2



**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP.HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

□□□□



**BÁO CÁO MÔN THỰC HÀNH KHAI THÁC DỮ LIỆU**

**ĐỀ TÀI:**

**PHÂN TÍCH HÀNH VI VÀ THÓI QUEN MUA SẮM CỦA NGƯỜI TIÊU DÙNG**

Giảng viên hướng dẫn: Ngô Dương Hà

Sinh viên thực hiện:

1. 2001215850 – Huỳnh Vĩ Khang

2. 2001215964 – Đào Quí Mùi

TP. HỒ CHÍ MINH – 2024

**MỤC LỤC**

[LỜI MỞ ĐẦU 1](#_Toc184240799)

[CHƯƠNG 1: GIỚI THIỆU 2](#_Toc184240800)

[1.1 BỐI CẢNH DỰ ÁN: 2](#_Toc184240801)

[1.2 MỤC TIÊU DỰ ÁN: 2](#_Toc184240802)

[1.3 PHẠM VI DỰ ÁN: 2](#_Toc184240803)

[1.4 KẾT CHƯƠNG: 3](#_Toc184240804)

[2.1 Phân tích yêu cầu 4](#_Toc184240805)

[2.2 Yêu cầu chức năng 4](#_Toc184240806)

[CHƯƠNG 3: THIẾT KẾ 4](#_Toc184240807)

[3.1 Các thuật toán sử dụng 4](#_Toc184240808)

[CHƯƠNG 4: THỰC HIỆN 4](#_Toc184240809)

[4.1 Các ứng dụng trong bài toán 4](#_Toc184240810)

[CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN 5](#_Toc184240811)

[5.1 HẠN CHẾ CỦA HỆ THỐNG: 5](#_Toc184240812)

[5.2 KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG: 5](#_Toc184240813)

[CHƯƠNG 6: KẾT LUẬN 6](#_Toc184240814)

[TÀI LIỆU THAM KHẢO 7](#_Toc184240815)

**DANH MỤC CÁC KÝ HIỆU VÀ CHỮ VIẾT TẮT**

| **Viết tắt** | **Tiếng Anh** | **Tiếng Việt** |
| --- | --- | --- |
| HUIT | Ho Chi Minh City University of Industry | Trường Đại Học Công Thương TP.HCM |

# LỜI MỞ ĐẦU

Dự báo và phân tích sức khỏe giấc ngủ đóng vai trò quan trọng trong việc nâng cao chất lượng cuộc sống và sức khỏe cộng đồng. Giấc ngủ là một yếu tố thiết yếu ảnh hưởng đến sức khỏe thể chất và tinh thần của con người. Sự phát triển của công nghệ dữ liệu lớn (Big Data) và trí tuệ nhân tạo đã mở ra nhiều cơ hội mới trong việc phân tích và cải thiện chất lượng giấc ngủ với độ chính xác cao hơn.

Chúng em lựa chọn thực hiện dự án “Phân tích sức khỏe giấc ngủ và đời sống” trong khuôn khổ khóa học Big Data tại trường Đại Học Công Thương (HUIT). Dự án nhằm tìm hiểu và ứng dụng các công nghệ xử lý dữ liệu lớn để phân tích các yếu tố ảnh hưởng đến chất lượng giấc ngủ, từ đó đưa ra các giải pháp cải thiện giấc ngủ cho mọi người.

Với khả năng thu thập và phân tích dữ liệu từ các thiết bị đeo thông minh (wearable devices) và ứng dụng di động, dự án này không chỉ giúp cải thiện sức khỏe giấc ngủ mà còn mở ra cơ hội ứng dụng trong các lĩnh vực như y tế, tâm lý học, và đời sống hàng ngày, góp phần nâng cao chất lượng cuộc sống cho cộng đồng.

# CHƯƠNG 1: GIỚI THIỆU

## BỐI CẢNH DỰ ÁN:

Trong kỷ nguyên số, dữ liệu đóng vai trò quan trọng trong việc đưa ra các quyết định chính xác và hiệu quả. Phân tích sức khỏe giấc ngủ là một trong những lĩnh vực được hưởng lợi từ việc ứng dụng công nghệ Big Data. Nhu cầu cải thiện chất lượng giấc ngủ ngày càng cao, đặc biệt trong bối cảnh căng thẳng và áp lực trong cuộc sống hiện đại. Việc hiểu rõ các yếu tố ảnh hưởng đến giấc ngủ sẽ giúp người dùng có những giải pháp phù hợp để nâng cao sức khỏe thể chất và tinh thần.

## MỤC TIÊU DỰ ÁN:

Dự án BigData-Phân-tích-sức-khỏe-giấc-ngủ được thực hiện nhằm xây dựng một hệ thống phân tích sức khỏe giấc ngủ dựa trên dữ liệu lớn, sử dụng các công nghệ phân tán tiên tiến. Mục tiêu cụ thể của dự án bao gồm:

