rã - Thuật toán Prophet .....	8
1.4. Nghiên cứu quy trình xây dựng hệ thống .....	10
<b>Chương 2: PHÂN TÍCH THIẾT KẾ.....</b>	<b>11</b>
2.1. Thu thập dữ liệu .....	11
2.2. Tiền xử lý dữ liệu .....	12
2.3. Phân tích, đánh giá dữ liệu .....	13
2.4. Áp dụng mô hình Prophet cho dữ liệu .....	<b>Error! Bookmark not defined.</b>
2.5. Đánh giá và lựa chọn mô hình phát triển hệ thống quản lý.....	21
2.6. Các IDE và ngôn ngữ lập trình sử dụng để phát triển hệ thống .....	22

2.7. Thiết kế cơ sở dữ liệu .....	24
<b>Chương 3: TRIỂN KHAI ỨNG DỤNG.....</b>	<b>29</b>
3.1. Giới thiệu tổng quan về mô hình hệ thống triển khai .....	29
3.2. Thực tế triển khai ứng dụng.....	30
<b>Chương 4: ĐÁNH GIÁ KẾT QUẢ THỬ NGHIỆM.....</b>	<b>53</b>
4.1. Kết quả thực nghiệm .....	53
4.2. Nhận xét kết quả thử nghiệm ứng dụng .....	62
<b>KẾT LUẬN.....</b>	<b>63</b>
<b>TÀI LIỆU THAM KHẢO .....</b>	<b>64</b>

## DANH MỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT

<b>Viết tắt</b>	<b>Tiếng Anh</b>	<b>Tiếng Việt</b>
DB	Database	Cơ sở dữ liệu
LSTM	Long Short-Term Memory	Bộ nhớ ngắn-dài hạn
UML	Unified Modeling Language	Ngôn ngữ mô hình hóa thống nhất
IDE	Integrated Development Environment	Môi trường phát triển tích hợp
SQL	Structured Query Language	Ngôn ngữ truy vấn mang tính cấu trúc
API	Application Programming Interface	Giao diện lập trình ứng dụng
UI	User interface	Giao diện người dùng
THPT	-	Trung học phổ thông
HS	-	Hồ sơ
DS	-	Danh sách
ĐGNL	-	Đánh giá năng lực
MAE	Mean Absolute Error	Sai số tuyệt đối trung bình
MAPE	Mean Absolute Percentage Error	Tỉ lệ phần trăm sai số tuyệt đối trung bình

## DANH SÁCH BẢNG

Bảng 2.1.1: Bảng thông tin thí sinh đăng ký xét tuyển năm 2020.....	11
Bảng 2.1.2: Bảng thông tin thí sinh trúng tuyển (1) .....	11
Bảng 2.1.3: Bảng thông tin thí sinh trúng tuyển (2) .....	11
Bảng 2.1.4: Bảng thông tin sinh viên nhập học .....	12
Bảng 2.2.1: Bảng thông tin sinh viên trúng tuyển đã format đúng định dạng .....	13
Bảng 2.2.2: Bảng thông tin sinh viên nhập học đã format đúng định dạng .....	13
Bảng 2.3.1: Bảng dữ liệu thông tin sinh viên trúng tuyển của từng ngành qua mỗi năm .....	14
Bảng 2.3.2: Bảng dữ liệu tổng số lượng sinh viên trúng tuyển qua các năm.....	15
Bảng 2.3.3: Bảng dữ liệu tổng số lượng sinh viên trúng tuyển ngành công nghệ đa phương tiện qua các năm.....	15
Bảng 2.4.1: Bảng dữ liệu chứa các thông số của thuật toán Prophet.....	20
Bảng 2.7.1: Bảng CSDL Area.....	24
Bảng 2.7.2: Bảng CSDL Provinces .....	25
Bảng 2.7.3: Bảng CSDL Districts .....	25
Bảng 2.7.4: Bảng CSDL Regions .....	25
Bảng 2.7.5: Bảng CSDL Race .....	25
Bảng 2.7.6: Bảng CSDL priority .....	26
Bảng 2.7.7: Bảng CSDL majors.....	26
Bảng 2.7.8: Bảng CSDL examinees.....	27
Bảng 2.7.9: Bảng CSDL xettuyenkethop .....	28
Bảng 2.7.10: Bảng CSDL xettuyenkythinl.....	28
Bảng 4.1: So sánh kết quả.....	61



## DANH SÁCH HÌNH VẼ

Hình 2.4.1: Mô hình Prophet Logistic Growth .....	17
Hình 2.4.2: Công thức tăng trưởng theo Logistic Growth .....	18
Hình 2.4.3: Mô hình tuyến tính từng mảnh .....	19
Hình 2.5.1: Mô hình phát triển hệ thống .....	21
Hình 2.6.1: Mô hình sơ đồ phát triển hệ thống.....	23
Hình 2.7.1: Mô hình tổng quan cơ sở dữ liệu .....	24
Hình 3.1.1: Source code backend trên IntelliJ .....	30
Hình 3.1.4: Source code font-end trên IDE visual studio code .....	30
Hình 3.2.1: Giao diện form đăng nhập .....	31
Hình 3.2.2: Giao diện trang quản lý tuyển sinh .....	31
Hình 3.2.3: Giao diện các chức năng trong menu của trang quản lý tuyển sinh .....	32
Hình 3.2.4: Giao diện của mục “Xét tuyển kết quả THPT” .....	33
Hình 3.2.5: Giao diện chức năng “Quản lý danh sách” thí sinh.....	33
Hình 3.2.6: Giao diện chức năng chọn nút “Tạo mới” .....	34
Hình 3.2.7: Giao diện form thêm mới thông tin một thí sinh.....	35
Hình 3.2.8: Giao diện update thông tin một thí sinh.....	36
Hình 3.2.9: Giao diện “Thêm ds trúng tuyển” .....	37
Hình 3.2.10: Format định dạng file Excel của “Thêm ds trúng tuyển” (Từ cột A tới cột U) .....	37
Hình 3.2.11: Format định dạng file Excel tiếp theo của “Thêm ds trúng tuyển” (Từ cột V tới cột AJ) .....	38
Hình 3.2.12: Giao diện chức năng “Thêm ds nhập học”.....	38
Hình 3.2.13: Format định dạng file Excel của “Thêm ds nhập học” .....	39
Hình 3.2.14: Giao diện của mục “Xét tuyển kết hợp” .....	39
Hình 3.2.15: Giao diện form “Nhập thông tin” của hình thức “Xét tuyển kết hợp” .....	40
Hình 3.2.16: Giao diện form “Thông tin chi tiết” của “Xét tuyển kết hợp” .....	40
Hình 3.2.17: Giao diện của mục “Xét tuyển kì thi ĐGNL” .....	41

Hình 3.2.18: Giao diện form “Nhập thông tin” của hình thức “Xét tuyển kì thi ĐGNL”.....	41
Hình 3.2.19: Giao diện form “Thông tin chi tiết” của “Xét tuyển kì thi ĐGNL” ....	42
Hình 3.2.20: Giao diện của mục “Dự báo” ở trang quản lý .....	42
Hình 3.2.21: Giao diện tổng quan gồm các mục của trang dự báo.....	43
Hình 3.2.22: Giao diện tổng quan xem dữ liệu và dự báo của mục “Xét tuyển dựa vào kết quả thi THPT” .....	44
Hình 3.2.23: Giao diện lựa chọn một tiêu chí để xem dữ liệu và dự báo .....	45
Hình 3.2.24: Giao diện xem dữ liệu theo Ngành năm 2016 của đối tượng trúng tuyển .....	46
Hình 3.2.25: Giao diện xem dữ liệu theo Ngành năm 2020 của tất cả đối tượng ....	46
Hình 3.2.26: Giao diện xem dữ liệu theo Ngành qua các năm của đối tượng trúng tuyển .....	47
Hình 3.2.27: Giao diện xem dữ liệu theo Ngành qua các năm của tất cả đối tượng	47
Hình 3.2.28: Giao diện tổng quan lựa chọn số năm, chọn đối tượng dự báo theo một ngành nào đó .....	48
Hình 3.2.29: Giao diện chọn dự báo một ngành trong các ngành của từng đối tượng .....	48
Hình 3.2.30: Giao diện kết quả dự báo hai năm tới (2022 và 2023) của đối tượng trúng tuyển theo ngành Công nghệ Đa phương tiện của mục “Xét tuyển dựa vào kết quả thi THPT” .....	49
Hình 3.2.31: Giao diện format file excel kết quả dự báo khi tải về.....	49
Hình 3.2.32: Giao diện tổng quan xem dữ liệu và dự báo của mục “Xét tuyển kết hợp” .....	50
Hình 3.2.33: Giao diện chọn một ngành để xem dữ liệu .....	51
Hình 3.2.34: Giao diện số liệu tổng quan theo ngành Công nghệ Đa phương tiện qua các năm của các đối tượng (đăng ký, trúng tuyển, nhập học) .....	51
Hình 3.2.35: Giao diện kết quả dự báo năm tới (2024) của đối tượng trúng tuyển theo ngành Công nghệ Đa phương tiện của mục “Xét tuyển kết hợp” .....	52

Hình 4.1.1 : Lệnh cài đặt môi trường .....	53
Hình 4.1.2: Các lệnh đọc dữ liệu từ google drive .....	54
Hình 4.1.3: Dữ liệu ngành Công nghệ đa phương tiện qua các năm .....	54
Hình 4.1.4: Biểu đồ trend dữ liệu.....	55
Hình 4.1.5: Đoạn mã lệnh chia dữ liệu thành 2 tập Cus_train và Cus_test.....	55
Hình 4.1.6: Tập Cus_train và tập Cus_test sau khi chia dữ liệu.....	56
Hình 4.1.7: Tập Train và tập Test được vẽ ra biểu đồ .....	56
Hình 4.1.8: Câu lệnh Thuật toán prophet được áp dụng .....	56
Hình 4.1.9: Câu lệnh hiển thị kết quả dự báo .....	57
Hình 4.1.10: Kết quả áp dụng thuật toán Prophet với mô hình growth=linear .....	57
Hình 4.1.11: Công thức và kết quả MAE tập Train.....	58
Hình 4.1.12: Công thức và kết quả MAE tập Test.....	58
Hình 4.1.13: Công thức và kết quả MAPE tập Train.....	59
Hình 4.1.14: Công thức và kết quả MAPE tập Test .....	59
Hình 4.1.15: Kết quả dự báo năm 2020 và 2021 .....	59
Hình 4.1.16: Bộ tham số của changepoint_prior_scale.....	60
Hình 4.1.17: Giá trị RMSE tương ứng với mỗi tham số .....	60
Hình 4.1.18: Thuật toán Prophet với tham số changepoint_prior_scale.....	61
Hình 4.1.19: Kết quả dự báo năm 2020 và 2021 với changepoint_prior_scale .....	61

## MỞ ĐẦU

### 1. Lý do chọn đề tài

Công nghệ thông tin là một trong những ngành khoa học ngày càng được quan tâm và sử dụng rộng rãi trong mọi lĩnh vực của cuộc sống. Với những ưu điểm mạnh có thể ứng dụng được nhiều trong cuộc sống thì công nghệ thông tin đã giúp cho công việc quản lý được dễ dàng hơn. Để quản lý một công việc cụ thể của một cơ quan nào đó thì người quản lý cần thực hiện rất nhiều công việc. Nhờ có các phần mềm giúp công việc đó trở nên đơn giản hơn rất nhiều.

Bộ giáo dục và đào tạo cũng đã và đang áp dụng công nghệ thông tin vào các hoạt động của mình từ việc xử lý, cập nhật đến các hoạt động cao hơn đặc biệt là công tác tuyển sinh. Hiện nay tất cả các trường đại học và cao đẳng nước ta đã tin học hoá hầu hết các công đoạn của tuyển sinh và đạt được những kết quả đáng kể đặc biệt là giảm thiểu mức độ sai sót trong hồ sơ của thí sinh, tính toán điểm số,... làm tăng độ tin cậy cho công tác tuyển sinh.

Tại Học viện Công nghệ Bưu chính Viễn thông, hằng năm có hàng ngàn thí sinh tham gia xét tuyển, nhưng công tác quản lý tuyển sinh vẫn còn đang được quản lý trên file excel; việc thống kê các kết quả như kết quả trúng tuyển, kết quả nhập học cần kết hợp từ nhiều file,... Bên cạnh đó để có sự chuẩn bị về cơ sở vật chất, về đội ngũ cán bộ giảng viên, nên Học viện cũng cần phỏng đoán trước các số liệu liên quan đến công tác tuyển sinh cho các năm sắp tới. Nên việc xây dựng một phần mềm quản lý và dự báo về số liệu tuyển sinh nhằm tránh rủi ro sai sót từ người quản lý, thuận lợi cho công tác quản lý, đồng thời góp phần cho việc ước tính và đưa ra các chiến lược tuyển sinh phù hợp cho các năm sắp tới là cần thiết.

Đề tài “Hệ thống quản lý và dự báo tuyển sinh tại Học viện Công nghệ Bưu chính Viễn thông” với mục tiêu xây dựng một hệ thống để quản lý kết hợp với dự báo số liệu giúp hạn chế sai sót cho cán bộ quản lý trong quá trình sử dụng và giúp cán bộ quản lý nắm bắt kịp thời tình hình tuyển sinh, các thông số liên quan đến tuyển

sinh để có cái nhìn tổng quan từ đó nhanh chóng đưa ra những giải pháp, chiến lược về tư vấn tuyển sinh sắp tới, về cơ sở vật chất, về đội ngũ giảng viên.

## **2. Tổng quan vấn đề nghiên cứu**

Để xây dựng hệ thống, trước hết cần nghiên cứu các công cụ, ngôn ngữ lập trình phù hợp để xây dựng một hệ thống tích hợp chức năng quản lý và dự báo, sau đó thu thập thông tin, dữ liệu liên quan đến các thông số cần thiết và đánh giá tầm quan trọng của thông số đầu vào, từ đó tìm hiểu mô hình, kịch bản đánh giá để đặt vấn đề cho thông số đầu vào.

Dự báo các số liệu tuyển sinh là điều cần thiết trong công tác tuyển sinh. Phục vụ cho nhiều hoạt động như tư vấn tuyển sinh (nếu số lượng sinh viên ở vùng miền nào còn ít, thì đẩy mạnh công tác tư vấn tuyển sinh xuống vùng miền đó), có sự chuẩn bị kịp thời về cơ sở vật chất, đội ngũ giảng viên (dựa vào số lượng sinh viên được dự báo theo ngành).

Dự báo theo chuỗi thời gian là một lớp mô hình quan trọng trong thống kê, kinh tế lượng và máy học. Mô hình là chuỗi thời gian (time series) được áp dụng trên các chuỗi đặc thù có yếu tố thời gian. Một mô hình chuỗi thời gian thường dự báo dựa trên giả định rằng các quy luật trong quá khứ sẽ lặp lại ở tương lai. Do đó xây dựng mô hình chuỗi thời gian là đang mô hình hóa mối quan hệ trong quá khứ giữa biến độc lập (biến đầu vào) và biến phụ thuộc (biến mục tiêu). Dựa vào mối quan hệ này để dự báo giá trị trong tương lai của biến phụ thuộc. Hiện nay có rất nhiều nghiên cứu dự báo chuỗi thời gian và dự báo phụ tải điện, tuy nhiên có thể phân loại thành 2 nhóm là phương pháp cổ điển dựa trên thống kê và phương pháp hiện đại dựa trên công nghệ trí tuệ nhân tạo.

Gần đây, các nghiên cứu dự báo chuỗi thời gian phi tuyến theo mô hình cộng (Additive Model) trong đó các xu thế phi tuyến khớp với tính thời vụ hàng năm, hàng tuần và hàng ngày, cộng với các ảnh hưởng ngày lễ cũng được sử dụng cho hiệu quả tốt. Nhiều công trình nghiên cứu dự báo theo chuỗi thời gian bằng thuật toán Facebook Prophet gần đây như:

Bài báo “Phân tích và dự báo hoạt động đầu tư tại khu công nghệ cao thành phố Hồ Chí Minh”, Huỳnh Trọng Thura và cộng sự [1]. Kết quả nghiên cứu của bài báo là cơ sở để xây dựng các mô hình dự báo thông minh cho các khu công nghệ cao khác trong cả nước. Trong nghiên cứu này, các tác giả đã thiết kế và xây dựng các chức năng quan trọng như web nhập liệu, mô hình phân tích và dự báo hoạt động đầu tư tại khu công nghệ cao dựa trên thuật toán Prophet của Facebook. Kết quả thu được là trang web thể hiện các thông tin thống kê và dự báo dòng vốn đầu tư vào khu công nghệ cao cho thời gian tới.

Bài báo “Time series facebook prophet model and python for covid-19 outbreak prediction”, Mashaal Khayyat, Kaouther Laabidi, Nada Almalki and Maysoon Al-zahrani[2]. Kết quả của bài báo đã dự báo sớm về sự lây lan của coronavirus để tạo điều kiện thuận lợi cho việc lập kế hoạch phòng chống dịch bệnh. Trong bài báo này các tác giả đã đưa ra dự báo về COVID-19 bằng cách sử dụng kỹ thuật dữ liệu chuỗi thời gian dựa trên tập dữ liệu hiện đang được đề xuất để phân tích dữ liệu bùng phát của vi rút COVID-19. Bài báo này đề xuất việc sử dụng phương pháp phân tích dữ liệu để dự báo đại dịch. Kết quả cho thấy mô hình được đề xuất có khả năng dự báo thấp về các trường hợp được khôi phục của bộ dữ liệu COVID-19. Ngược lại, mô hình đề xuất các trường hợp tử vong có khả năng dự báo cao trong tập dữ liệu COVID-19.

### **3. Mục tiêu nghiên cứu**

#### ***3.1. Mục tiêu tổng quát***

Mục tiêu của đề tài là xây dựng hệ thống để hỗ trợ cho công tác quản lý tuyển sinh, cung cấp giao diện dễ dùng, dễ dàng trong việc xem các số liệu, tránh rủi ro sai sót từ người quản lý. Hỗ trợ để đưa ra con số dự báo cho chỉ tiêu từng ngành, từng vùng góp phần vào công tác tư vấn tuyển sinh và sớm có sự chuẩn bị về cơ sở vật chất, về đội ngũ cán bộ giảng viên cho kì tuyển sinh sắp tới.

### **3.2. Mục tiêu cụ thể**

Xây dựng hệ thống gồm các chức năng sau:

Chức năng quản lý gồm:

- Chức năng nhập thông tin hồ sơ thí sinh:
  - Nhập qua form: thông tin cá nhân, thông tin học tập, kết quả học tập,
  - Import thông tin từ file excel vào hệ thống.
- Chức năng tìm kiếm thông tin: Tìm kiếm và hiển thị thông tin theo năm, ngành, vùng hoặc tìm kiếm theo mã sinh viên.

Chức năng dự báo gồm:

- Dự báo số lượng hồ sơ trúng tuyển, số lượng sinh viên nhập học, theo từng ngành, từng khu vực, từng vùng, hoặc theo đối tượng ưu tiên.
- Xuất ra đồ thị thể hiện số liệu tương ứng với các thông số lựa chọn.

## **4. Đối tượng và phạm vi nghiên cứu**

### **4.1. Đối tượng nghiên cứu:**

Quy trình, nghiệp vụ quản lý tuyển sinh đại học và nhu cầu thống kê, dự báo các thông tin về số lượng hồ sơ trúng tuyển, số lượng sinh viên thực học,... của từng ngành tại Học viện Công nghệ Bưu chính Viễn thông.

Các tài liệu, dữ liệu tuyển sinh từ năm 2016 đến 2021 tại Học viện Công nghệ Bưu chính Viễn thông.

Thuật toán mã nguồn mở Prophet của Facebook, các công cụ, các framework, các kỹ thuật để xây dựng trang web.

### **4.2. Phạm vi nghiên cứu:**

Nghiên cứu về quy trình tuyển sinh đại học tại Học viện Công nghệ Bưu chính Viễn thông.

Dựa vào số liệu liên quan đến vấn đề tuyển sinh để xây dựng hệ thống quản lý và dự báo về các chỉ tiêu (cụ thể là số hồ sơ đăng ký xét tuyển, số lượng hồ sơ trúng tuyển, số lượng sinh viên thực học) tại Học viện Công nghệ Bưu chính Viễn thông.

## **5. Phương pháp nghiên cứu**

Đề tài này sử dụng phương pháp nghiên cứu lý thuyết kết hợp với xây dựng ứng dụng thực nghiệm:

- Nghiên cứu lý thuyết quy trình tuyển sinh.
- Tìm hiểu, nghiên cứu thuật toán Prophet.
- Thu thập các tài liệu có liên quan tới đề tài, các dữ liệu ở hiện tại và quá khứ.
- Đánh giá tầm quan trọng của thông số đầu vào, tìm hiểu mô hình, kịch bản đánh giá để đặt vấn đề cho thông số đầu vào.
- Tìm hiểu các mô hình trong thuật toán Prophet, đánh giá và lựa chọn mô hình phù hợp với dữ liệu đầu vào.
- Tiến hành đánh giá kết quả thực nghiệm, đưa ra hướng phát triển mở rộng của đề tài để đáp ứng những nhu cầu triển khai thực tế.

## **6. Cấu trúc luận văn**

Ngoài phần mở đầu, mục lục, kết luận và kiến nghị, danh mục hình vẽ, danh mục bảng biểu, tài liệu tham khảo, phụ lục, phần chính của luận văn gồm 4 chương như sau:

Chương 1: NGHIÊN CỨU TỔNG QUAN

Chương 2: PHÂN TÍCH THIẾT KẾ

Chương 3: TRIỂN KHAI ỨNG DỤNG

Chương 4: ĐÁNH GIÁ KẾT QUẢ THỬ NGHIỆM



## Chương 1: NGHIÊN CỨU TỔNG QUAN

### 1.1. Nghiên cứu quy trình tuyển sinh

Sau khi tìm hiểu hoạt động của quy trình quản lý tuyển sinh tại trường và nhận thấy quy trình đang thực hiện theo nhiều công đoạn. Đầu tiên trường công bố chỉ tiêu mỗi ngành nghề, tổ hợp các môn xét tuyển mỗi ngành. Sau đó, tiến hành nhận hồ sơ dự tuyển và kiểm tra hồ sơ, nhập liệu thông tin hồ sơ, lưu trữ dữ liệu vào file excel (thông tin sinh viên, điểm các môn, điểm ưu tiên, nguyện vọng ngành,...). Tiếp theo là thống kê dữ liệu (bảng thống kê thông tin thí sinh, bảng điểm, nguyện vọng ngành,...).

Hiện tại Học viện có các phương pháp xét tuyển sau:

- Phương thức 1: Xét tuyển thẳng và ưu tiên xét tuyển là xét tuyển thẳng đối với thí sinh đạt 27,0 điểm trong kỳ thi tốt nghiệp THPT.
- Phương thức 2: Xét tuyển dựa vào kết quả thi THPT. Có khoảng 85% chỉ tiêu cho xét tuyển dựa vào kết quả thi THPT. Phương thức này xét trúng tuyển theo từng ngành và không có điểm chênh lệch giữa các tổ hợp bài thi/môn thi trong cùng một ngành.
- Phương thức 3: Xét tuyển kết hợp là sự kết hợp giữa kết quả học tập ở bậc THPT với chứng chỉ quốc tế hoặc thành tích cá nhân trong kỳ thi tuyển chọn học sinh giỏi hoặc thành tích học tập tại các trường THPT chuyên (phương thức này có chức năng nộp hồ sơ online, có 2 nguyện vọng) [3].
- Phương thức 4: Xét tuyển dựa vào kỳ thi đánh giá năng lực, kỳ thi đánh giá năng lực được tổ chức với mục tiêu đánh giá năng lực học đại học của thí sinh, đa dạng hóa hình thức tuyển sinh và đồng thời tăng thêm cơ hội trúng tuyển vào Học viện. Kết quả thi đánh giá năng lực là một trong những phương thức xét tuyển vào Học viện năm 2021 vừa qua.

Cán bộ quản lý thực hiện kết nối với cổng thông tin tuyển sinh của bộ giáo dục và đào tạo để truy xuất dữ liệu, sau đó tải về và lưu trữ các tập tin excel để xem báo

cáo, thống kê số lượng, bảng điểm, số lượng hồ sơ, sau đó ra quyết định điểm trúng tuyển, báo cáo danh sách trúng tuyển. Tiếp theo Học viện sẽ công bố danh sách trúng tuyển lên cổng thông tin Học viện. Cuối cùng là khâu xuất giấy báo nhập học và gửi giấy báo.

Quy trình tuyển sinh tại Học viện qua rất nhiều công đoạn, luận văn tập trung vào để hỗ trợ giải quyết công đoạn từ việc tập hợp các tập tin excel để thống kê các kết quả và từ đó lấy dữ liệu đầu vào để hỗ trợ cho chức năng dự báo số liệu.

## **1.2. Nghiên cứu tổng quan về dự báo**

Dự báo là khoa học và nghệ thuật tiên đoán các sự việc có thể sẽ xảy ra trong tương lai. Tính khoa học của dự báo thể hiện ở chỗ khi tiến hành dự báo người ta phải căn cứ trên các dữ liệu phản ánh tình hình thực tế trong quá khứ và hiện tại, căn cứ vào xu thế trên cơ sở khoa học để dự báo những sự việc có thể sẽ xảy ra trong tương lai. Tính nghệ thuật trong dự báo thể hiện ở chỗ phải sử dụng tài phán đoán, kinh nghiệm trong những điều kiện thiếu thông tin hoặc nhu cầu của khách hàng biến động mạnh. Từ những đặc điểm này cho thấy dự báo bao giờ cũng có sai số, chỉ ngẫu nhiên nếu chúng ta dự báo đúng hoàn toàn, tính chính xác của dự báo càng thấp khi thời gian dự báo càng dài.

Các phương pháp dự báo cơ bản được phân thành hai loại cơ bản: Các phương pháp định tính và định lượng, và việc lựa chọn loại thích hợp chủ yếu phụ thuộc vào dữ liệu có sẵn. Các phương pháp dự báo định tính mang tính chủ quan, sự kiện tương lai được dự báo chủ quan dựa trên việc sử dụng ý kiến của các chuyên gia; tuy nhiên, chúng không hoàn toàn là phỏng đoán, nhưng chúng là các phương pháp tiếp cận có cấu trúc được phát triển để thu được các dự báo tốt mà không cần sử dụng dữ liệu lịch sử. Do đó, các phương pháp này hữu ích và được triển khai khi dữ liệu lịch sử không có sẵn hoặc khan hiếm. Mặt khác, các phương pháp dự báo định lượng dựa trên toán học và công thức thống kê. Chúng được áp dụng khi có sẵn dữ liệu, nhưng phải thỏa mãn hai điều kiện: thông tin về quá khứ có sẵn và giả định rằng một số các khía cạnh của mô hình trong quá khứ sẽ tiếp tục trong tương lai. Các phương pháp dự báo định lượng liên quan đến một loạt các phương pháp và mỗi phương pháp có

các thuộc tính, độ chính xác và chi phí riêng phải được cân nhắc khi lựa chọn một phương pháp cụ thể trong các ngành cụ thể, cho các mục đích cụ thể.

### 1.3. Nghiên cứu về mô hình phân rã - Thuật toán Prophet

Ý tưởng của mô hình phân rã là chuỗi thời gian có thể được phân tách thành ba thành phần: xu hướng (hướng dài hạn), theo mùa (có hệ thống, chuyển động liên quan đến lịch) và bất thường (biến động không hệ thống, ngắn hạn). Mô hình phân rã thông thường chia thành 2 loại: Mô hình cộng và Mô hình nhân

Trong mô hình cộng, chuỗi thời gian quan sát ( $O_t$ ) được coi là tổng của ba thành phần độc lập: theo mùa  $St$ , xu hướng  $Tt$  và bất thường  $It$ :

$$O_t = T_t + S_t + I_t \quad (1.3.1)$$

Trong mô hình nhân, chuỗi thời gian quan sát ( $O_t$ ) được biểu thị như là tích của các thành phần xu hướng  $Tt$ , theo mùa  $St$  và bất thường  $It$ :

$$O_t = T_t \times S_t \times I_t \quad (1.3.2)$$

Gần đây, các nghiên cứu dự báo chuỗi thời gian phi tuyến theo mô hình cộng (Additive Model) trong đó các xu thế phi tuyến khớp với tính thời vụ hàng năm, hàng tuần và hàng ngày, cộng với các ảnh hưởng ngày lễ cũng được sử dụng cho hiệu quả tốt. Sean J. Taylor và Benjamin Letham [4] đã giới thiệu phương pháp phân tích chuỗi thời gian theo mô hình cộng bằng thuật toán Facebook Prophet trong dự báo số sự kiện trên Facebook. Huỳnh Trọng Thừa và các đồng sự [1] đã công bố kết quả nghiên cứu sử dụng thuật toán phân tích chuỗi thời gian Facebook Prophet trong dự báo hoạt động đầu tư ở Khu Công nghệ cao TP.HCM. Zar Zar Oo và Sabai Phyu [5] sử dụng thuật toán Prophet để dự báo nhiệt độ ở tỉnh Myintkyina, Myanmar cho kết quả tốt so với các phương pháp truyền thống. Toni Toharudin và các đồng sự đã công bố báo cáo so sánh kết quả dự báo nhiệt độ không khí hàng ngày trong 5 năm ở Bandung, Taiwan giữa 2 mô hình dự báo bởi LSTM và Facebook Prophet; Các kết quả cho thấy rằng, Prophet cho kết quả tốt hơn đối với nhiệt độ tối đa, trong khi LSTM cho kết quả tốt hơn đối với nhiệt độ tối thiểu. Tuy nhiên, sự khác biệt về giá trị của RMSE không quá lớn đáng kể. Emir Žunić và các đồng sự [6] đề xuất bộ khung ứng dụng thuật

toán Prophet trong dự báo doanh số bán hàng trong thực tế để dự báo doanh số hàng tháng, hàng quý sắp tới.

Trên nền tảng mã nguồn mở có tên là Prophet do Facebook phát triển, nền tảng này hỗ trợ cho cả ngôn ngữ Python và ngôn ngữ R bên cạnh việc cung cấp các thông số trực quan, dễ điều chỉnh. Prophet là một mô hình dự báo dữ liệu chuỗi thời gian dựa trên mô hình cộng (Additive model) trong đó các xu hướng phi tuyến tính phù hợp với thời vụ hàng năm, hàng tuần và hàng ngày, cộng với các ảnh hưởng ngày lễ. Cơ sở của mô hình Prophet là phân rã chuỗi thời gian thành 4 thành phần đại diện cho xu hướng, tính chu kỳ, ảnh hưởng của ngày lễ và sai số mô hình theo phương trình:

$$y(t) = g(t) + s(t) + h(t) + \epsilon t \quad (1.3.3)$$

Trong đó:

- $g(t)$  là hàm đại diện cho xu hướng, sự thay đổi không có tính chu kỳ của chuỗi thời gian.
- $s(t)$  là hàm đại diện cho sự thay đổi có tính chu kỳ (tính mùa vụ) của chuỗi thời gian (ví dụ: hàng tuần, hàng tháng, hàng năm).
- $h(t)$  là hàm đại diện cho sự ảnh hưởng của ngày nghỉ, ngày lễ (do người dùng cung cấp) xảy ra theo lịch một hoặc vài ngày.
- $\epsilon t$  là sai số mang tính ngẫu nhiên không xác định được của mô hình.

Ưu điểm của thuật toán Prophet:

- Rất nhanh, vì nó được xây dựng bằng Stan, một ngôn ngữ lập trình để suy luận thống kê được viết bằng C ++.
- Một mô hình hồi quy cộng hưởng trong đó các xu hướng phi tuyến phù hợp với tính thời vụ hàng năm, hàng tuần và hàng ngày.
- Mạnh mẽ ngay cả khi thiếu dữ liệu và thay đổi theo xu hướng, có khả năng xử lý tốt cho các trường hợp ngoại lệ.
- Dễ dàng thay đổi quy trình và điều chỉnh dự báo cũng như bổ sung thông tin chi tiết về dữ liệu doanh nghiệp.

#### **1.4. Nghiên cứu quy trình xây dựng hệ thống**

- Thu thập thông tin, số liệu liên quan đến các thông số cần thiết cho việc đánh giá, áp dụng thuật toán để dự báo.
- Đánh giá tầm quan trọng của thông số đầu vào, tìm hiểu mô hình, kịch bản đánh giá để đặt vấn đề cho thông số đầu vào.
- Nghiên cứu tổng quan về thuật toán Prophet, tính ứng dụng của thuật toán, công cụ, ngôn ngữ để triển khai thuật toán.
- Nghiên cứu các công cụ, ngôn ngữ lập trình phù hợp để xây dựng một hệ thống tích hợp hai chức năng quản lý và dự báo.

## Chương 2: PHÂN TÍCH THIẾT KẾ

### 2.1. Thu thập dữ liệu

Tuyển sinh là công tác hằng năm của Học viện, từ năm 2014 trở về trước thì thi quy chế tuyển sinh dựa vào kỳ thi đại học, cao đẳng. Từ năm 2016 trở lại Bộ giáo dục và đào tạo có thay đổi, không tổ chức kỳ thi đại học, cao đẳng, thay vào đó sẽ tuyển sinh dựa vào kết quả xét tuyển.

Công việc tiếp theo là thu thập dữ liệu tuyển sinh của Học viện Công nghệ Bưu chính Viễn thông từ năm 2016 đến năm 2021 của bộ phận quản lý tuyển sinh. Dữ liệu là các file excel bao gồm:

- Thông tin chi tiết thí sinh đăng ký xét tuyển năm 2020:

**Bảng 2.1.1: Bảng thông tin thí sinh đăng ký xét tuyển năm 2020**

STT	Số hồ sơ	SBD	Họ Tên	CMND	Ngày sinh	Giới tính	ĐTUT	KVUT	Năm TN THPT	Tổng điểm xét tuyển	KQ sơ tuyển	KQ xét tuyển sinh	Ngành TT	Tên ngành
1	1	02015944	LÊ ĐỨC ANH	079200006257	11/12/2000	Nam		3	2019			Trượt		
2	1	29031128	BÙI TRỌNG ANH	187862358	06/10/2001	Nam		1	2019	22.55		Trúng tuyển	7480201	Công nghệ thông tin
3	1	34005901	TRẦN THỊ CAM	206279213	14/11/2001	Nữ		1	2019	21.25		Trúng tuyển	7329001	Công nghệ đa phương tiện
4	1	37000396	NGUYỄN THIÊN BÁO	215531276	24/08/2001	Nam		2	2019					
5	1	37005626	HỒ THÁI PHONG BA	215529626	12/11/2001	Nam		2	2019			Trượt		
6	1	37013828	DƯƠNG QUANG ANH	215501758	07/03/2001	Nam		2NT	2019			Trượt		

- Thông tin chi tiết thí sinh trúng tuyển gồm 2 bảng nối nhau:

**Bảng 2.1.2: Bảng thông tin thí sinh trúng tuyển (1)**

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG -																								
STT	SBD	Họ tên	Ngày sinh	Giới tính	Mã ngành trúng	Tên ngành trúng tuyển	Mã tổ hợp	Thứ tự NV	Điểm trúng tuyển	Điểm khuyến khích	Mã môn 1	Điểm môn 1	Mã môn 2	Điểm môn 2	Mã môn 3	Điểm môn 3	ĐTUT	KVUT	Năm tốt nghiệp THPT	Học lực	Hạng kiểm	Điểm TB lớp 12	TN cao đẳng	TN trung cấp
1	40014088	ĐẶNG LÊ TUẤN ANH	30/11/2002	Nam	7329001	Công nghệ đa phương tiện	A00	5	24.9	0	TO	8.4	LI	8	HO	7.75		1	2020	Khá	Tốt	7.7		
2	44005414	TẠ THỊ HOÀNG ANH	10/02/2002	Nữ	7329001	Công nghệ đa phương tiện	D01	1	23.8	0	TO	8	VA	8.75	N1	6.8		2	2020	Khá	Tốt	7.7		

**Bảng 2.1.3: Bảng thông tin thí sinh trúng tuyển (2)**

DANH SÁCH THÍ SINH TRÚNG TUYỂN												
TN cao đẳng	TN trung cấp	HK Tỉnh	HK Huyện	HK Xã	Mã tỉnh lớp 12	Mã trường lớp 12	Điện thoại	Địa chỉ Email	Địa chỉ nhận giấy báo	Nơi sinh	Dân tộc	Số CMND
		40 - Đắk Lắk	01 - Thành phố Buôn Ma Thuột	03 - Phường Tân An	40	074	0946233777		Trường THCS và THPT Đông Du, Hẻm 129 Nguyễn Chí Thanh, Khóm 8, Phường Tân An, Tp. Buôn Ma Thuột, Đắk Lắk	ĐẮK LẮK	Kinh	241864654
		44 - Bình Dương	11 - Thành phố Dĩ An		44	081	0798294608	<a href="mailto:sunvmon102@gmail.com">sunvmon102@gmail.com</a>	Tà Thị Hoàng Anh - SĐT 0798294608 - Lớp 12C12 - Trường THPT Bình An, KP Bình Thung 1, P Bình An, TP Dĩ An, Tỉnh Bình Dương	BÌNH DƯƠNG	Kinh	231440623

- Thông tin chi tiết sinh viên nhập học:

**Bảng 2.1.4: Bảng thông tin sinh viên nhập học**

DANH SÁCH SINH VIÊN HOÀN TẤT THỦ TỤC NỘP HỌC PHÍ, NHẬP HỌC, ...		
KHÓA 2020		
Số TT	Số báo danh	Họ và Tên sinh viên
1	02051360	TRẦN GIA LONG
2	35009705	VÕ DUY TÙNG
3	36001580	PHẠM THỊ NGÀ
4	61001461	NGUYỄN PHƯỚC DUY THỊNH
5	52002654	ĐỖ THU GIANG
6	43002281	NGUYỄN HOÀNG THIÊN
7	49006719	HỒ GIA HUY

Dữ liệu thu thập được ở mỗi năm có định dạng không đồng nhất. Thứ nhất là về sự sắp xếp các cột không giống nhau giữa các file ở các năm. Thứ 2 là định dạng file có sự khác biệt về cách trình bày, thứ tự hàng bắt đầu. Thứ ba là mã ngành, ở năm 2016 và năm 2017 tuy cùng một ngành với những năm 2018 đến 2021 nhưng mã ngành khác (ví dụ ngành công nghệ thông tin, năm 2016 có mã ngành là D480201, nhưng năm 2017 trở về sau mã ngành có sửa đổi thành 7480201),...

## 2.2. Tiền xử lý dữ liệu

Sau khi thu thập nhận thấy dữ liệu không thống nhất, qua các năm đều có sự thay đổi về cấu trúc file dữ liệu nên cần bước tiền xử lý dữ liệu trước khi tiến hành viết các chức năng import cho hệ thống. Các file khác nhau nên đưa về cùng một định dạng để phục vụ cho chức năng import. Các file cần cho luận văn gồm 2 file: file thông tin sinh viên trúng tuyển và thông tin nhập học.

Để phục vụ chức năng import, tác giả chỉnh sửa các file dữ liệu thu thập được về cùng một cấu trúc như sau: Format lại file thông tin sinh viên trúng tuyển như bảng 2.2.1 bên dưới:

**Bảng 2.2.1: Bảng thông tin sinh viên trúng tuyển đã format đúng định dạng**

STT	SBD	Họ tên	Ngày sinh	Giới tính	Mã ngành trúng tuyển	Tên ngành trúng tuyển	Mã tổ hợp trúng tuyển	Thứ tự NV trúng tuyển	Điểm trúng tuyển	Điểm khuyến khích	Môn 1	Môn 2	Môn 3	ĐTUT		
1	40014088	ĐẶNG LÊ TUẤN ANH	30/11/2002	Nam	7329001	Công nghệ đa phương tiện	A00	5	24.9		TO	8.4	LI	8	HO	7.75
2	44005414	TẠ THỊ HOÀNG ANH	10/02/2002	Nữ	7329001	Công nghệ đa phương tiện	D01	1	23.8		TO	8	VA	8.75	N1	6.8
3	44009298	LUỖNG THỊ NGỌC ANH	29/06/2002	Nữ	7329001	Công nghệ đa phương tiện	A00	1	23.95		TO	7.2	LI	7.75	HO	8.5

ĐTUT	KV UT	Năm tốt nghiệp THPT	Học lực	Hạng kiểm	Điểm TB lớp 12	TN cao đẳng	TN trung cấp	HK Tỉnh	HK Huyện	HK Xã	Mã tỉnh lớp 12	Mã trường lớp 12	Điện thoại	Địa chỉ Email	Địa chỉ nhận giấy báo	Nơi sinh	Dân tộc	Số CMND
	1	2020	Khá	Tốt	7.7			40 - Đắk Lắk	01 - Thành phố Buôn Ma Thuột	03 - Phường Tân An, Tp. Buôn Ma Thuột, Đắk Lắk	40	074	0946233777		Trường THCS và THPT Đông Du, Hẻm 129 Nguyễn Chí Thanh, Khối 8, Phường Tân An, Tp. Buôn Ma Thuột, Đắk Lắk	ĐẮK LẮK	Kinh	241864654
	2	2020	Khá	Tốt	7.7			44 - Bình Dương	11 - Thành phố Dĩ An		44	081	0798294608	suraymon102@gmail.com	Tạ Thị Hoàng Anh - SĐT 0798294608 - Lp 12C12 - Trường THPT Bình An, KP Bình Thang 1, P. Bình An, TP Dĩ An, Tỉnh Bình Dương	BÌNH DƯƠNG	Kinh	231440623
	2NT	2020	Khá	Tốt	7.6			44 - Bình Dương	06 - Huyện Phú Giáo		44	030	0353934169	hoangthangocanh12a@tayson.sgd.binhduong.edu.vn	Lương Thị Ngọc Anh - Lp12A, Trường THPT Tây Sơn - Ấp 2, Tân Lông, Phú Giáo, Bình Dương	BÌNH DƯƠNG	Kinh	281292785
	2NT	2020	Giỏi	Tốt	8.1			53 - Tiền Giang	07 - Huyện Gò Công Tây		53	022	0975478325	vobao412002@gmail.com	Có Đoàn Lê Bạch Thảo, giáo viên Trường TH Nguyễn Văn Trứ, xã Thạnh Nhứt, huyện Gò Công Tây, Tiền Giang.	TỈNH TIỀN GIANG	Kinh	312502909

Tiếp theo là format lại file thông tin sinh viên nhập học như bảng 2.2.2 bên dưới:

**Bảng 2.2.2: Bảng thông tin sinh viên nhập học đã format đúng định dạng**

Số TT	Số báo danh	Họ và Tên sinh viên
1	02051360	TRẦN GIA LONG
2	35009705	VÕ DUY TÙNG
3	36001580	PHẠM THỊ ANH
4	61001461	NGUYỄN PHƯỚC DUY THỊNH
5	52002654	ĐỖ THU GIANG
6	43002281	NGUYỄN HOÀNG THIÊN

### 2.3. Phân tích, đánh giá dữ liệu

Tiếp theo sau tiền xử lý dữ liệu thì ta tiến hành truy vấn, tổng hợp dữ liệu để hỗ trợ việc phân tích và đánh giá dữ liệu. Kết quả truy vấn dữ liệu sinh viên trúng tuyển của từng ngành qua mỗi năm thể hiện ở bảng sau:



**Bảng 2.3.1: Bảng dữ liệu thông tin sinh viên trúng tuyển từng ngành qua mỗi năm**

<b>Năm</b>	<b>Số lượng</b>	<b>Ngành</b>
2016	127	An toàn thông tin
2016	111	Công nghệ đa phương tiện
2016	318	Công nghệ thông tin
2016	183	Điện - Điện tử
2016	149	Điện tử - Viễn thông
2016	133	Kế toán
2016	118	Marketing
2016	116	Quản trị kinh doanh
2017	82	An toàn thông tin
2017	83	Công nghệ đa phương tiện
2017	204	Công nghệ thông tin
2017	103	Điện - Điện tử
2017	139	Điện tử - Viễn thông
2017	65	Kế toán
2017	72	Marketing
2018	112	An toàn thông tin
2018	98	Công nghệ đa phương tiện
2018	268	Công nghệ thông tin
2018	85	Điện - Điện tử
2018	93	Điện tử - Viễn thông
2018	103	Kế toán
2018	99	Marketing
2018	96	Quản trị kinh doanh
2019	106	An toàn thông tin
2019	99	Công nghệ đa phương tiện
2019	241	Công nghệ thông tin
2019	59	Điện - Điện tử
2019	91	Điện tử - Viễn thông
2019	112	Kế toán
2019	99	Marketing
2019	108	Quản trị kinh doanh
2020	66	An toàn thông tin
2020	107	Công nghệ đa phương tiện
2020	169	Công nghệ thông tin
2020	42	Điện - Điện tử
2020	89	Điện tử - Viễn thông
2020	71	Kế toán
2020	90	Marketing
2020	58	Quản trị kinh doanh
2021	63	An toàn thông tin
2021	96	Công nghệ đa phương tiện
2021	190	Công nghệ thông tin
2021	106	Điện - Điện tử
2021	126	Điện tử - Viễn thông
2021	76	Kế toán
2021	71	Marketing
2021	50	Quản trị kinh doanh

Kết quả lấy tổng số lượng sinh viên trúng tuyển qua các năm như bên dưới:

**Bảng 2.3.2: Bảng dữ liệu tổng số lượng sinh viên trúng tuyển qua các năm**

	ds	y
▶	2016-01-01	1255
	2017-01-01	748
	2018-01-01	954
	2019-01-01	915
	2020-01-01	692
	2021-01-01	778

Kết quả lấy tổng số lượng sinh viên trúng tuyển ngành công nghệ đa phương tiện qua các năm như bên dưới:

**Bảng 2.3.3: Bảng dữ liệu tổng số lượng sinh viên trúng tuyển ngành công nghệ đa phương tiện qua các năm**

	ds	y
▶	2016-01-01	111
	2017-01-01	83
	2018-01-01	98
	2019-01-01	99
	2020-01-01	107
	2021-01-01	96

Nhìn chung dữ liệu tuyển sinh phụ thuộc vào quy chế, quy định của Bộ giáo dục và đào tạo. Vì vậy ta thấy dữ liệu có sự thay đổi rất lớn, nhìn về con số tổng quát, như trên bảng 2.3.2 dữ liệu có sự biến đổi lớn nhất là ở năm 2020, do năm 2020 có chính sách thay đổi từ Bộ giáo dục và đào tạo.

Nhưng nhìn riêng ngành Công nghệ đa phương tiện tuy năm 2020 số lượng sinh viên trúng tuyển giảm mạnh, nhưng số lượng sinh viên trúng tuyển ngành Công nghệ Đa phương tiện lại tăng, và qua các năm ngành này không có sự biến động nhiều.

## 2.4. Áp dụng mô hình Prophet cho dữ liệu

Prophet có tính khoa học vì đã được nhiều nhà khoa học nghiên cứu để vận dụng vào dự báo dữ liệu chuỗi thời gian. Về sự phù hợp với bài toán dự báo số liệu tuyển sinh thì có thể thấy Prophet có nhiều lợi thế đặc biệt hữu ích cho một dự án dự báo định hướng thực tế. Một trong số đó là tính dễ sử dụng, chỉ cần một vài dòng mã và hầu như không cần thêm kỹ thuật phức tạp nào để có một đường dự báo cơ sở tốt.

Về đặc tính thời gian, xu hướng và tính thời vụ được tạo ra một cách rất tự nhiên – loại bỏ các tính năng phức tạp được yêu cầu bởi các mô hình khác như đường trung bình, độ trễ,... Một lợi thế lớn so với các mô hình tự động phức hồi (ví dụ ARIMA) là Prophet không yêu cầu chuỗi thời gian cố định, mà là một thành phần xu hướng được tạo tự nhiên.

Qua phân phân tích thì dữ liệu tuyển sinh là dữ liệu đặc thù của ngành, chỉ với sáu năm liên tục từ năm 2016 đến 2021 có cùng hình thức xét tuyển, và chỉ với sáu dòng dữ liệu thì không đủ để đáp ứng tính mùa vụ trong thuật toán prophet vì dữ liệu mùa vụ phải thể hiện qui luật lặp đi lặp lại trong chuỗi thời gian. Chúng ta cần quan sát qui luật đó trong một khoảng thời gian dài để nhận định được tính mùa vụ sẽ lặp lại theo một số chu kỳ phổ biến chẳng hạn như:

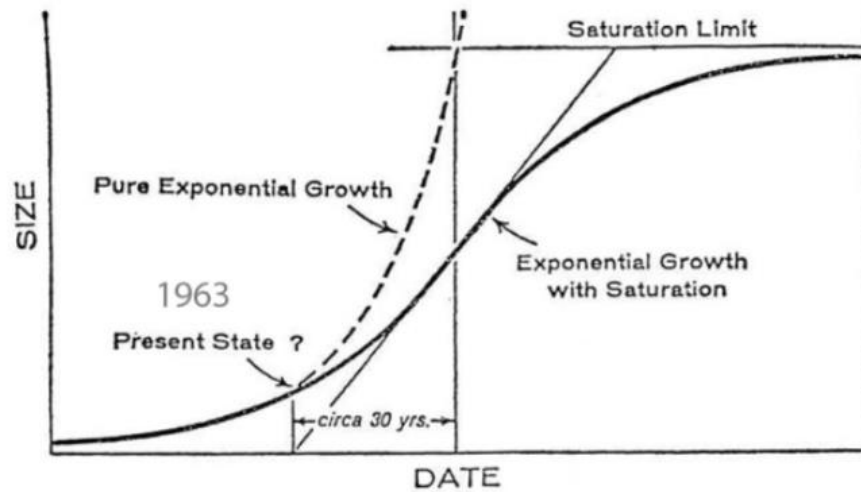
- 4 quý trong 1 năm: Thường thể hiện ở các chuỗi như GPD
- 12 tháng trong 1 năm: Các chuỗi liên quan tới doanh thu, doanh số, du lịch
- 7 ngày trong 1 tuần: Các chuỗi liên quan tới qui luật mua sắm, tiêu dùng, giải trí

Vì dữ liệu không có tính chu kỳ và dữ liệu cũng không bị ảnh hưởng bởi ngày nghỉ, ngày lễ. Bên cạnh đó prophet có thành phần xu hướng  $g(t)$  được tạo tự nhiên, không yêu cầu chuỗi thời gian cố định, nên khi áp dụng thuật toán Prophet thành phần cần quan tâm là  $g(t)$  - là hàm đại diện cho xu hướng.

$g(t)$  có thể thực hiện 2 mô hình xu hướng là mô hình Logistic growth và Linear growth. Trong mô hình Logistic growth cần biết trước điểm bão hòa và xử lý dữ liệu

tăng trưởng phi tuyến tính với độ bão hòa [7]. Tăng trưởng phi tuyến tính có bão hòa

- Giai đoạn tăng trưởng ban đầu xấp xỉ theo cấp số nhân (hình học), khi bắt đầu bão hòa, tăng trưởng chậm dần đến tuyến tính (số học) và khi đến thời gian cuối cùng trong chuỗi thời gian dự báo, tốc độ tăng trưởng sẽ dừng lại [8].



**Hình 2.4.1: Mô hình Prophet Logistic Growth**

Công thức tăng trưởng theo Logistic Growth:

$$f(x) = \frac{L}{1 + e^{-k(x-x^0)}}$$

#### Hình 2.4.2: Công thức tăng trưởng theo Logistic Growth

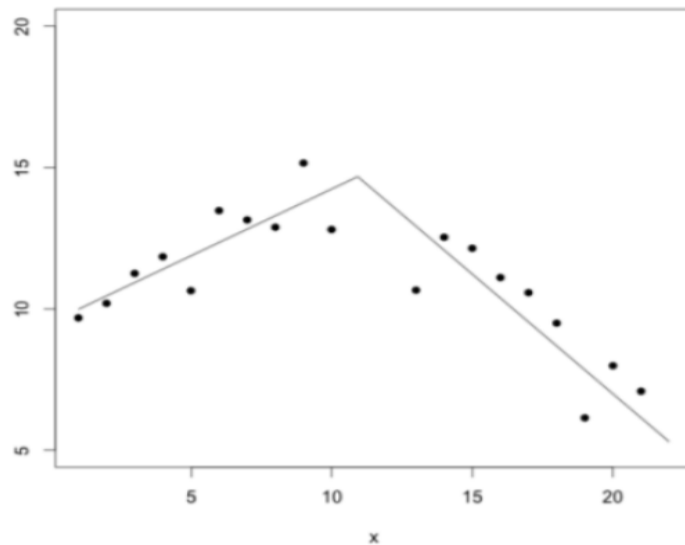
Trong đó:

- $x^0$ : giá trị của Hàm sigmoid.
- $L$ : giới hạn, giá trị tối đa của đường cong.
- $k$ : tốc độ tăng trưởng logistic, độ dốc của đường cong.

So với mô hình Logistic, Linear growth được điều chỉnh đơn giản hơn, nhưng rất hữu ích. Các phạm vi khác nhau của  $x$ , các mối quan hệ tuyến tính khác nhau xảy ra:

- Một mô hình tuyến tính duy nhất có thể không cung cấp giải thích hoặc mô tả đầy đủ.
- Breakpoints là giá trị của  $x$  khi độ dốc thay đổi.
- Giá trị của các breakpoints có thể được biết trước hoặc có thể không được biết trước khi phân tích, khi chưa biết thì breakpoints phải được ước tính

Theo mặc định Prophets sử dụng Linear growth. Với đặc điểm là tốc độ tăng trưởng không đổi thì mô hình này là phù hợp nhất khi không có sự tăng trưởng bão hòa.



**Hình 2.4.3: Mô hình tuyến tính từng mảnh**

Công thức mô hình tuyến tính từng mảnh:

$$y = \beta_0 + \beta_1 x + \beta_2 (x-c)^+ + \varepsilon$$

Trong đó:  $c$  là giá trị của breakpoints.

Có thể viết:

$$y = \begin{cases} \beta_0 + \beta_1 x & ; x \leq c \\ \beta_0 - \beta_2 c + (\beta_1 + \beta_2)x & ; x > c \end{cases}$$

Chúng ta có thể điều chỉnh các thông số này (thành phần xu hướng) trong mô hình Prophet bằng cách đặt các điểm ngắt (còn được gọi là changepoints).

Đầu vào của mô hình là một dataframe có 2 cột ds và cột y. Cột ds (viết tắt của datestamp) là cột được định dạng kiểu thời gian, thường là YYYY-MM-DD hoặc YYYY-MM-DD HH: MM: SS. Cột y phải là giá trị số thực tương ứng với mốc thời gian.

Kết quả của thuật toán Prophet là bảng chứa các thông số sau:

**Bảng 2.4.1: Bảng dữ liệu chứa các thông số của thuật toán Prophet**

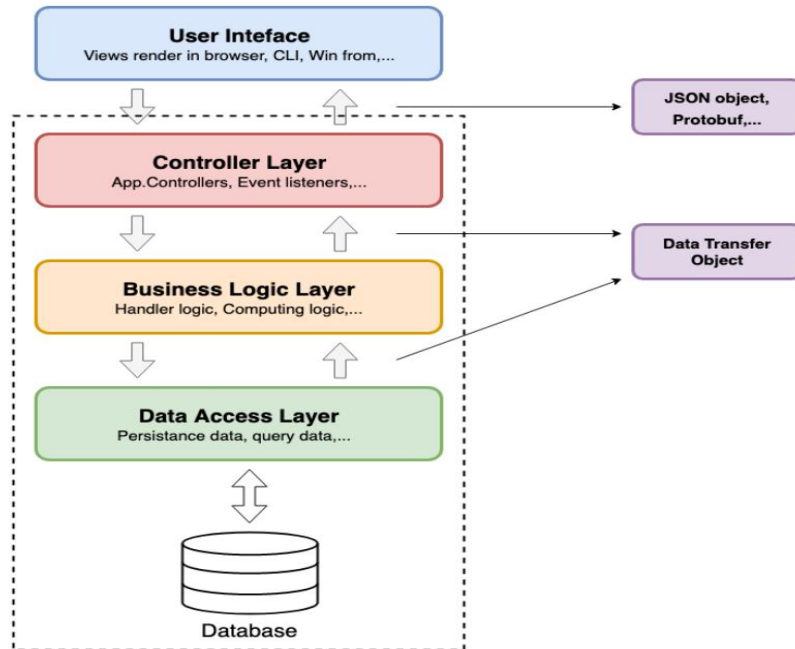
	DS	YHAT	YHAT_LOWER	YHAT_UPPER
<b>3265</b>	2017-01-15	8.211542	7.444742	8.903545
<b>3266</b>	2017-01-16	8.536553	7.847804	9.211145
<b>3267</b>	2017-01-17	8.323968	7.541829	9.035461
<b>3268</b>	2017-01-18	8.156621	7.404457	8.830642
<b>3269</b>	2017-01-19	8.168561	7.438865	8.908668

Trong đó:

- ds: đại diện cho thời gian tương ứng với giá trị dự báo được
- yhat: giá trị dự báo được
- yhat lower: giá trị lớn nhất có thể xảy ra
- yhat upper: giá trị nhỏ nhất có thể xảy ra

## 2.5. Đánh giá và lựa chọn mô hình phát triển hệ thống quản lý

Giới thiệu tổng quan về mô hình ba lớp:



**Hình 2.5.1: Mô hình phát triển hệ thống**

Trong đó:

- User interface: là phần giao diện tương tác với người dùng như web browser, winform,... Đây là nơi người dùng gửi các yêu cầu RESTful tới hệ thống.
- Controller Layer: tương tự như ở trên, là một điểm đầu vào của service, là nơi nhận các gói tin yêu cầu và phản hồi về User interface. Layer chịu trách nhiệm xử lý các logic routing, kiểm tra tham số, chuyển tiếp request,...
- Business Logic Layer: là lớp xử lý chính các business của hệ thống. Khi nhận các yêu cầu từ Controller layer, tùy vào loại yêu cầu sẽ có cách xử lý, tính toán khác nhau. Những yêu cầu cần đến dữ liệu hay thay đổi dữ liệu sẽ được lớp này đẩy xuống Data Access Layer tính toán.
- Data Access Layer: là lớp duy nhất có thể truy vấn đến database của service, layer thực hiện các thao tác có liên quan đến dữ liệu như (select, insert, update, delete,...).



Việc phân chia thành các lớp khác nhau giúp cho chúng ta dễ dàng phát triển và bảo trì hệ thống. Các lớp đảm nhận vai trò khác nhau, giảm sự phụ thuộc giữa các lớp khi hệ thống phát triển. Đồng thời các lớp được phân tách giúp chúng ta có thể tái sử dụng lại khá nhiều, tiết kiệm được thời gian xây dựng. Vì vậy mô hình ba lớp phù hợp để lựa chọn phát triển hệ thống.

## 2.6. Các IDE và ngôn ngữ lập trình sử dụng để phát triển hệ thống

IntelliJ là một IDE hỗ trợ nhiều hệ thống, đồng thời hỗ trợ mở rộng database editor và UML designer [9], cho các tính năng chạy thử UI, bảo vệ code, tích hợp Git, tích hợp kiểm soát các phiên bản, hỗ trợ nhiều công cụ phát triển và gỡ lỗi cho hầu hết những servers. Java là một trong những ngôn ngữ lập trình hướng đối tượng. Nó được sử dụng trong phát triển phần mềm cụ thể như các trang web, các ứng dụng trên windows, game hay ứng dụng trên các thiết bị di động. Vì vậy sự kết hợp của ngôn ngữ Java dùng trong IDE IntelliJ là lựa chọn thuận lợi cho việc xây dựng API backend.

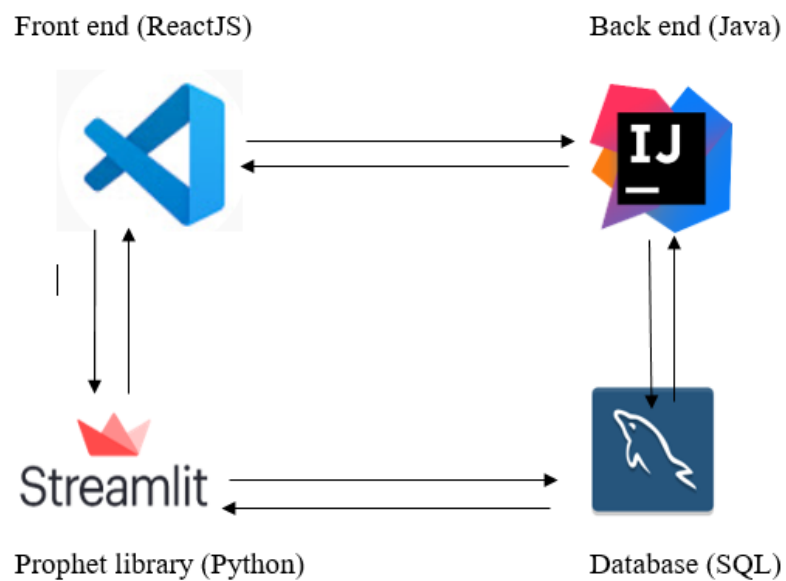
ReactJs là một thư viện Javascript được phát triển bởi Facebook[10], dùng để xây dựng giao diện người dùng. Nó giúp xây dựng các ứng dụng lớn, có nhiều thao tác phức tạp từ người dùng. Rất nhiều trang web lớn đã áp dụng React vào việc xây dựng UI như: facebook.com, wordpress.com, airbnb.com, dropbox.com,... Bên cạnh đó, thư viện này không chỉ hoạt động trên phía client, mà còn được render trên server và có thể kết nối với nhau. Sự kết hợp của reactJs và IDE Visual Code là một lựa chọn tốt để xây dựng giao diện của hệ thống

MySQL là một hệ thống quản trị cơ sở dữ liệu mã nguồn mở hoạt động theo mô hình client-server[11]. MySQL quản lý dữ liệu thông qua các cơ sở dữ liệu. Mỗi cơ sở dữ liệu có thể có nhiều bảng quan hệ chứa dữ liệu. MySQL cũng có cùng một cách truy xuất và mã lệnh tương tự với ngôn ngữ SQL, do đó luận văn chọn MySQL để xây dựng cơ sở dữ liệu cho hệ thống.

Streamlit là công cụ được xây dựng với mục đích dành cho Machine Learning Engineer[12], tạo ra giao diện web như Jupyter notebook. Điểm đặc biệt khác với Jupyter notebook là Streamlit không phải hiển thị code, giúp cho bạn có thể tạo ra

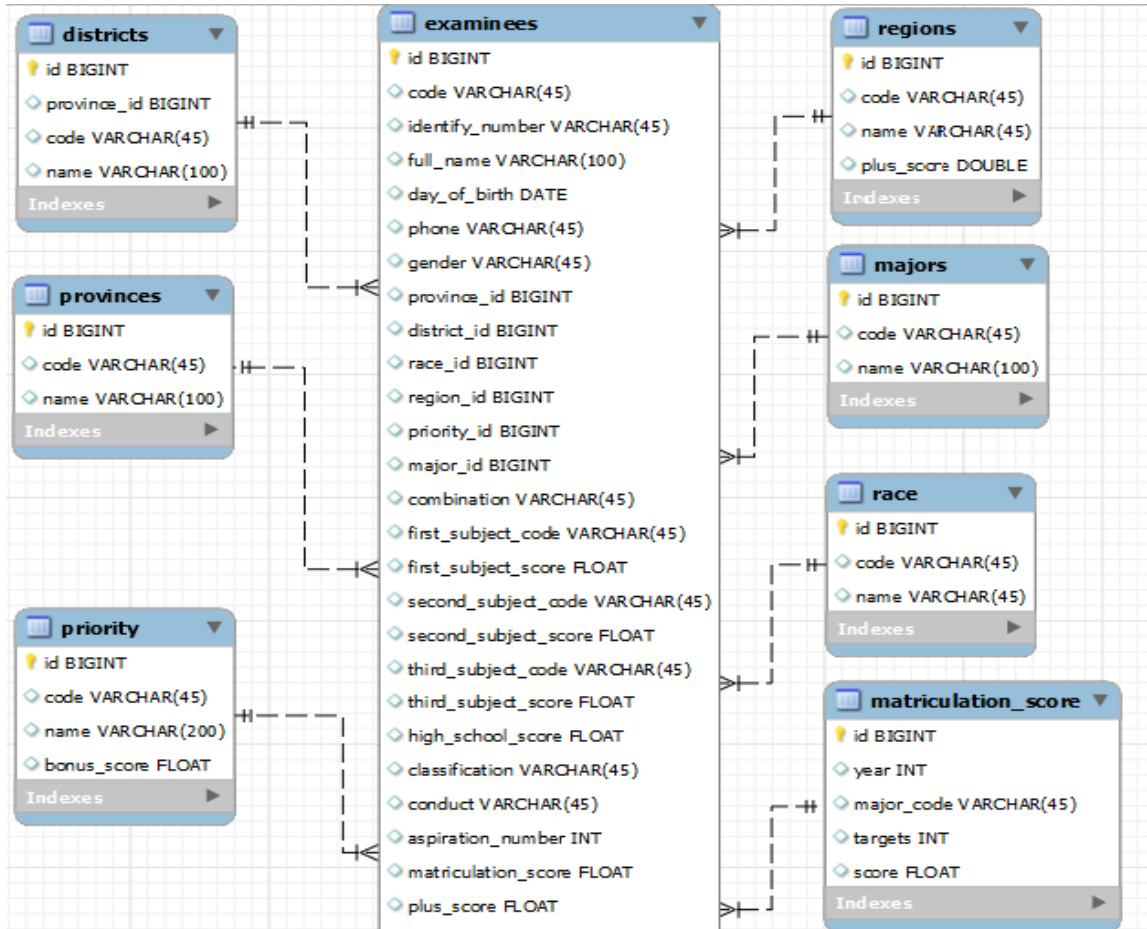
sản phẩm có tính hoàn thiện cao, là một lựa chọn thích hợp để thao tác lựa chọn và xem kết quả dự báo.

Sự kết hợp của các IDE và các ngôn ngữ lập trình trên trong mô hình lập trình ba lớp vừa giúp người lập trình quản lý source nhanh hơn, đặc biệt là công tác duy trì, phát triển hệ thống rất dễ dàng về sau. Qua quá trình phân tích những công cụ và ngôn ngữ ở trên ta xây dựng được sơ đồ phát triển hệ thống như bên dưới:



**Hình 2.6.1: Mô hình sơ đồ phát triển hệ thống**

## 2.7. Thiết kế cơ sở dữ liệu



Hình 2.7.1: Mô hình tổng quan cơ sở dữ liệu

Từ mô hình cơ sở dữ liệu ta có chi tiết các bảng trong CSDL như sau:

- Bảng area: lưu thông tin 8 vùng miền của cả nước ta:

Bảng 2.7.1: Bảng CSDL Area

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(5)	Mã vùng miền
3	name	VARCHAR(100)	Tên vùng miền của nước ta

- Bảng provinces: lưu thông tin các tỉnh theo vùng miền của nước ta:

**Bảng 2.7.2: Bảng CSDL Provinces**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(5)	Mã tỉnh
3	name	VARCHAR(100)	Tên tỉnh
4	area_id	BIGINT	Mã vùng miền

- Bảng districts: lưu thông tin các quận, huyện theo tỉnh:

**Bảng 2.7.3: Bảng CSDL Districts**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	province_id	VARCHAR(5)	Mã tỉnh
3	code	VARCHAR(5)	Mã quận, huyện
4	name	VARCHAR(100)	Tên quận, huyện

- Bảng regions: lưu thông tin khu vực và điểm cộng từng khu vực:

**Bảng 2.7.4: Bảng CSDL Regions**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(5)	Mã khu vực gồm: 1, 2, 2NT, 3
3	name	VARCHAR(100)	Tên khu vực gồm: Khu vực 1, Khu vực 2, Khu vực 2 nông thôn,...
4	plus_score	DOUBLE	Điểm cộng tương ứng với từng khu vực

- Bảng race: lưu thông tin tất cả dân tộc trên nước ta:

**Bảng 2.7.5: Bảng CSDL Race**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(5)	Mã dân tộc: DT01, DT02,...
3	name	VARCHAR(100)	Tên dân tộc

- Bảng priority: lưu thông tin đối tượng ưu tiên:

**Bảng 2.7.6: Bảng CSDL priority**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(5)	Mã đối tượng ưu tiên
3	name	VARCHAR(100)	Tên đối tượng ưu tiên
4	bonus_score	DOUBLE	Điểm cộng theo từng đối tượng ưu tiên

- Bảng majors: lưu thông tin tất cả các ngành đang tuyển sinh tại Học viện:

**Bảng 2.7.7: Bảng CSDL majors**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(5)	Mã ngành
3	name	VARCHAR(100)	Tên ngành

- Bảng examinees: lưu thông tin chi tiết của từng thí sinh/ sinh viên, bao gồm thông tin cá nhân, thông tin tuyển sinh, dữ liệu kết quả học tập:

**Bảng 2.7.8: Bảng CSDL examinees**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	code	VARCHAR(20)	Số báo danh
3	identify_number	VARCHAR(20)	Số CMND/ CCCD
4	full_name	VARCHAR(20)	Họ và tên
5	day_of_birth	DATE	Ngày tháng năm sinh
6	gender	VARCHAR(5)	Giới tính
7	province_id	BIGINT	Mã tỉnh
8	district_id	BIGINT	Mã huyện
9	race_id	BIGINT	Mã dân tộc
10	region_id	BIGINT	Mã khu vực
11	priority_id	BIGINT	Mã thứ tự ưu tiên
12	major_id	BIGINT	Mã ngành
13	combination	VARCHAR(45)	Mã tổ hợp
14	first_subject_code	VARCHAR(45)	Mã môn 1
15	first_subject_score	FLOAT	Điểm môn 1
16	second_subject_code	VARCHAR(45)	Mã môn 2
17	second_subject_score	FLOAT	Điểm môn 2
18	third_subject_code	VARCHAR(45)	Mã môn 3
19	third_subject_score	FLOAT	Điểm môn 3
20	high_school_score	FLOAT	Điểm trung bình lớp 12
21	classification	VARCHAR(45)	Học lực
22	conduct	VARCHAR(45)	Hành kiểm
23	aspiration_number	INT	Thứ tự nguyện vọng trúng tuyển
24	matriculation_score	FLOAT	Điểm trúng tuyển
25	year	DATETIME	Năm trúng tuyển
26	admission	TINYINT(1)	Đánh dấu sinh viên này có nhập học hay không
27	email	VARCHAR(255)	Đại chỉ Email
28	address	TEXT	Địa chỉ

Ngoài ra còn có 2 bảng chứa thông tin của hình thức xét tuyển kết hợp và xét tuyển dựa vào kỳ thi đánh giá năng lực:

- Bảng xettuyenkethop: lưu thông tin hình thức xét tuyển kết hợp mỗi năm của từng ngành, sau khi người dùng điền thông tin vào form thêm thông tin xét tuyển kết hợp và bấm nút “Thêm”, tất cả dữ liệu trên form sẽ được insert vào bảng 2.7.9

**Bảng 2.7.9: Bảng CSDL xettuyenkethop**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	year	DATETIME	Năm xét tuyển
3	major_id	BIGINT	Mã ngành
4	dangki	BIGINT	Số lượng sinh viên đăng ký
5	trungtuyen	BIGINT	Số lượng sinh viên trúng tuyển
6	nhaphoc	BIGINT	Số lượng sinh viên nhập học

- Bảng xettuyenkythinl: lưu thông tin hình thức xét tuyển dựa vào kỳ thi đánh giá năng lực mỗi năm của từng ngành, sau khi người dùng điền thông tin vào form thêm thông tin và bấm nút “Thêm”, tất cả dữ liệu trên form sẽ được insert vào bảng 2.7.10

**Bảng 2.7.10: Bảng CSDL xettuyenkythinl**

STT	Trường dữ liệu	Kiểu dữ liệu	Mô tả
1	id	BIGINT	Khóa chính của bảng, là mã tự tăng
2	year	DATETIME	Năm xét tuyển
3	major_id	BIGINT	Mã ngành
4	dangki	BIGINT	Số lượng sinh viên đăng ký
5	trungtuyen	BIGINT	Số lượng sinh viên trúng tuyển
6	nhaphoc	BIGINT	Số lượng sinh viên nhập học

## Chương 3: TRIỂN KHAI ỨNG DỤNG

### 3.1. Giới thiệu tổng quan về mô hình hệ thống triển khai

Mô hình hệ thống website gồm các chức năng chính là quản lý 3 hình thức xét tuyển và dự báo. Trong chức năng quản lý, có chia 3 hình thức xét tuyển:

- Xét tuyển kết quả thi THPT:
  - Quản lí danh sách.
  - Thêm danh sách trúng tuyển.
  - Thêm danh sách nhập học.
- Xét tuyển kết hợp.
- Xét tuyển dựa vào kỳ thi đánh giá năng lực

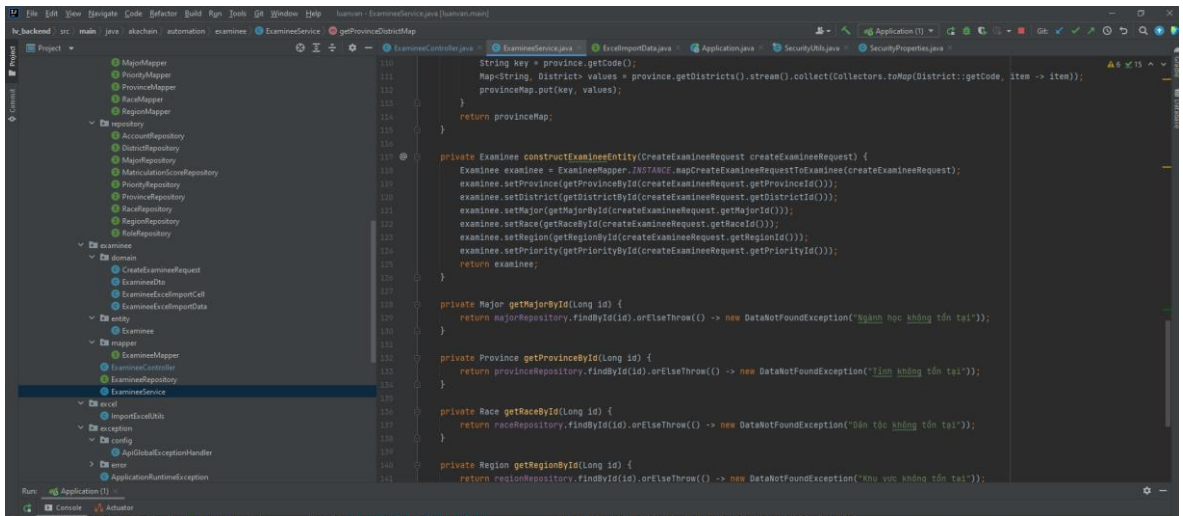
Tiếp theo là chức năng dự báo, chi tiết chức năng gồm:

- Dự báo số lượng hồ sơ trúng tuyển, số lượng sinh viên nhập học cho 3 hình thức xét tuyển nêu trên theo ngành hoặc theo vùng hoặc theo đối tượng ưu tiên,... tùy chọn ở menu.
- Xuất ra đồ thị thể hiện các thông tin thống kê, dự báo.

Sau khi áp dụng mô hình 3 lớp và xây dựng cơ sở dữ liệu, việc tiếp theo là tiến hành xây dựng phần backend. Gồm các API:

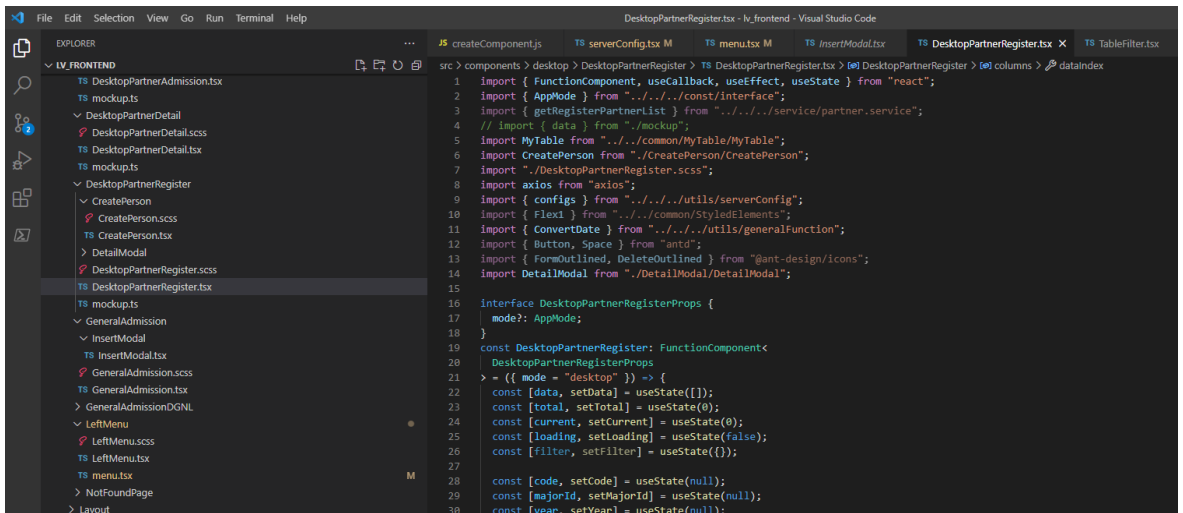
- API get token: response của API này là chuỗi accessToken, dùng làm authorization cho các API khác.
- API get by id: response của API này là json thông tin của sinh viên theo id (mã sinh viên), API này được gọi trong phần xem thông tin sinh viên trên UI.
- API create: body API là chuỗi Json thể hiện thông tin của sinh viên, khi người dùng cần thêm từng sinh viên trên giao diện, API này sẽ được gọi.
- API import: body của API import là cặp key: file tương ứng với value: “file excel cần import”, chức năng này dùng để insert hoặc update tất cả các record trong file excel xuống cơ sở dữ liệu.





Hình 3.1.1: Source code backend trên IntelliJ

Phần front end được viết bằng ngôn ngữ reactJs trên IDE visual studio code:



Hình 3.1.2: Source code font-end trên IDE visual studio code

## 3.2. Thực tế triển khai ứng dụng

Form đăng nhập: cho phép người quản lý đăng nhập vào hệ thống:

**ĐĂNG NHẬP**

\* Tài khoản

babui

\* Mật khẩu

.....

Remember me

**Đăng nhập**

**Hình 3.2.1: Giao diện form đăng nhập**

Sau khi đăng nhập thành công thì vào trang quản lý tuyển sinh:

Quản lý tuyển sinh

Quản lý danh sách

Tạo mới

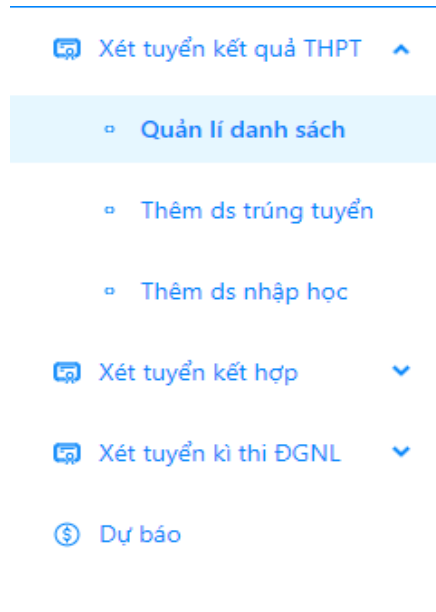
Mã sinh viên: Mã sinh viên Ngành Khu vực Năm Nhập học: Tìm kiếm

Tổng số sinh viên: 3723 sinh viên

STT	SBD	Họ tên	Ngày sinh	Giới tính	Ngành trúng tuyển	Năm trúng tuyển	Tổ hợp	Nguyên vọng	Điểm trúng tuyển	Điểm cộng	Mã môn 1	Điểm môn 1	Mã môn 2	Điểm
1	40014088	ĐẶNG LÊ TUẤN ANH	30/11/2002	Nam	Công nghệ đa phương tiện	2020	A00	5	24.9	0	TO	8.4	LI	
2	44005414	TÀ THỊ HOÀNG ANH	10/02/2002	Nữ	Công nghệ đa phương tiện	2020	D01	1	23.8	0	TO	8	VA	6
3	44009298	LƯƠNG THỊ NGỌC ANH	29/06/2002	Nữ	Công nghệ đa phương tiện	2020	A00	1	23.95	0	TO	7.2	LI	7
4	53011001	VŨ ĐOÀN GIA BẢO	04/01/2002	Nam	Công nghệ đa phương tiện	2020	A00	1	25.35	0	TO	8.6	LI	
5	57009947	PHẠM THỊ HỒNG CẨM	27/04/2002	Nữ	Công nghệ đa phương tiện	2020	A00	1	25.3	0	TO	8.8	LI	5
6	02026367	NGUYỄN TÂM CHIẾN	26/07/2002	Nam	Công nghệ đa phương tiện	2020	A01	1	26	0	TO	9.2	LI	

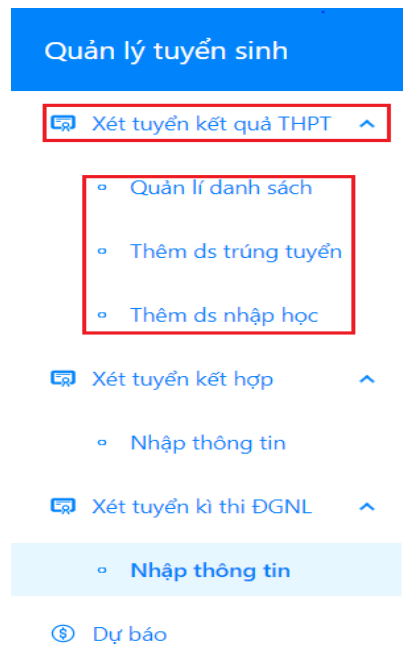
**Hình 3.2.2: Giao diện trang quản lý tuyển sinh**

Giới thiệu các chức năng trong menu của trang quản lý quản lý:



**Hình 3.2.3: Giao diện các chức năng trong menu của trang quản lý tuyển sinh**

Menu gồm thông tin 3 hình thức xét tuyển và mục dự báo. Đầu tiên là mục “Xét tuyển kết quả THPT” gồm 3 chức năng bên trong:



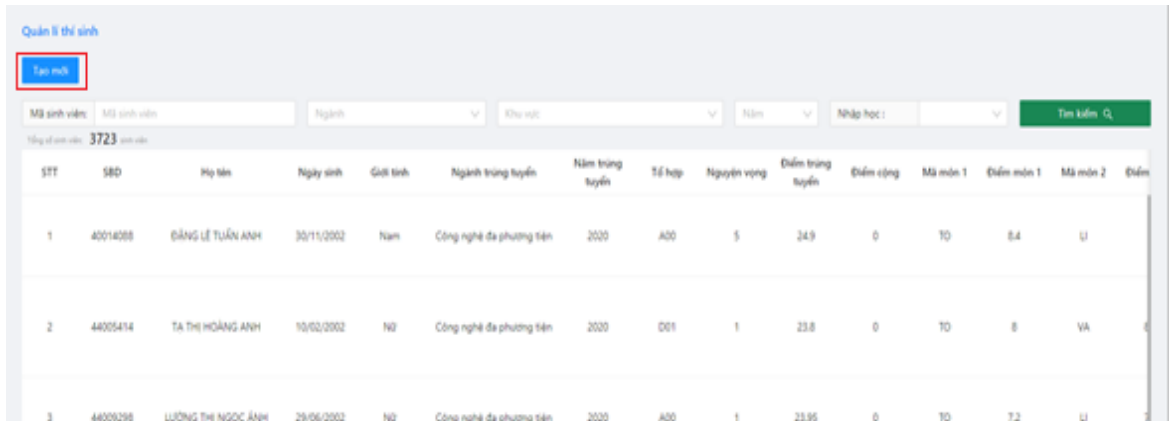
**Hình 3.2.4: Giao diện của mục “Xét tuyển kết quả THPT”**

Chức năng đầu tiên của mục “Xét tuyển kết quả THPT” là “Quản lý danh sách”, người dùng có thể xem dữ liệu, và tìm kiếm dữ liệu mong muốn xem, có thể tìm kiếm theo số báo danh, theo ngành, theo năm, theo khu vực,...

STT	SBD	Họ tên	Ngày sinh	Giới tính	Ngành trúng tuyển	Năm trúng tuyển	Tổ hợp	Nguyên vọng	Điểm trúng tuyển	Điểm cộng	Mã môn 1	Điểm môn 1	Mã môn 2	Điểm
1	40014088	ĐẶNG LÊ TUẤN ANH	30/11/2002	Nam	Công nghệ đa phương tiện	2020	A00	5	24.9	0	TO	8.4	LI	
2	44005414	TA THI HOÀNG ANH	10/02/2002	Nữ	Công nghệ đa phương tiện	2020	D01	1	23.8	0	TO	8	VA	8
3	44009298	LƯƠNG THI NGOC ANH	29/06/2002	Nữ	Công nghệ đa phương tiện	2020	A00	1	23.95	0	TO	7.2	LI	7

**Hình 3.2.5: Giao diện chức năng “Quản lý danh sách” thí sinh**

- Ngoài ra, người dùng có thể tạo thêm dữ liệu khi có nhu cầu bổ sung dữ liệu, để tạo mới, chọn nút “Tạo mới”:



The screenshot shows a web interface for student management. At the top left, there is a blue button labeled 'Tạo mới' (Create new), which is highlighted with a red rectangular box. Below this button is a search bar with several input fields: 'Mã sinh viên' (Student ID), 'Ngày' (Date), 'Khóa học' (Course), 'Năm' (Year), and 'Nhập học' (Enrollment). To the right of these fields is a green 'Tìm kiếm' (Search) button. Below the search bar, there is a table with the following columns: STT, SBD, Họ tên, Ngày sinh, Giới tính, Ngành trình độ, Năm trình độ, Tổ hợp, Nguyên vọng, Điểm trung bình, Điểm cộng, Mã môn 1, Điểm môn 1, Mã môn 2, and Điểm. The table contains three rows of student data.

STT	SBD	Họ tên	Ngày sinh	Giới tính	Ngành trình độ	Năm trình độ	Tổ hợp	Nguyên vọng	Điểm trung bình	Điểm cộng	Mã môn 1	Điểm môn 1	Mã môn 2	Điểm
1	4001408	ĐẶNG LÊ TUẤN ANH	30/11/2002	Nam	Công nghệ đa phương tiện	2020	A00	5	24.9	0	T0	8.4	LI	
2	44005414	TA THỊ HOÀNG ANH	10/02/2002	Nữ	Công nghệ đa phương tiện	2020	D01	1	23.8	0	T0	8	VA	
3	44009298	LƯƠNG THỊ NGỌC ANH	29/06/2002	Nữ	Công nghệ đa phương tiện	2020	A00	1	23.95	0	T0	7.2	LI	

**Hình 3.2.6: Giao diện chức năng chọn nút “Tạo mới”**

- Sau khi nhấn “Tạo mới” thì form nhập thông tin sẽ hiển thị và nhập đầy đủ thông tin theo mẫu, sau đó chọn nút “Tạo mới”

The image shows a web form titled "Tạo mới" (Create New) with a close button (X) in the top right corner. The form contains the following fields:

- Mã thí sinh:
- Họ và tên:
- Ngày sinh:
- Giới tính:
- Số điện thoại:
- Mã tổ hợp:
- Phân loại:
- Hành kiểm:
- Nhập học:
- CMND/CCCD:
- Ngành:
- Tỉnh:
- Quận huyện:
- Khu vực:
- Dân tộc:
- Đối tượng ưu tiên:
- Mã môn 1:
- Điểm môn 1:
- Mã môn 2:
- Điểm môn 2:
- Mã môn 3:
- Điểm môn 3:
- Điểm TB THPT:
- Nguyên vọng thứ:
- Điểm trúng tuyển:
- Điểm cộng:

At the bottom of the form is a blue button labeled "Tạo mới".

**Hình 3.2.7: Giao diện form thêm mới thông tin một thí sinh**

- Một thông báo “Tạo thành công” sẽ xuất hiện và hiện lên hàng đầu tiên trong bảng nếu nhập đầy đủ dữ liệu hoặc một thông báo lỗi thiếu dữ liệu hay dữ liệu không đúng định dạng nếu người dùng nhập sai thông tin.

- Ngoài ra hệ thống cũng cho phép người dùng chỉnh sửa thông tin hoặc xóa thông tin trên giao diện. Để sửa hoặc xóa, người dùng chọn số báo danh của sinh viên cần xóa, hệ thống sẽ hiện lên form để sửa hoặc xóa dữ liệu của số báo danh đã chọn. Nếu cần sửa thông tin thì nhập thông tin cần sửa sau đó bấm nút “Lưu thông tin thay đổi” hoặc xóa thông tin thì bấm nút “Xóa”:

The screenshot shows a web form titled "Thông tin chi tiết" (Detailed Information) with a close button (X) in the top right corner. The form contains the following fields and values:

- \* Họ tên: PHAM TRƯỜNG AN
- \* Ngày sinh: 10/12/1998
- \* Giới tính: Nam
- \* Ngành trúng tuyển: Quản trị kinh doanh
- Địa chỉ: Trường THCS và THPT Đông Du, Hẻm 129 Nguyễn Chí Thanh, Khối 8, Phường Tân An, Tp. Buôn Ma Thuột, Đắk Lắk
- \* Số điện thoại: 0946233777
- Email: (empty field)
- \* Nhập học: Có

At the bottom of the form, there are two buttons: a red button labeled "Xóa" (Delete) and a blue button labeled "Lưu thông tin thay đổi" (Save information change).

**Hình 3.2.8: Giao diện update thông tin một thí sinh**

Chức năng thứ hai trong mục “Xét tuyển kết quả THPT” là “Thêm ds trúng tuyển”:

**Hình 3.2.9: Giao diện “Thêm ds trúng tuyển”**

Chức năng này để người dùng import file excel kết quả trúng tuyển vào hệ thống, và dữ liệu sẽ được ghi vào cơ sở dữ liệu. Để import, cần chọn năm (sẽ có từ năm 2016 đến năm 2021 khi chọn combobox), sau đó cần chọn file excel danh sách trúng tuyển và chọn nút “Import”. File import “Thêm ds trúng tuyển” cần phải đúng định dạng như hình bên dưới:

STT	SBD	Họ tên	Ngày sinh	Giới tính	Mã ngành trúng tuyển	Tên ngành trúng tuyển	Mã tổ hợp trúng tuyển	Thứ tự NV trúng tuyển	Điểm trúng tuyển	Điểm khuyến khích	Môn 1	Môn 2	Môn 3	ĐT UT	KV UT	Năm tốt nghiệp THPT	Học lực	Hạng kiểm		
1	40014088	ĐẶNG LÊ TUẤN ANH	30/11/2002	Nam	7329001	Công nghệ đa phương tiện	A00	5	24.9		TO	8.4	LI	8	HO	7.75	1	2020	Khá	Tốt
2	44005414	TẠ THỊ HOÀNG ANH	10/02/2002	Nữ	7329001	Công nghệ đa phương tiện	D01	1	23.8		TO	8	VA	8.75	N1	6.8	2	2020	Khá	Tốt
3	44009298	LƯƠNG THỊ NGỌC ANH	29/06/2002	Nữ	7329001	Công nghệ đa phương tiện	A00	1	23.95		TO	7.2	LI	7.75	HO	8.5	2NT	2020	Khá	Tốt
4	53011001	VÕ ĐOÀN GIA BẢO	04/01/2002	Nam	7329001	Công nghệ đa phương tiện	A00	1	25.35		TO	8.6	LI	8	HO	8.25	2NT	2020	Giỏi	Tốt

**Hình 3.2.10: Format định dạng file Excel của “Thêm ds trúng tuyển” (Từ cột A tới cột U)**

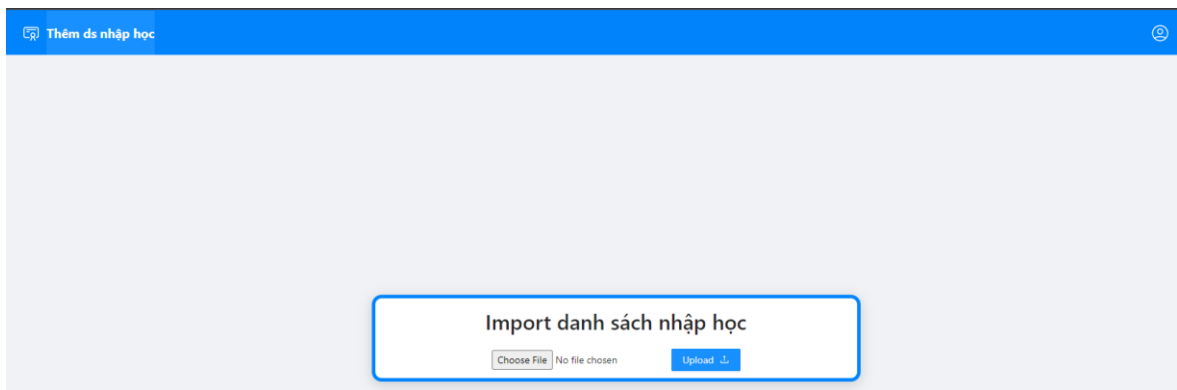


	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
1	<b>DANH SÁCH THÍ SINH TRÚNG TUYỂN</b>															
2																
3	Học lực	Hạng kiểm	Điểm TB lớp 12	TN cao đẳng	TN trung cấp	HK Tỉnh	HK Huyện	HK Xã	Mã tỉnh lớp 12	Mã trường lớp 12	Điện thoại	Địa chỉ Email	Địa chỉ nhận giấy báo	Not sinh	Dân tộc	Số CMND
4	Khá	Tốt	7.7			40 - Đắk Lắk	01 - Thành phố Buôn Ma Thuột	03-Phường Tân An	40	074	0946233777		Trường THCS và THPT Đông Du, Hẻm 129 Nguyễn Chí Thanh, Khối 8, Phường Tân An, Tp. Buôn Ma Thuột, Đắk Lắk	ĐẮK LẮK	Kinh	241864654
5	Khá	Tốt	7.7			44 - Bình Dương	11 - Thành phố Dĩ An		44	081	0798294608	sunymon102@gmail.com	Tạ Thị Hoàng Anh - SĐT 0798294608 - Lớp 12C12 - Trường THPT Bình An, KP Bình Thung 1, P Bình An, TP Dĩ An, Tỉnh Bình Dương	BÌNH DƯƠNG	Kinh	231440623
6	Khá	Tốt	7.6			44 - Bình Dương	06 - Huyện Phú Giáo		44	030	0353934169	luongthangscah12a@taysonsgbinhduong.edu.vn	Lương Thị Ngọc Anh - Lớp12A, Trường THPT Tây Sơn - Ấp 2, Tân Long, Phú Giáo, Bình Dương	BÌNH DƯƠNG	Kinh	281292785
7	Giỏi	Tốt	8.1			53 - Tiền Giang	07 - Huyện Gò Công Tây		53	022	0975478325	vobao+12002@gmail.com	Có Đoàn Lê Bạch Thảo, giáo viên Trường TH Nguyễn Văn Trứ, xã Thạnh Nhựt, huyện Gò Công Tây, Tiền Giang.	TỈNH TIỀN GIANG	Kinh	312502909

**Hình 3.2.11: Format định dạng file Excel tiếp theo của “Thêm ds trúng tuyển” (Từ cột V tới cột AJ)**

Nếu file không đúng định dạng, hệ thống sẽ báo lỗi định dạng file không đúng, hoặc import nhưng không chọn file sẽ báo lỗi chưa chọn file. Sau khi có kết quả trúng tuyển, sinh viên sẽ liên hệ với học viên để làm thủ tục và hồ sơ nhập học, thông tin về sinh viên nhập học được lưu ở một file excel khác gồm các cột: STT, Số báo danh, Họ và tên. Để quản lý danh sách sinh viên nhập học thì sẽ có chức năng là “Thêm ds nhập học”.

Chức năng thứ ba “Thêm ds nhập học” giúp cập nhật cơ sở dữ liệu thông tin sinh viên nhập học:



**Hình 3.2.12: Giao diện chức năng “Thêm ds nhập học”**

Chức năng này cũng thao tác import file excel giống như chức năng “Thêm ds trúng tuyển”, nhưng format định dạng dữ liệu import của “Thêm ds nhập học” sẽ khác so với format định dạng dữ liệu của “Thêm ds trúng tuyển”. File import “Thêm ds nhập học” cần phải đúng định dạng như hình bên dưới:

	A	B	C	D
1				
2	DANH SÁCH SINH VIÊN HOÀN TẤT THỦ TỤC NỘP HỌC PHÍ, NHẬP HỌC, ...			
3	<b>KHÓA 2020</b>			
4				
5	<b>Số TT</b>	<b>Số báo danh</b>	<b>Họ và Tên sinh viên</b>	
6	1	02051360	TRẦN GIA LONG	
7	2	35009705	VÕ DUY TÙNG	
8	3	36001580	PHẠM THỊ NGA	
9	4	61001461	NGUYỄN PHƯỚC DUY THỊNH	
10	5	52002654	ĐỖ THU GIANG	
11	6	43002281	NGUYỄN HOÀNG THIÊN	
12	7	49006719	HỒ GIA HUY	

**Hình 3.2.13: Format định dạng file Excel của “Thêm ds nhập học”**

Tiếp theo trong menu là mục “Xét tuyển kết hợp”

Quản lý tuyển sinh Nhập thông tin

- Xét tuyển kết quả THPT
  - Quản lý danh sách
  - Thêm ds trúng tuyển
  - Thêm ds nhập học
  - Xét tuyển kết hợp**
  - Nhập thông tin
- Xét tuyển kì thi ĐGNL
- Dự báo

STT	Năm học	Chuyên ngành	Số lượng đăng kí	Số lượng trúng tuyển	Số lượng nhập học
1	2020	Công nghệ đa phương tiện	100	80	70
2	2021	Công nghệ đa phương tiện	100	80	10
3	2022	Công nghệ đa phương tiện	90	70	9
4	2023	Công nghệ đa phương tiện	120	80	15

**Hình 3.2.14: Giao diện của mục “Xét tuyển kết hợp”**

Do hình thức xét tuyển kết hợp này chỉ có ở năm 2020 và 2021 trở đi, nên chức năng này cho phép người dùng nhập dữ liệu tạm để có dữ liệu phục vụ cho chức năng dự báo.

Nhập thông tin					
Thêm xét tuyển kết hợp					
STT	Năm học	Chuyên ngành	Số lượng đăng kí	Số lượng trúng tuyển	Số lượng nhập học
1	2020	Công nghệ đa phương tiện	100	80	70
2	2021	Công nghệ đa phương tiện	100	80	10
3	2022	Công nghệ đa phương tiện	90	70	9
4	2023	Công nghệ đa phương tiện	120	80	15

**Hình 3.2.15: Giao diện form “Nhập thông tin” của hình thức “Xét tuyển kết hợp”**

Để nhập thông tin “Xét tuyển kết hợp” thì nhấn vào nút “Thêm xét tuyển kết hợp” sẽ hiển thị ra form “Thông tin chi tiết” và nhập vào các thông tin như form, sau đó bấm nút “Thêm” thì một dòng dữ liệu về hình thức “Xét tuyển kết hợp” sẽ lưu vào và hiển thị lên giao diện.

Thông tin chi tiết ×

**Xét tuyển kết hợp**

\* Năm:

\* Ngành:

\* Số lượng đăng kí:

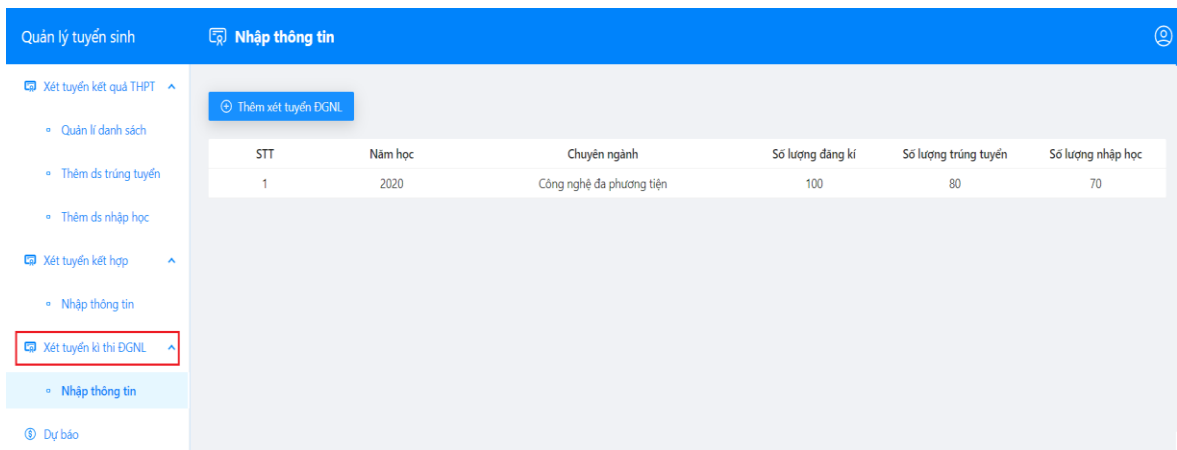
\* Số lượng trúng tuyển:

\* Số lượng nhập học:

Thêm

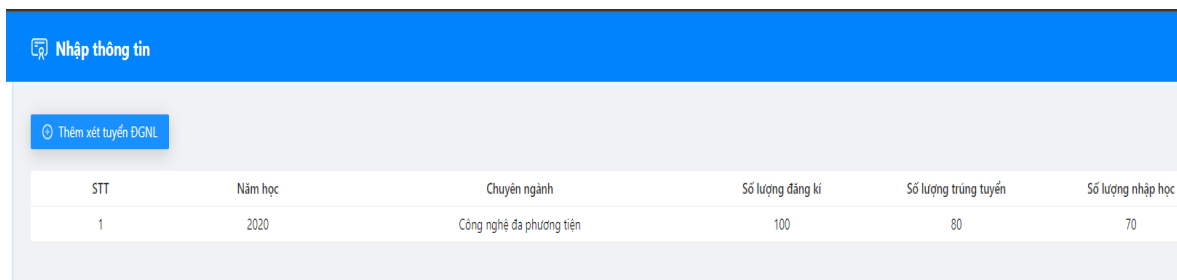
**Hình 3.2.16: Giao diện form “Thông tin chi tiết” của “Xét tuyển kết hợp”**

Tiếp theo trong menu là mục “Xét tuyển kì thi ĐGNL”



**Hình 3.2.17: Giao diện của mục “Xét tuyển kì thi ĐGNL”**

Tương tự như mục “Xét tuyển kết hợp” thì là chức năng này cũng chỉ do phép người dùng nhập thông tin cho hình thức xét tuyển dựa vào kỳ thi đánh giá năng lực.



**Hình 3.2.18: Giao diện form “Nhập thông tin” của hình thức “Xét tuyển kì thi ĐGNL”**

Để nhập thông tin “Xét tuyển kì thi ĐGNL” thì cũng giống như nhập thông tin của “Xét tuyển kết hợp” là nhấn vào nút “Thêm xét tuyển ĐGNL” sẽ hiển thị ra form “Thông tin chi tiết” và nhập vào các thông tin như form, sau đó bấm nút “Thêm” thì một dòng dữ liệu về hình thức “Xét tuyển kì thi ĐGNL” sẽ lưu vào và hiển thị lên giao diện.

**Thông tin chi tiết**

**Xét tuyển ĐGNL**

\* Năm:

\* Ngành:

\* Số lượng đăng kí:

\* Số lượng trúng tuyển:

\* Số lượng nhập học:

**Hình 3.2 19: Giao diện form “Thông tin chi tiết” của “Xét tuyển kì thi ĐGNL”**

Mục cuối cùng trong menu là chức năng “Dự báo”

Quản lý tuyển sinh | Quản lý danh sách

Xét tuyển kết quả THPT  
Xét tuyển kết hợp  
Xét tuyển kì thi ĐGNL  
**Dự báo**

Quản lý thí sinh

Tạo mới

Mã sinh viên:  Mã sinh viên:  Ngành:  Khu vực:  Năm:  Nhập học:

Tổng số sinh viên: 5449 sinh viên

STT	SBD	Họ tên	Ngày sinh	Giới tính	Ngành trúng tuyển	Năm trúng tuyển	Tổ hợp	Nguyên vọng	Điểm trúng tuyển	Điểm c
1	HUI000089	PHAM TRƯỜNG AN	10/12/1998	Nam	Quản trị kinh doanh	2016	A00	0	22.25	0

**Hình 3.2.20: Giao diện của mục “Dự báo” ở trang quản lý**

Sau khi chọn mục “Dự báo” hệ thống sẽ mở một trang mới để người dùng thao tác trên chức năng dự báo này. Chức năng dự báo cũng gồm có 3 mục như trang quản lý: “Xét tuyển dựa vào kết quả thi THPT”, “Xét tuyển dựa vào kì thi đánh giá năng lực”, “Xét tuyển kết hợp”.



## DỰ BÁO SỐ LIỆU TUYỂN SINH ĐẠI HỌC

TẠI HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Xét tuyển dựa vào kết quả thi THPT

Xét tuyển dựa vào kì thi đánh giá năng lực

Xét tuyển kết hợp

### Xét tuyển dựa vào kết quả thi THPT

✓ Chọn xem dữ liệu theo:

Ngành|

#### Xem dữ liệu theo Ngành

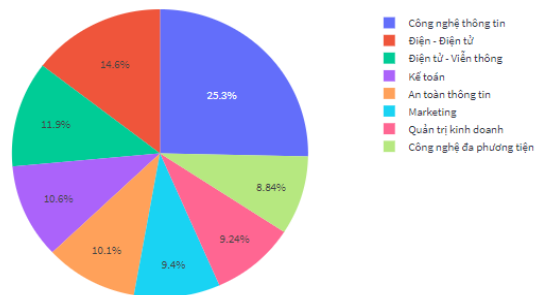
✓ Xem dữ liệu từng ngành theo năm:

2016

✓ Chọn đối tượng theo ngành:

Trúng tuyển

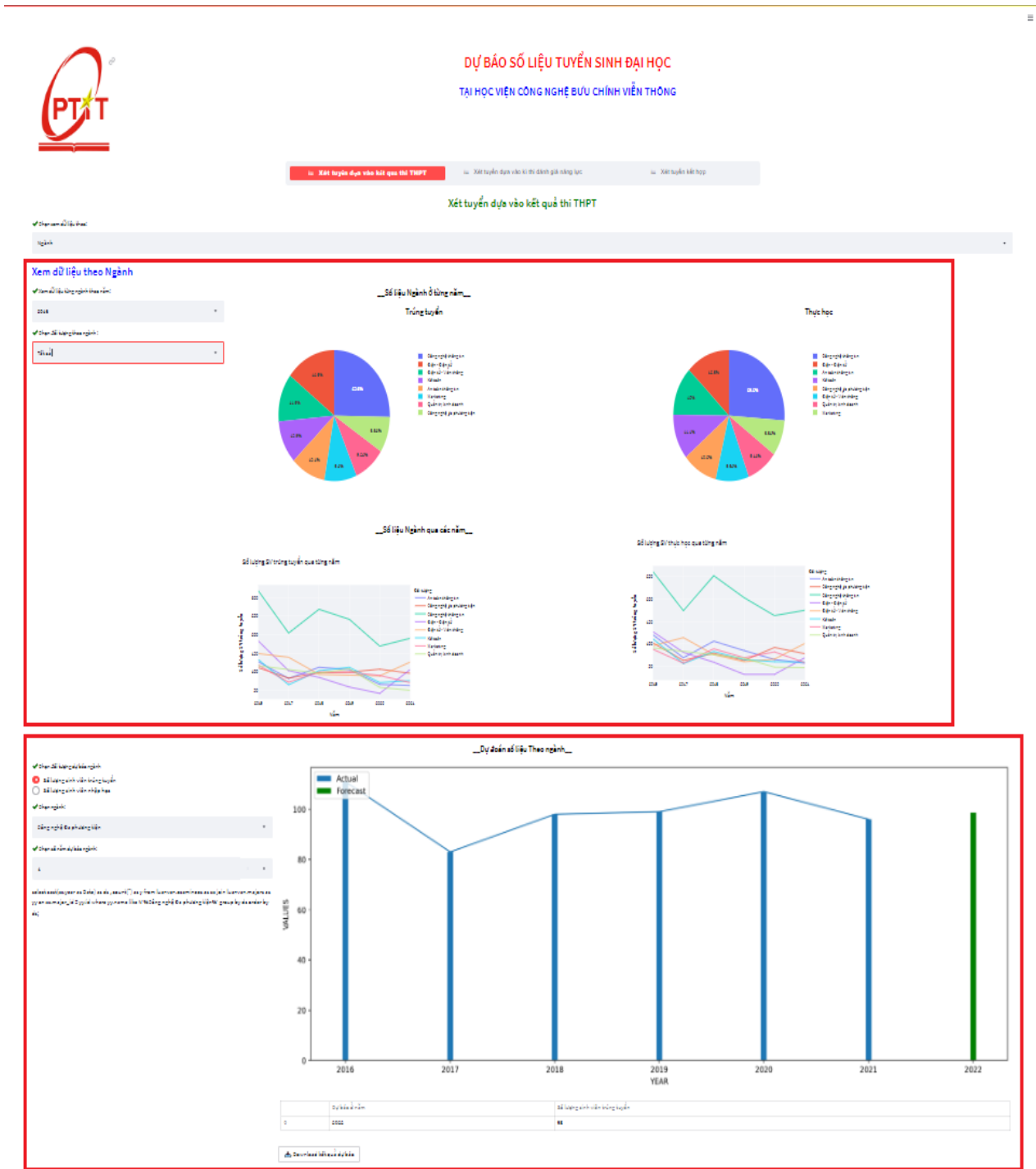
\_\_Số liệu Ngành ở từng năm\_\_



**Hình 3.2.21: Giao diện tổng quan gồm các mục của trang dự báo**

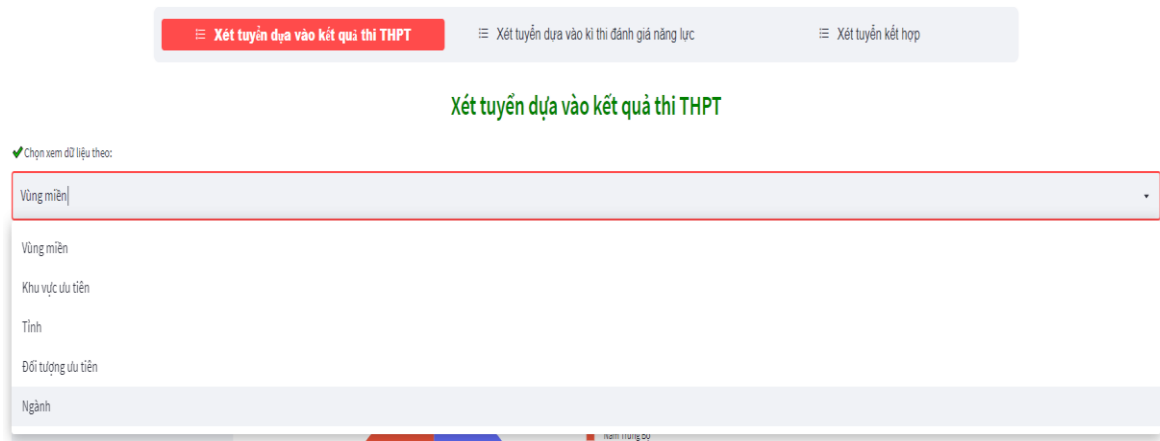
Ở mỗi mục của trang dự báo đều có chung chức năng xem dữ liệu và dự báo số liệu theo từng hình thức tuyển sinh.

- Đầu tiên là mục dự báo “Xét tuyển dựa vào kết quả thi THPT”



**Hình 3.2.22: Giao diện tổng quan xem dữ liệu và dự báo của mục “Xét tuyển dựa vào kết quả thi THPT”**

Ở mục này thì người dùng có thể xem và dự báo theo nhiều tiêu chí như (vùng miền, khu vực ưu tiên, ngành,...) và hiển thị dưới dạng biểu đồ tròn. Nên đầu tiên người dùng phải chọn xem dữ liệu theo một tiêu chí trong các tiêu chí như: vùng miền, khu vực ưu tiên, tỉnh, đối tượng ưu tiên, ngành,...



**Hình 3.2.23: Giao diện lựa chọn một tiêu chí để xem dữ liệu và dự báo**

Ở từng tiêu chí được chọn sẽ cho phép người dùng lựa chọn thêm là xem dữ liệu theo các tiêu chí con khác như: xem dữ liệu theo từng năm hoặc qua các năm của các đối tượng (trúng tuyển, nhập học, tất cả).

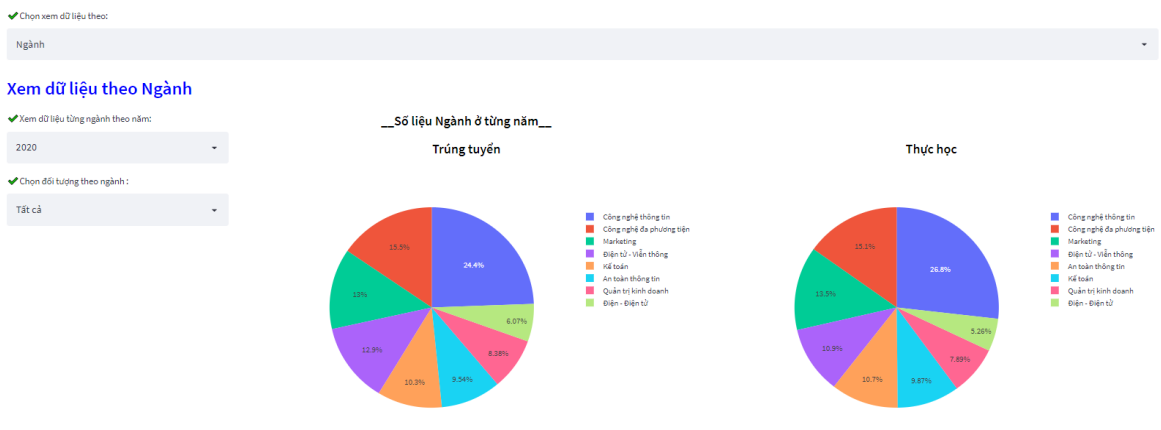
Ví dụ ở đây ta chọn xem dữ liệu theo Ngành thì có thể chọn thêm xem dữ liệu theo từng năm của từng đối tượng (trúng tuyển, nhập học,...) hoặc của tất cả đối tượng.



### Xét tuyển dựa vào kết quả thi THPT



**Hình 3.2.24: Giao diện xem dữ liệu theo Ngành năm 2016 của đối tượng trúng tuyển**



**Hình 3.2.25: Giao diện xem dữ liệu theo Ngành năm 2020 của tất cả đối tượng**

Hoặc là xem dữ liệu theo Ngành qua các năm của từng đối tượng (trúng tuyển, nhập học,...) hoặc của tất cả đối tượng.

\_\_Số liệu Ngành qua các năm\_\_

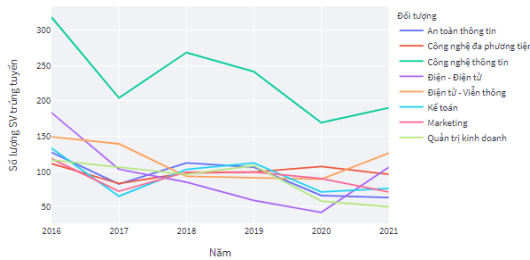
Số lượng SV trúng tuyển qua từng năm



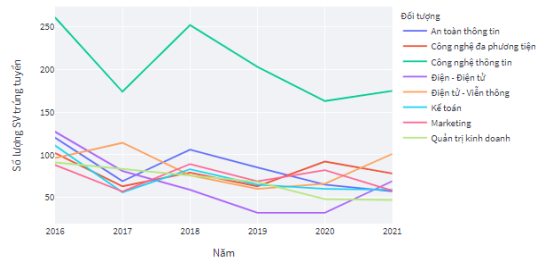
**Hình 3.2.26: Giao diện xem dữ liệu theo Ngành qua các năm của đối tượng trúng tuyển**

\_\_Số liệu Ngành qua các năm\_\_

Số lượng SV trúng tuyển qua từng năm

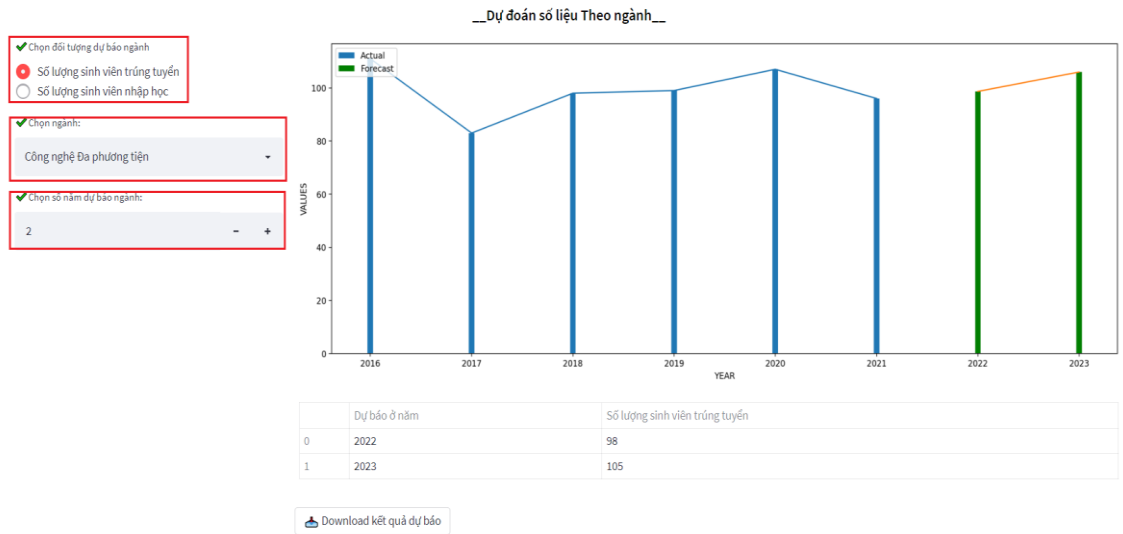


Số lượng SV thực học qua từng năm



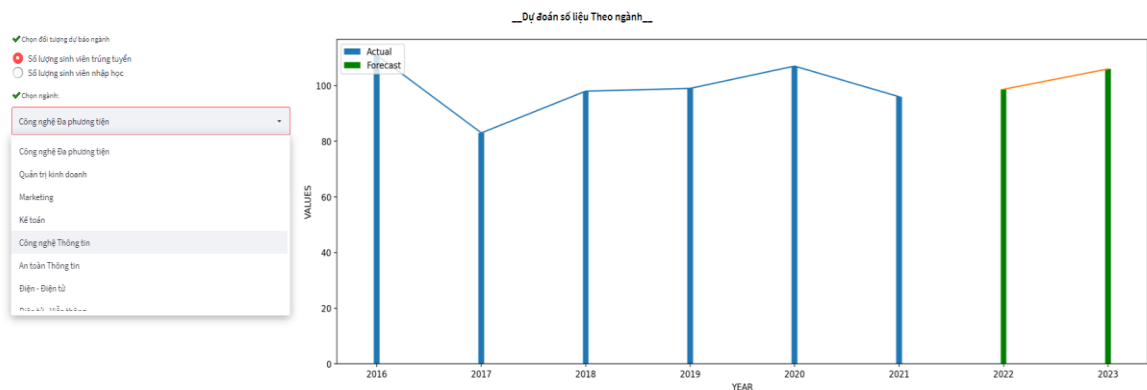
**Hình 3.2.27: Giao diện xem dữ liệu theo Ngành qua các năm của tất cả đối tượng**

Tiếp theo là phần dự báo thi tùy ở từng tiêu chí (vùng miền, ngành,...) mà người dùng đã chọn ở bước đầu tiên thì sẽ cho phép chọn tiếp số năm dự báo sắp tới và chọn đối tượng (trúng tuyển, nhập học,...) theo một lĩnh vực nhỏ của tiêu chí đó ví dụ như ngành gồm có (Công nghệ Đa phương tiện, Quản trị kinh doanh, Marketing,...), tương tự nếu tiêu chí Vùng (Bắc Bộ, Đông Nam Bộ,...).



**Hình 3.2.28: Giao diện tổng quan lựa chọn số năm, chọn đối tượng dự báo theo một ngành nào đó**

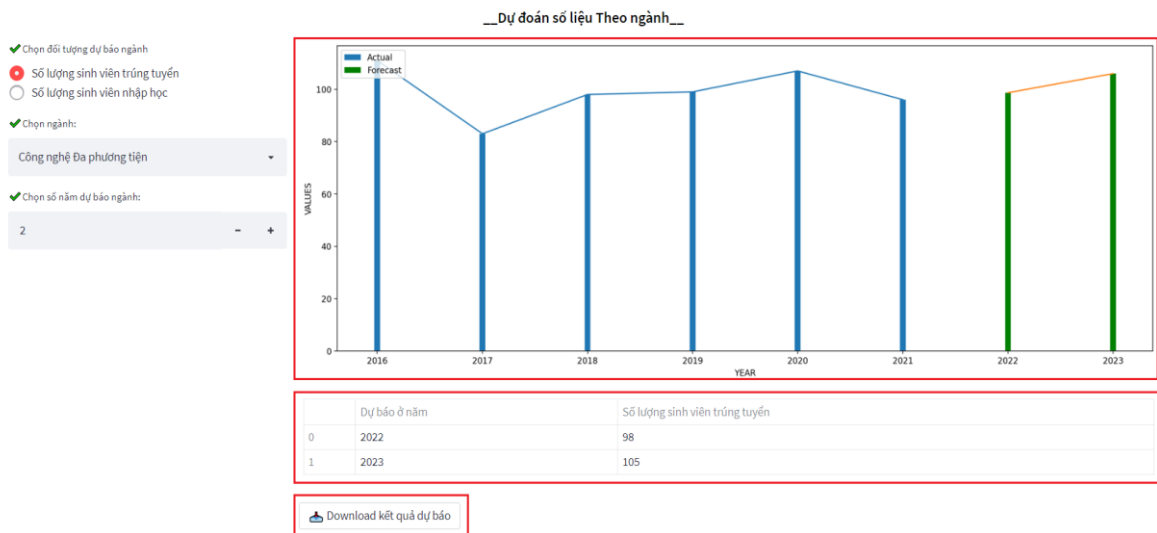
Khi đối tượng dự báo là số lượng sinh viên trúng tuyển hoặc nhập học thì người dùng có thể xem chi tiết cho từng ngành.



**Hình 3.2.29: Giao diện chọn dự báo một ngành trong các ngành của từng đối tượng**

Sau khi chọn xong thì hệ thống sẽ chạy và kết quả dự báo sẽ được hiện ra bên phải menu gồm có:

- Biểu đồ thể hiện số liệu và đường tăng trưởng theo dữ liệu thực và dữ liệu dự báo
- Bảng thống kê số liệu dự báo ở các năm tiếp theo




**Hình 3.2.30: Giao diện kết quả dự báo hai năm tới (2022 và 2023) của đối tượng trúng tuyển theo ngành Công nghệ Đa phương tiện của mục “Xét tuyển dựa vào kết quả thi THPT”**

Ngoài ra người dùng còn có thể tải kết quả dự báo về, khi tải về dữ liệu sẽ được hiển thị trong file excel theo format như hình bên dưới:

	A	B	C	D	E	F
1	Dự báo ở năm	Số lượng sinh viên trúng tuyển				
2	2022	98				
3	2023	105				
4						
5						


**Hình 3.2.31: Giao diện format file excel kết quả dự báo khi tải về**

Tiếp theo là mục dự báo “Xét tuyển kết hợp” và “Xét tuyển dựa vào kì thi đánh giá năng lực”. Khác với mục “Xét tuyển dựa vào kết quả thi THPT” thì ở hai mục này sẽ giống nhau là đều chỉ cho xem dữ liệu theo một ngành nào đó biểu diễn dưới dạng biểu đồ cột gồm các đối tượng (đăng ký, trúng tuyển, nhập học) và dự báo dữ liệu theo từng đối tượng của một ngành nào đó.



## DỰ BÁO SỐ LIỆU TUYỂN SINH ĐẠI HỌC

### TẠI HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



☰ Xét tuyển dựa vào kết quả thi THPT

☰ Xét tuyển dựa vào kì thi đánh giá năng lực

☰ **Xét tuyển kết hợp**

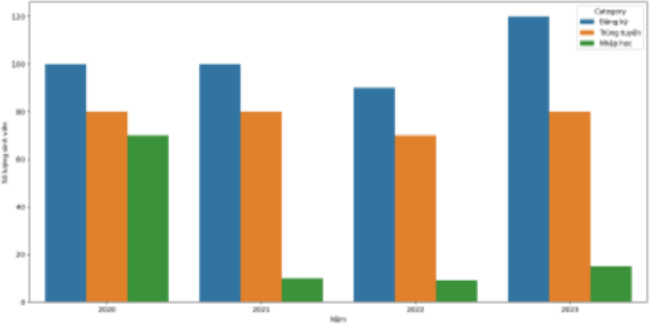
### Xét tuyển kết hợp

\* Số liệu tổng quan theo ngành \*

✔ Chọn theo ngành xét tuyển kết hợp:

Công nghệ Đa phứ... ▾

	Năm	Đăng ký
0	2020	100
1	2021	100
2	2022	90
3	2023	120



\_\_Dự đoán số liệu theo ngành phương pháp xét tuyển kết hợp\_\_

✔ Chọn đối tượng dự báo xét tuyển kết hợp:

Số lượng sinh viên trúng tuyển

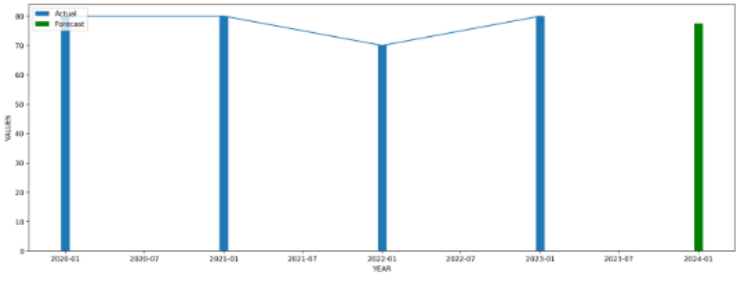
Số lượng sinh viên nhập học

✔ Chọn ngành xét tuyển kết hợp:

Công nghệ Đa phương... ▾

✔ Chọn số năm dự báo xét tuyển kết hợp:

1    -    +

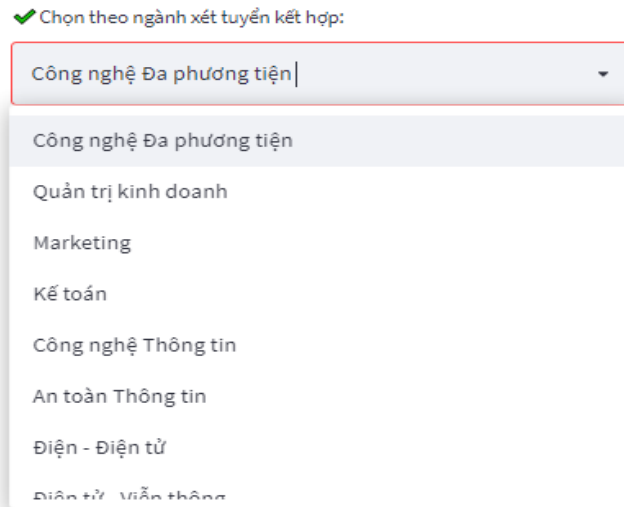


	Dự báo ở năm	Số lượng sinh viên trúng tuyển
0	2024	77

Download kết quả dự báo

**Hình 3.2.32: Giao diện tổng quan xem dữ liệu và dự báo của mục “Xét tuyển kết hợp”**

Do hai mục này giống nhau nên chỉ trình bày “Xét tuyển kết hợp”. Đầu tiên là xem số liệu tổng quan thì người dung phải lựa chọn một ngành trong tất cả các ngành để xem dữ liệu.



**Hình 3.2.33: Giao diện chọn một ngành để xem dữ liệu**

Sau đó dữ liệu của ngành được chọn sẽ hiển thị ra ở bảng bên dưới ngành được chọn và biểu diễn bằng biểu đồ cột phía bên phải.



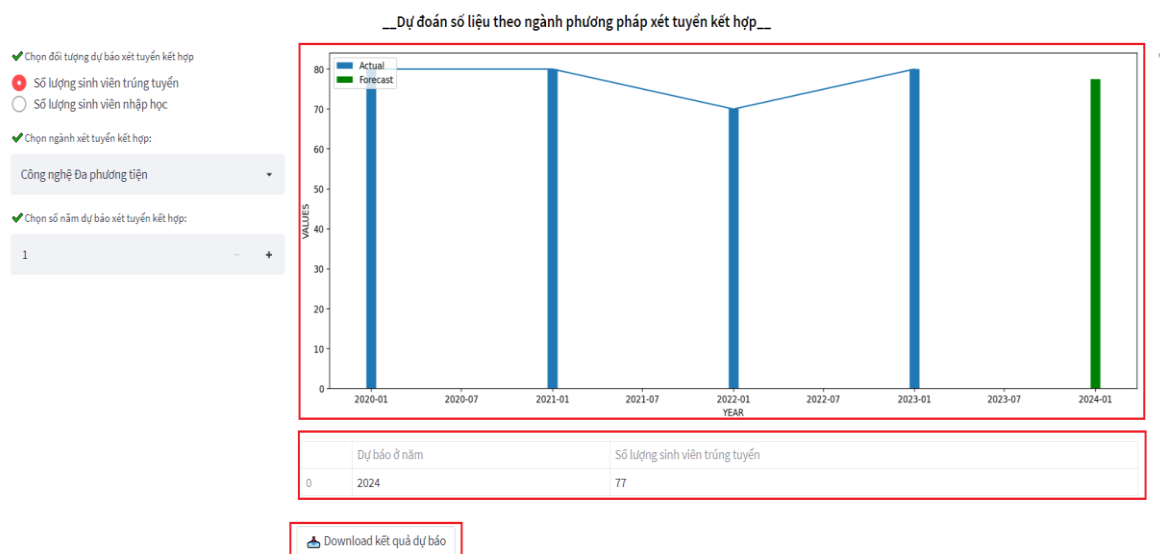
**Hình 3.2.34: Giao diện số liệu tổng quan theo ngành Công nghệ Đa phương tiện qua các năm của các đối tượng (đăng ký, trúng tuyển, nhập học)**

Tiếp theo phần dự báo cũng giống như mục “Xét tuyển dựa vào kết quả THPT” thì dự báo này cũng cho phép người dùng chọn đối tượng (trúng tuyển, nhập học,...),

chọn tiếp số năm dự báo sắp tới theo một ngành nào đó trong tổng các ngành của trường có tuyển sinh như ngành Công nghệ Đa phương tiện, Quản trị kinh doanh, Marketing,...

Sau khi lựa chọn xong kết quả dự báo sẽ được hiện ra bên phải menu gồm có:

- Biểu đồ thể hiện số liệu và đường tăng trưởng theo dữ liệu thực và dữ liệu dự báo.
- Bảng thống kê số liệu dự báo ở các năm tiếp theo và có thể tải về biểu diễn dưới dạng file excel.



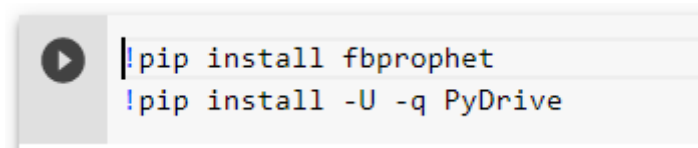
**Hình 3.2.35: Giao diện kết quả dự báo năm tới (2024) của đối tượng trúng tuyển theo ngành Công nghệ Đa phương tiện của mục “Xét tuyển kết hợp”**

## Chương 4: ĐÁNH GIÁ KẾT QUẢ THỬ NGHIỆM

### 4.1. Kết quả thực nghiệm

Phần sau là quá trình kiểm chứng kết quả dự báo khi áp dụng thuật toán Facebook Prophet trên số liệu của ngành Công nghệ đa phương tiện từ năm 2016 đến năm 2021.

Phần phân tích dự báo được xây dựng trên Google Colaboratory. Mục đích của Google Colab đó chính là hỗ trợ chạy code Python trực tiếp thông qua trình duyệt web, một công cụ thích hợp cho việc phân tích dữ liệu, machine learning và giáo dục[13].



```
!pip install fbprophet
!pip install -U -q PyDrive
```

**Hình 4.1.1 : Lệnh cài đặt môi trường**

Hình 4.1.1 là đoạn mã nguồn Python để cài đặt thuật toán Prophet và cài đặt PyDrive để lấy dữ liệu từ googleDrive.

Sau khi cài đặt môi trường và import các thư viện cần thiết. Sau đó tiến hành tạo file có đuôi .CSV để lưu thông tin số lượng sinh viên trúng tuyển của ngành Công nghệ đa phương tiện từ năm 2016 đến năm 2021. File gồm 2 cột ds và y, tương ứng với thời gian và số lượng sinh viên. Việc tiếp theo là upload file lên google drive để đọc data từ Google Colab

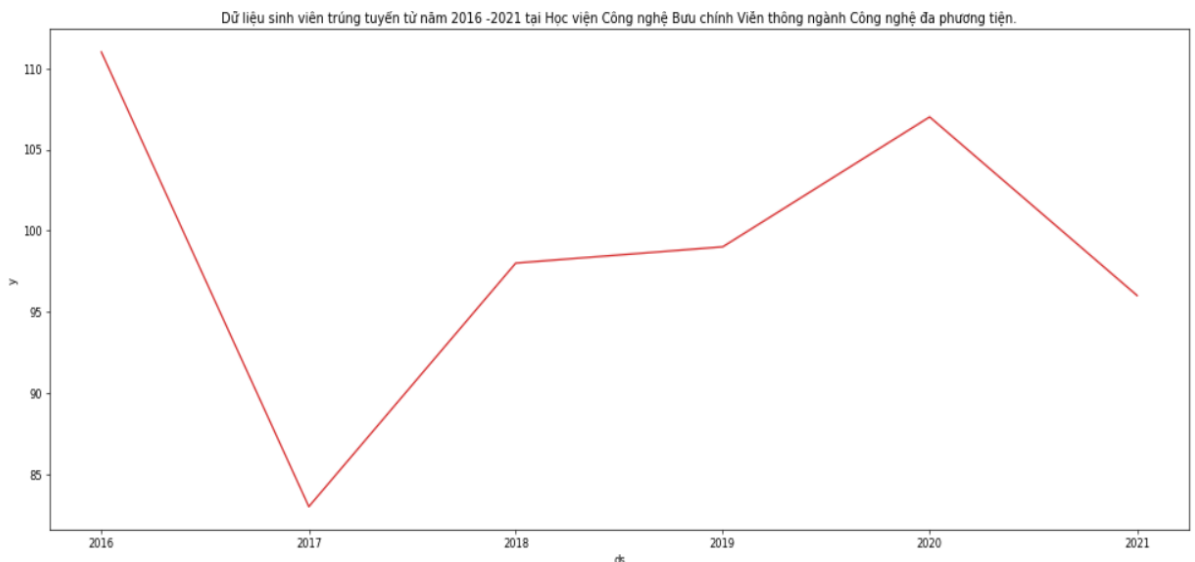


```
[16] # 1. đọc file csv từ Google Colab
auth.authenticate_user()
gauth = GoogleAuth()
gauth.credentials = GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)
#https://drive.google.com/file/d/1J2ekt8pa90QXdB0sP8TNlGbHXNmjJvsG/view?usp=sharing
fileDownloaded = drive.CreateFile({'id':'1J2ekt8pa90QXdB0sP8TNlGbHXNmjJvsG'})
fileDownloaded.GetContentFile('nganhCNDPT.csv')
# 2. Lấy dữ liệu từ file csv đưa vào DataFrame
df = read_csv('nganhCNDPT.csv')
df['ds'] = to_datetime(df['ds'])
df
```

	ds	y
0	2016-01-01	111
1	2017-01-01	83
2	2018-01-01	98
3	2019-01-01	99
4	2020-01-01	107
5	2021-01-01	96

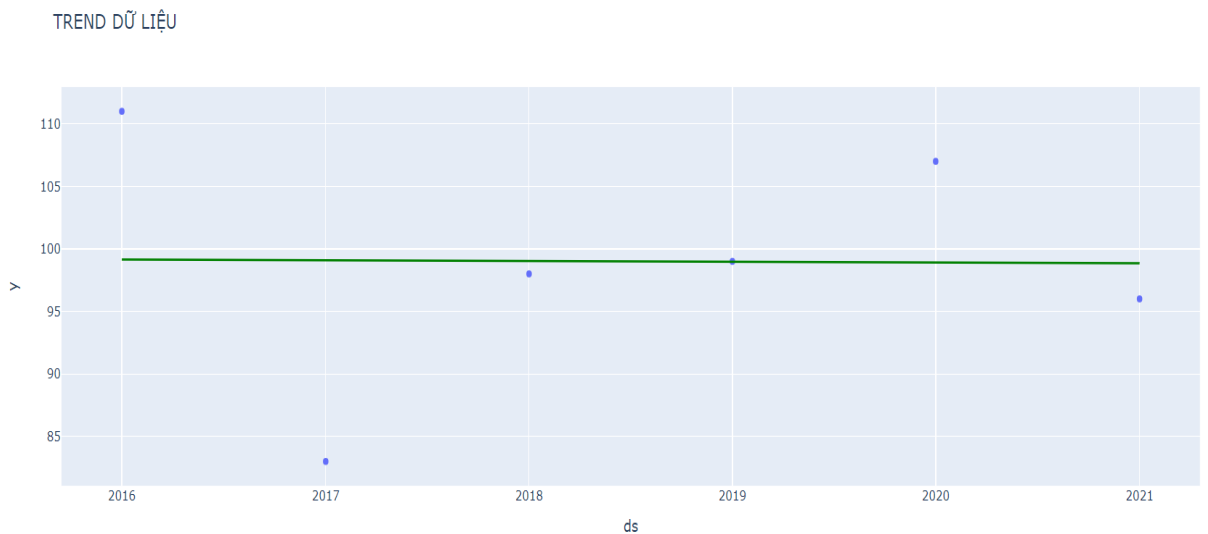
**Hình 4.1.2: Các lệnh đọc dữ liệu từ google drive**

- Sau khi đọc dữ liệu như hình 4.1.2 và hệ thống sẽ thể hiện số liệu qua các năm dưới dạng đường gấp khúc như hình 4.1.3 bên dưới:



**Hình 4.1.3: Dữ liệu ngành Công nghệ đa phương tiện qua các năm**

- Sau đó tiến hành vẽ biểu đồ Trend của dữ liệu bằng Phương pháp Ordinary least square (OLS)[14] như hình 4.1.4 bên dưới:



**Hình 4.1.4: Biểu đồ trend dữ liệu**

**Nhận xét:** Qua biểu đồ Trend, dữ liệu của biểu đồ theo năm (Year) có chiều hướng có giảm nhẹ.

Để đánh giá mức độ hiệu quả của thuật toán cần chia dữ liệu thành 2 phần:

- Phần dữ liệu huấn luyện (tập Train): từ năm 2016 đến năm 2019 (4 năm)
- Phần dữ liệu kiểm thử (tập Test): từ năm 2020 đến năm 2021 (2 năm)

```
# cắt dữ liệu để test
# 2. Lấy dữ liệu từ file csv đưa vào DataFrame
df = read_csv('nganhCNDPT.csv')
split_date = '2019-01-01'
Cus_train = df.loc[df.ds <= split_date].copy()
Cus_test = df.loc[df.ds > split_date].copy()
```

**Hình 4.1.5: Đoạn mã lệnh chia dữ liệu thành 2 tập Cus\_train và Cus\_test**

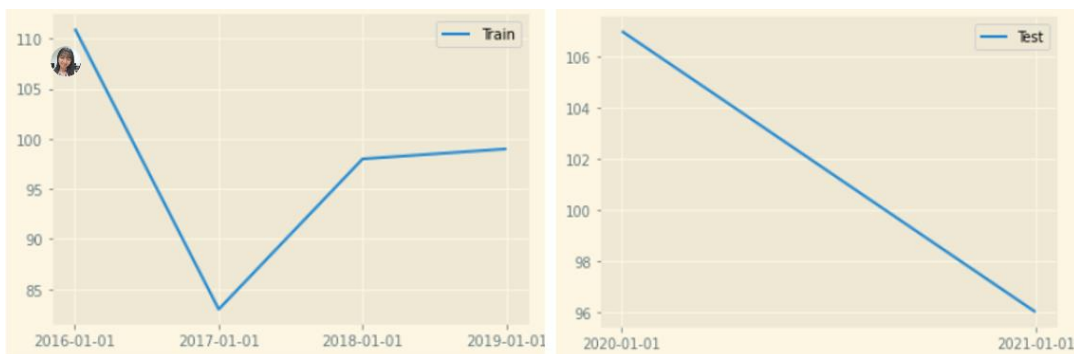
Kết quả dữ liệu sau khi chia

	ds	y
0	2016-01-01	111
1	2017-01-01	83
2	2018-01-01	98
3	2019-01-01	99

	ds	y
4	2020-01-01	107
5	2021-01-01	96

**Hình 4.1.6: Tập Cus\_train và tập Cus\_test sau khi chia dữ liệu**



**Hình 4.1.7: Tập Train và tập Test được vẽ ra biểu đồ**

Sau đó áp dụng thuật toán Prophet với mô hình growth='linear' cho tập Train để dự báo phản hồi cho cả tập train và tập test như sau:

```
# Test dự báo bằng mô hình Prophet(growth='linear')
model_growthlinear = Prophet(growth='linear')
model_growthlinear.fit(Cus_train)
Cus_train_forecast_growth = model_growthlinear.predict(Cus_train)
Cus_test_forecast_growth = model_growthlinear.predict(Cus_test)
```

**Hình 4.1.8: Câu lệnh Thuật toán prophet được áp dụng**

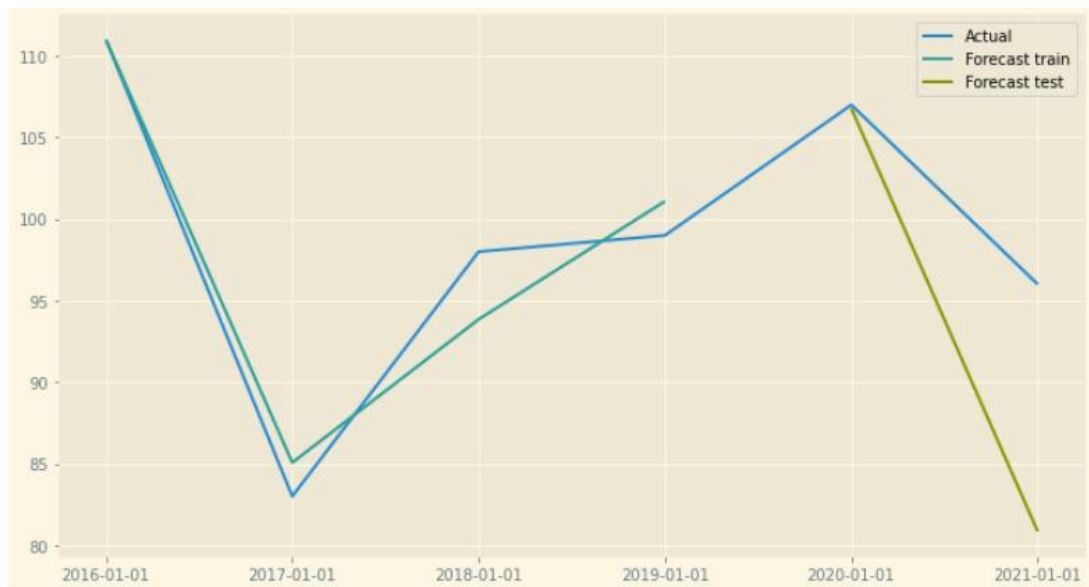
Dùng hàm plot để thể hiện kết quả. Hình sau là đoạn mã truy xuất dữ liệu, thực hiện dự báo bằng thuật toán Prophet với mô hình Growth = Linear và cho hiển thị kết quả dự báo trực tiếp trên trình duyệt web.

```

y_forecast_linear = np.array(Cus_train_forecast_growth['yhat']).reshape(-1)
y_forecast_linear_test = np.array(Cus_test_forecast_growth['yhat']).reshape(-1)
y_cus_test = Cus_test['y'].values
y_cus_train = Cus_train['y'].values
with plt.style.context('Solarize_Light2'):
    plt.figure(figsize=(12, 6.5))
    plt.plot(x,y, label='Actual')
    plt.plot(x2, y_forecast_linear, label='Forecast train')
    plt.plot(x1, y_forecast_linear_test, label='Forecast test')
    plt.legend(loc="upper right")

```

**Hình 4.1.9: Câu lệnh hiển thị kết quả dự báo**



**Hình 4.1.10: Kết quả áp dụng thuật toán Prophet với mô hình growth=linear**

Kết quả thể hiện trên hình gồm 3 biểu đồ:

- Biểu đồ có line là “Actual” màu xanh dương vẽ ra số liệu thực tế từ năm 2016 đến năm 2021
- Biểu đồ có line là “Forecast train” màu xanh lục vẽ ra số liệu khi dùng hàm predict() cho tập Train
- Biểu đồ có line là “Forecast test” màu xanh non vẽ ra số liệu khi dùng hàm predict() cho tập Test

Sai số tuyệt đối trung bình MAE (Mean Absolute Error) của mô hình:

+ Với predict() cho tập Train

```
mean_absolute_error(y_true=Cus_train['y'],
                    y_pred=Cus_train_forecast_growth['yhat'])
= 2.0784390587289465
```

**Hình 4.1.11: Công thức và kết quả MAE tập Train**

+ Với predict() cho tập Test

```
mean_absolute_error(y_true=Cus_test['y'],
                    y_pred=Cus_test_forecast_growth['yhat'])
= 7.684307588200809
```

**Hình 4.1.12: Công thức và kết quả MAE tập Test**

Tỉ lệ phần trăm trung bình sai số tuyệt đối được tính bằng công thức[15]:

$$\text{MAPE} = \frac{\sum_{t=1}^n \left( \frac{|y_t - y_d|}{|y_t|} \right)}{n} \quad (4.1)$$

Trong đó:

yt: giá trị kiểm chứng (test value)

yd: giá trị dự báo (yhat)

n: tổng thời gian dự báo

t: thời gian dự báo

Tỉ lệ phần trăm trung bình sai số tuyệt đối MAPE:

+ Với predict() cho tập Train

```
def mean_absolute_percentage_error(y_true, y_pred):
    """Calculates MAPE given y_true and y_pred"""
    y_true, y_pred = np.array(y_true), np.array(y_pred)
    return np.mean(np.abs((y_true - y_pred) / y_true)) * 100
mean_absolute_percentage_error(y_true=Cus_train['y'],
                               y_pred=Cus_train_forecast_growth['yhat'])
= 2.2104851042067954
```

**Hình 4.1.13: Công thức và kết quả MAPE tập Train**

+ Với predict() cho tập Test

```
def mean_absolute_percentage_error(y_true, y_pred):
    """Calculates MAPE given y_true and y_pred"""
    y_true, y_pred = np.array(y_true), np.array(y_pred)
    return np.mean(np.abs((y_true - y_pred) / y_true)) * 100
# tỉ lệ phần trăm sai số tuyệt đối mô hình prophet growth
mean_absolute_percentage_error(y_true=Cus_test['y'],
                               y_pred=Cus_test_forecast_growth['yhat'])
= 7.992689576548176
```

**Hình 4.1.14: Công thức và kết quả MAPE tập Test**

In ra kết quả giá trị dự báo cho năm 2020 và 2021

```
Cus_test_forecast_growth.head()[['ds', 'yhat']]
```

	ds	yhat
0	2020-01-01	106.779666
1	2021-01-01	80.851719

**Hình 4.1.15: Kết quả dự báo năm 2020 và 2021**

Prophet cho phép điều chỉnh tham số để kết quả dự báo chính xác hơn bằng cách sử dụng Cross-validation để chọn tham số phù hợp cho bộ dữ liệu trên bằng cách thay đổi giá trị cho tham số `changepoint_prior_scale`[16].

Trước tiên, cần chọn 1 bộ tham số cho `changepoint_prior_scale`

```
param_grid = {
    'changepoint_prior_scale': [0.00001, 0.0001, 0.001,
                                0.01, 0.1, 0.2, 0.3]}
```

**Hình 4.1.16: Bộ tham số của `changepoint_prior_scale`**

Ứng với mỗi tham số trong bộ tham số trên, tiến hành đi tính giá trị RMSE để chọn ra một tham số tối ưu nhất là tham số có giá trị RMSE nhỏ nhất

```
# Find the best parameters
tuning_results = pd.DataFrame(all_params)
tuning_results['rmse'] = rmses
print(tuning_results)
best_params = all_params[np.argmin(rmses)]
print('Tham số được chọn:')
print(best_params)
```

	changepoint_prior_scale	rmse
0	0.00001	33.890352
1	0.00010	8.980455
2	0.00100	9.960649
3	0.01000	9.886925
4	0.10000	11.190394
5	0.20000	11.099764
6	0.30000	11.208809

Tham số được chọn:  
{'changepoint\_prior\_scale': 0.0001}

**Hình 4.1.17: Giá trị RMSE tương ứng với mỗi tham số**

Giá trị tối ưu nhất của RMSE là 8.980455 tương ứng với tham số `changepoint_prior_scale` có giá trị là 0.0001. Sau đó áp dụng lại thuật toán Prophet với `changepoint_prior_scale` có giá trị là 0.0001 để đánh giá và so sánh với kết quả khi chưa áp dụng tham số

```
# Test dự báo bằng mô hình growth='logistic', changepoint_prior_scale=0.0001 trong Prophet
model_linear = Prophet(growth='linear', changepoint_prior_scale=0.0001)
model_linear.fit(Cus_train)
Cus_test_forecast_linear = model_linear.predict(Cus_test)
y_forecast_linear = np.array(Cus_test_forecast_linear['yhat']).reshape(-1)
y_cus_test = Cus_test['y'].values
plt.figure(figsize=(12, 6.5))
plt.plot(y_forecast_linear, label='Dự đoán')
plt.plot(y_cus_test, label='Thực tế')
plt.legend(loc="upper right")
```

**Hình 4.1.18: Thuật toán Prophet với tham số changepoint\_prior\_scale**

Kết quả giá trị dự báo cho năm 2020 và 2021 khi có áp dụng tham số changepoint\_prior\_scale=0.0001 như sau:

```
Cus_test_forecast_linear.head()[['ds', 'yhat']]
```

	ds	yhat
0	2020-01-01	107.826864
1	2021-01-01	82.222795

**Hình 4.1.19: Kết quả dự báo năm 2020 và 2021 với changepoint\_prior\_scale**

Sai số tuyệt đối trung bình MAE = 7.302034465032428

Tỉ lệ phần trăm trung bình sai số tuyệt đối MAPE = 7.56201263762127

**Bảng 4.1: So sánh kết quả**

Giá trị	Prophet(growth='linear')	Prophet(growth='linear', changepoint_prior_scale =0.0001)	Giá trị thực tế
yhat(2020)	106.779666	107.826864	107
yhat(2021)	80.851719	82.222795	96
MAE	7.302034465032428	7.684307588200809	-
MAPE	7.562012637162127	7.992689576548176	-



So sánh kết quả của 2 độ đo Sai số tuyệt đối trung bình MAE của 2 trường hợp trên. Kết quả cho thấy đối với bộ dữ liệu này, sai số tuyệt đối của mô hình khi có điều chỉnh tham số changepoint\_prior\_scale sẽ thấp hơn so với sai số tuyệt đối của mô hình khi không áp dụng tham số. Và tỉ lệ sai số tuyệt đối này là nhỏ.

#### **4.2. Nhận xét kết quả thử nghiệm ứng dụng**

Qua kết quả dự báo cho thấy ưu điểm của thuật toán prophet rất linh hoạt, có thể điều chỉnh các tham số cho phù hợp với bộ dữ liệu đầu vào để có được kết quả tối ưu nhất.

Do dữ liệu được lựa chọn để đánh giá là dữ liệu ngành Công nghệ đa phương tiện, nhìn chung bộ dữ liệu này tương đối ổn định, ít biến động, nên cho ra kết quả dự báo khá sát với thực tế. Nhưng đối với bộ dữ liệu khác, có thể có kết quả và sai số lớn hơn, do dữ liệu ít và có sự thay đổi lớn qua các năm.

## KẾT LUẬN

Hiện nay có rất nhiều hướng nghiên cứu để triển khai hệ thống quản lý và dự báo tuyển sinh. Luận văn trình bày một hướng tiếp cận nhằm đáp ứng nhu cầu quản lý và dự báo số liệu cho công tác tuyển sinh, đó là nghiên cứu xây dựng hệ thống dự báo số liệu dựa trên phương pháp dự báo chuỗi thời gian dựa trên thuật toán Facebook Prophet. Kết quả thực nghiệm của mô hình dự báo trên dữ liệu phụ thuộc vào dữ liệu đầu vào. Do đặc thù của ngành tuyển sinh thay đổi từng năm, nên để lấy dữ liệu đồng bộ thì chỉ thu thập từ năm 2016 đến nay, dẫn đến dữ liệu ít, và dữ liệu phụ thuộc nhiều vào quy chế, quy định hằng năm của Bộ giáo dục và đào tạo.

Đề tài mở ra hướng phát triển về dự báo trong lĩnh vực giáo dục và đào tạo, góp phần hỗ trợ cán bộ quản lý có cái nhìn tổng quan giúp cho việc lên kế hoạch, chiến lược cụ thể nhằm chủ động trong công tác tuyển sinh, giúp công tác tuyển sinh phát triển ngày một tốt hơn.

## DANH MỤC TÀI LIỆU THAM KHẢO

### Tiếng Việt

[1] HT Thừa, T Hạnh, LNK Thư, HT Trụ, LB Loan, HL Minh, LTK Vân, “Phân tích nhu cầu và dự báo hoạt động đầu tư tại Khu công nghệ cao Thành phố Hồ Chí Minh,” 2021.

[3] <https://tuyensinh.ptit.edu.vn>. Truy cập ngày 22/04/2021

### Tiếng Anh

[2] Laabidi, Nada Almalki and Maysoun Al-zahrani, năm 2020, “Time series facebook prophet model and python forcovid-19 outbreak prediction”

[4] Sean J. Taylor and Benjamin Letham, "Forecasting at Scale," The American Statistician, vol. 72, no. 1, pp. 37-45, 2018.

[5] Zar Zar Oo and Sabai Phyu, "Time Series Prediction Based on Facebook Prophet: A Case Study, Temperature Forecasting in Myintkyina," International Journal of Applied Mathematics, Electronics and Computers, vol. 8, no. 4, pp. 263-267, 2020.

[6] Emir Žunić, Kemal Korjenić, Kerim Hodžić and Dženana Đonko, "Application of Facebook's Prophet Algorithm for Successful Sales Forecasting Based on Real-World Data," International Journal of Computer Science & Information Technology, vol. 12, no. 2, pp. 23-36, 2020.

[7] <https://medium.com/analytics-vidhya/how-does-prophet-work>. Truy cập ngày 22/04/2022.

[8] Surya Srikar Sirigineedi, “A quick start of time series forecasting : using facebook prophet”, pp. 37-45, 2019.

[9] <https://www.jetbrains.com/help/idea/discover-intellij-idea.html>. Truy cập ngày 20/05/2021

[10] <https://www.javatpoint.com/react-introduction/>. Truy cập ngày 20/04/2022

[11] <https://dev.mysql.com/doc/refman/8.0/en/introduction.html>. Truy cập ngày 20/04/2022

[12] <https://streamlit.io/>. Truy cập ngày 25/05/2021.

- [13] <https://www.mygreatlearning.com/blog/google-colab-tutorial/>. Truy cập ngày 20/04/2021
- [14] <https://www.datarobot.com/blog/ordinary-least-squares-in-python/>. Truy cập ngày 20/04/2022
- [15] <https://www.statisticshowto.com/mean-absolute-percentage-error-mape/>. Truy cập ngày 20/04/2022
- [16] <https://facebook.github.io/prophet/docs/diagnostics.html>. Truy cập ngày 20/04/2022

## **BẢNG CAM ĐOAN**

Tôi cam đoan đã thực hiện việc kiểm tra mức độ tương đồng nội dung luận văn “**Hệ thống quản lý và dự báo tuyển sinh tại học viện công nghệ bưu chính viễn thông**” qua phần mềm Kiểm Tra Tài Liệu một cách trung thực và đạt kết quả mức độ tương đồng 5% toàn bộ nội dung luận văn. Bản luận văn kiểm tra qua phần mềm là bản cứng luận văn đã nộp để bảo vệ trước hội đồng. Nếu sai tôi xin chịu các hình thức kỷ luật theo quy định hiện hành của Học viện.

*TP.HCM, ngày 16 tháng 07 năm 2022*

**HỌC VIÊN CAO HỌC**

**BÙI THỊ BÉ BA**



## BÁO CÁO KIỂM TRA TRÙNG LẶP

### Thông tin tài liệu

Tên tài liệu: BeBa\_LuanVan\_saubaove\_doit  
Tác giả: BuiThiBeBa  
Điểm trùng lặp: 5  
Thời gian tải lên: 13:13 16/07/2022  
Thời gian sinh báo cáo: 13:15 16/07/2022  
Các trang kiểm tra: 68/68 trang



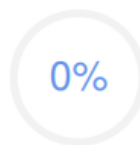
### Kết quả kiểm tra trùng lặp



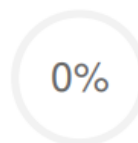
Có 5% nội dung trùng lặp



Có 95% nội dung không trùng lặp



Có 0% nội dung người dùng loại trừ



Có 0% nội dung hệ thống bỏ qua

### Nguồn trùng lặp tiêu biểu

123docz.net tailieu.vn tapchiktvtv.vn

**HỌC VIÊN**

**NGƯỜI HƯỚNG DẪN KHOA HỌC**

**BÙI THỊ BÉ BA**

**TS. HUỖNH TRỌNG THỨA**

**BIÊN BẢN  
HỌP HỘI ĐỒNG CHẤM LUẬN VĂN THẠC SĨ**

Căn cứ quyết định số Quyết định số 712-54/QĐ-HV ngày 10 tháng 06 năm 2022 của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm luận văn thạc sĩ đợt tháng 07 năm 2022. Hội đồng đã họp vào hồi 11h00-11h45, ngày 02 tháng 07 năm 2022 tại Học viện Công nghệ Bưu chính Viễn thông để chấm luận văn thạc sĩ cho:

Học viên: **Bùi Thị Bé Ba**

Tên luận văn: **Hệ thống quản lý và dự báo tuyển sinh tại Học viện Công nghệ Bưu chính Viễn thông**

Chuyên ngành: **Hệ thống Thông tin**

Mã số: 8.48.01.04

Các thành viên của Hội đồng chấm luận văn có mặt: 05 / 05

TT	HỌ VÀ TÊN	TRÁCH NHIỆM TRONG HĐ	GHI CHÚ
1	PGS.TS. Đinh Đức Anh Vũ	Chủ tịch	
2	PGS.TS. Trần Trung Duy	Thư ký	
3	PGS.TS. Nguyễn Đình Thuận	Phản biện 1	
4	PGS.TS. Lê Hoàng Thái	Phản biện 2	
5	PGS.TS. Trần Vĩnh Phước	Ủy viên	

**Các nội dung thực hiện:**

1. Chủ tịch Hội đồng điều khiển buổi họp. Công bố quyết định của Giám đốc Học Viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm luận văn thạc sĩ
2. Người hướng dẫn khoa học hoặc thư ký đọc lý lịch khoa học và các điều kiện bảo vệ luận văn của học viên. (có bản lý lịch khoa học và kết quả các môn học cao học của học viên kèm theo).
3. Học viên trình bày tóm tắt luận văn.
4. Phản biện 1 đọc nhận xét (có văn bản kèm theo)
5. Phản biện 2 đọc nhận xét (có văn bản kèm theo)
6. Các câu hỏi của thành viên Hội đồng:

Câu hỏi 1: N.hững thuật toán nào phổ biến? tại sao HV chọn thuật toán trong luận văn? Tên phi tuyến và mức độ của thuật toán?  
Câu hỏi 2: Thuật toán có đáp ứng yêu cầu dự báo? từ số liệu năm 2016 - 2020?  
Câu hỏi 3: làm rõ mô hình, dữ liệu huấn luyện?  
Câu hỏi 4: làm rõ tính chọn lọc? Ở VN hiểu dự báo số lượng nộp vào trường?

**7. Trả lời của học viên:**

Trả lời 1: lý do học viên chọn thuật toán trong LV của HV. Thuật toán đã triển khai, đạt hiệu quả cao. Thuật toán có

..... nhiệm vụ, dùng lại trên lại, hoàn phần xử lý tạo ra từ nhiên  
..... Trả lời: Thuật toán có tính mùa vụ, thích hợp và dễ hiểu, hiệu quả

8. Thư ký đọc nhận xét về quá trình thực hiện luận văn của học viên (có văn bản kèm theo).

9. Hội đồng họp riêng:

- Bầu Ban kiểm phiếu:

1. Trưởng Ban kiểm phiếu: PGS.TS. Lê Hoàng Thái
2. Ủy viên Ban kiểm phiếu: PGS.TS. Nguyễn Đình Thuận
3. Ủy viên Ban kiểm phiếu: PGS.TS. Trần Trung Duy

- Hội đồng chấm luận văn bằng bỏ phiếu kín.

- Ban kiểm phiếu làm việc:

- Trưởng Ban kiểm phiếu báo cáo kết quả kiểm phiếu (có Biên bản họp Ban kiểm phiếu kèm theo)

- Điểm trung bình của luận văn: 8.4

**Kết luận:**

1. Các nội dung cần chỉnh sửa, hoàn thiện sau bảo vệ luận văn:

Học viên cần liên chỉnh định dạng, công thức toán học  
theo định dạng quy định.  
Học viên cần làm rõ về sơ sử dụng thuật toán "logit".

2. Đề nghị Học viện công nhận (hoặc không) và cấp bằng (hoặc không) thạc sĩ cho học viên: **Bùi Thị Bé Ba**

3. Luận văn có thể phát triển thành đề tài nghiên cứu cho NCS... không  
Buổi làm việc kết thúc vào... 8h.45... cùng ngày.

**Chủ tịch**

PGS.TS. Đinh Đức Anh Vũ

**Thư ký**

PGS.TS. Trần Trung Duy

**Phản biện 1**

PGS.TS. Nguyễn Đình Thuận

**Phản biện 2**

PGS.TS. Lê Hoàng Thái

**Ủy viên**

PGS.TS. Trần Vĩnh Phước



**BẢN NHẬN XÉT LUẬN VĂN TỐT NGHIỆP THẠC SĨ**  
(Dùng cho người phản biện)

Tên đề tài luận văn:

**HỆ THỐNG QUẢN LÝ VÀ DỰ BÁO TUYỂN SINH TẠI HỌC VIỆN CÔNG  
NGHỆ BUU CHÍNH VIỄN THÔNG**

Chuyên ngành: Hệ thống thông tin

Mã số: 8.48.01.04

Tên học viên: **Bùi Thị Bé Ba**

Họ và tên người nhận xét: **Lê Hoàng Thái**

Học hàm, học vị: Phó giáo sư, Tiến sĩ

Chuyên ngành: Khoa học máy tính

Cơ quan công tác: trường Đại học Khoa học tự nhiên, ĐHQG- HCM.

**NỘI DUNG NHẬN XÉT**

**I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:**

Xuất phát từ mục tiêu hỗ trợ cán bộ quản lý tại Học viện Công nghệ Bưu chính Viễn thông trong các công tác tuyển sinh. Cụ thể, làm sao nắm bắt tình hình tuyển sinh (cơ sở vật chất; đội ngũ giảng viên, khả năng tuyển sinh); từ đó, có cái nhìn tổng quan để kịp thời đưa ra giải pháp chiến lược về tư vấn tuyển sinh.

Để thực hiện được mục tiêu trên, Luận văn hướng tới việc xây dựng hệ thống để hỗ trợ cho công tác quản lý tuyển sinh với giao diện thân thiện, cho phép xem xét nhanh các số liệu, tránh rủi ro sai sót cho người quản lý.

Nhận thấy: công việc thực hiện trong luận văn là cần thiết và có giá trị thực tiễn cao.

**II/ Về nội dung, chất lượng của luận văn, các kết quả đã đạt được (so với đề cương đã được duyệt):**

Xây dựng được một hệ thống hỗ trợ cho công tác quản lý tuyển sinh với các chức năng:

1. Hỗ trợ phục vụ quản lý.
2. Truy tìm dữ liệu.

3. Dự báo: Dự báo số lượng sinh viên trúng tuyển và nhập học ứng với từng ngành, vùng, khu vực; trực quan hóa các dữ liệu dự báo thông qua đồ thị.

Luận văn đã hoàn thành các kết quả đạt được so với đề cương đã duyệt ban đầu.

**III/ Những vấn đề cần giải thích thêm:**

Luận văn thiên về ứng dụng. Làm rõ hơn tính khoa học của luận văn? Cụ thể, Làm rõ hơn: hoạt động của mô hình dự báo chuỗi thời gian Prophet cho dự báo dữ liệu tuyển sinh?

**IV/ Kết luận:**

Tôi đồng ý đề học viên Bùi Thị Bé Ba được bảo vệ luận văn trước Hội đồng chấm luận văn thạc sĩ

Ngày 25 tháng 06 năm 2022  
NGƯỜI NHẬN XÉT



**PGS.TS.Lê Hoàng Thái**

**BẢN NHẬN XÉT LUẬN VĂN TỐT NGHIỆP THẠC SĨ**  
(Dùng cho người phản biện)

Tên đề tài luận văn: **Hệ thống quản lý và dự báo tuyển sinh tại Học viện Công nghệ  
Bưu chính Viễn thông**

Chuyên ngành: Hệ thống thông tin

Mã số: 60 48 01 04

Tên học viên: **Bùi Thị Bé Ba**

Họ và tên người nhận xét: **Nguyễn Đình Thuận**

Học hàm, học vị: PGS. TS

Chuyên ngành: Công nghệ thông tin

Cơ quan công tác: **Trường Đại học Công nghệ thông tin- ĐHQGTP.HCM**

**NỘI DUNG NHẬN XÉT**

**I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:**

Luận văn nghiên cứu và ứng dụng thuật toán Prophet, xây dựng hệ thống quản lý và dự báo các chỉ tiêu về tuyển sinh như số hồ sơ đăng ký, số lượng trúng tuyển, số lượng sinh viên nhập học tại Học viện Công nghệ Bưu chính Viễn thông.

**II/ Về nội dung, chất lượng của luận văn, các kết quả đã đạt được (so với đề cương đã được duyệt):**

Luận văn đã cài đặt ứng dụng gồm:

- Chức năng nhập hồ sơ thí sinh.
- Tìm kiếm một số loại thông tin về thí sinh.
- Chức năng dự báo số hồ sơ đăng ký, số lượng trúng tuyển, số lượng sinh viên nhập học.
- Xuất ra một số đồ thị hiển thị dữ liệu trên.

### **III/ Những vấn đề cần giải thích thêm:**

- Cần làm rõ thêm với tập dữ liệu từ năm 2016 đến 2021, có đủ đáp ứng tính mùa vụ để áp dụng thuật toán Prophet hay không.

- Hiện nay có khá nhiều thuật toán dự báo, vì sao học viện chỉ sử dụng thuật toán Prophet.

### **IV/ Kết luận:**

Luận văn đáp ứng các yêu cầu của luận văn thạc sĩ chuyên ngành Hệ thống thông tin.

Tôi đồng ý đề học viên Bùi Thị Bé Ba được bảo vệ luận văn trước Hội đồng chấm luận văn thạc sĩ.

Ngày 25 tháng 06 năm 2022  
NGƯỜI NHẬN XÉT



**PGS. TS. Nguyễn Đình Thuận**

**BÁO CÁO GIẢI TRÌNH**  
**SỬA CHỮA, HOÀN THIỆN LUẬN VĂN THẠC SĨ**

Họ và tên học viên: **Bùi Thị Bé Ba**

Chuyên ngành: **Hệ thống thông tin**

Khóa: **2020-2022**

Tên đề tài: **Hệ thống quản lý và dự báo tuyển sinh tại Học viện Công nghệ Bru chính Viễn thông**

Người hướng dẫn khoa học: **TS. Huỳnh Trọng Thura**

Ngày bảo vệ: **02/07/2022**

Các nội dung học viên đã sửa chữa, bổ sung trong luận văn theo ý kiến đóng góp của Hội đồng chấm luận văn:

TT	Ý kiến hội đồng	Sửa chữa của học viên
1	Học viên cần chỉnh sửa định dạng của luận văn: công thức cần đánh số, viết bằng word equation hoặc Mathtype.	Học viên xin tiếp thu ý kiến của hội đồng và đã chỉnh sửa nội dung yêu cầu ở chương 1, mục 1.3, trang 9; chương 2, mục 2.4, trang 18, 19 của luận văn.
2	Cần làm rõ vì sao sử dụng thuật toán Prophet	Học viên xin tiếp thu ý kiến của hội đồng và đã chỉnh sửa bổ sung nội dung yêu cầu ở chương 2, mục 2.4, trang 16 của luận văn.

*Tp.HCM, ngày 15 tháng 07 năm 2022*

**Ký xác nhận của**

**CHỦ TỊCH HỘI ĐỒNG  
CHẤM LUẬN VĂN**

**THƯ KÝ HỘI ĐỒNG**

**NGƯỜI HƯỚNG DẪN  
KHOA HỌC**

**HỌC VIÊN**

**PGS.TS. Đinh Đức Anh Vũ**

**PGS.TS. Trần Trung Duy**

**TS. Huỳnh Trọng Thura**

**Bùi Thị Bé Ba**