1. **Xây dựng kiến trúc hệ thống:** Thiết kế và triển khai một hệ thống phân tích sức khỏe giấc ngủ hiệu quả, có khả năng xử lý lượng dữ liệu lớn, đảm bảo tính ổn định và khả năng mở rộng.
2. **Thu thập và xử lý dữ liệu:** Thu thập dữ liệu về giấc ngủ từ nhiều nguồn khác nhau, bao gồm các cuộc khảo sát về sức khỏe giấc ngủ ở các bệnh viện lớn, ứng dụng di động, và các nghiên cứu khoa học. Xử lý và tiền xử lý dữ liệu để đảm bảo chất lượng và tính nhất quán cho quá trình phân tích.
3. **Đào tạo và đánh giá mô hình:** Xây dựng và huấn luyện các mô hình học máy tiên tiến như Random Forest, Logistic Regression, và các mô hình mạng nơ-ron để phân tích các yếu tố ảnh hưởng đến chất lượng giấc ngủ và đưa ra các khuyến nghị cải thiện. Sử dụng các phương pháp đánh giá phù hợp để xác định độ chính xác của các mô hình và lựa chọn mô hình tối ưu.
4. **Phát triển giao diện web:** Xây dựng một giao diện web thân thiện với người dùng, cung cấp thông tin phân tích sức khỏe giấc ngủ một cách trực quan, rõ ràng và dễ hiểu. Giao diện web sẽ hiển thị các thông tin liên quan đến chất lượng giấc ngủ, xu hướng giấc ngủ, và các khuyến nghị cải thiện.
5. **Thực hiện và đánh giá hệ thống:** Triển khai hệ thống phân tích sức khỏe giấc ngủ trong môi trường thực tế, thu thập dữ liệu và đánh giá hiệu quả của hệ thống dựa trên các chỉ số phù hợp.

## PHẠM VI DỰ ÁN:

Dự án tập trung vào việc xây dựng hệ thống phân tích sức khỏe giấc ngủ cho một nhóm người dùng cụ thể, sử dụng dữ liệu lịch sử và các thông tin sức khỏe hiện tại. Phạm vi dự án bao gồm:

1. **Thiết kế và triển khai kiến trúc hệ thống:** Bao gồm việc lựa chọn các công nghệ phù hợp, cấu hình và tích hợp các thành phần của hệ thống để đảm bảo khả năng mở rộng và xử lý dữ liệu hiệu quả.
2. **Xây dựng và huấn luyện các mô hình học máy:** Bao gồm việc lựa chọn các thuật toán học máy phù hợp, tối ưu hóa siêu tham số, và đánh giá hiệu quả của mô hình để phân tích các yếu tố ảnh hưởng đến chất lượng giấc ngủ.
3. **Phát triển giao diện web:** Bao gồm thiết kế giao diện, lập trình chức năng, và tích hợp với hệ thống phân tích để cung cấp thông tin sức khỏe giấc ngủ một cách trực quan và dễ hiểu cho người dùng.
4. **Thực hiện các bài kiểm tra và đánh giá hệ thống:** Bao gồm đánh giá độ chính xác của các mô hình phân tích, đánh giá hiệu suất của hệ thống, và đánh giá sự hài lòng của người dùng thông qua các khảo sát và phản hồi.

## KẾT CHƯƠNG:

Dự án BigData-Phân-tích-sức-khỏe-giấc-ngủ nhằm mục tiêu cung cấp một hệ thống phân tích sức khỏe giấc ngủ hiện đại, chính xác và hiệu quả, sử dụng các công nghệ Big Data tiên tiến. Hệ thống này hứa hẹn mang lại lợi ích thiết thực cho cộng đồng, hỗ trợ người dùng cải thiện chất lượng giấc ngủ và sức khỏe tổng thể, đồng thời thúc đẩy nhận thức về tầm quan trọng của giấc ngủ trong đời sống hàng ngày.

CHƯƠNG 2: PHÂN TÍCH ĐỀ TÀI

## Phân tích yêu cầu

## Yêu cầu chức năng

# CHƯƠNG 3: THIẾT KẾ

## Các thuật toán sử dụng

# CHƯƠNG 4: THỰC HIỆN

## Các ứng dụng trong bài toán

# CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## HẠN CHẾ CỦA HỆ THỐNG:

Hệ thống hiện chỉ sử dụng dữ liệu từ một số nguồn, có thể dẫn đến hạn chế về độ chính xác của mô hình dự đoán.

Hệ thống chưa được tối ưu hóa về hiệu suất, có thể dẫn đến thời gian xử lý dữ liệu và thời gian dự đoán chậm.

Hệ thống chưa khảo sát thông tin thực tế, có thể dẫn đến các vấn đề về độ ổn định và khả năng mở rộng.

## KẾ HOẠCH PHÁT TRIỂN HỆ THỐNG:

Mở rộng nguồn dữ liệu: Thêm nhiều nguồn dữ liệu để tăng độ chính xác của mô hình dự đoán.

Tối ưu hóa hiệu suất: Tối ưu hóa hệ thống để cải thiện thời gian xử lý dữ liệu, thời gian dự đoán, và tốc độ truyền thông.

Khảo sát tình hình thực tế: Thử nghiệm hệ thống từ thông tin thực tế để đánh giá độ ổn định và khả năng mở rộng của hệ thống.

# CHƯƠNG 6: KẾT LUẬN

Dự án BigData-Dự-đoán-chất-lượng-giấc-ngủ đã thành công trong việc xây dựng một hệ thống dự đoán chất lượng giấc ngủ hiệu quả, sử dụng các công nghệ Big Data tiên tiến. Hệ thống đã đạt được các mục tiêu đề ra, bao gồm:

Kiến trúc hệ thống: Hệ thống được thiết kế và triển khai với kiến trúc phân tán, sử dụng các công nghệ Big Data phù hợp, đảm bảo tính ổn định và khả năng mở rộng.

Thu thập và xử lý dữ liệu: Hệ thống đã thu thập dữ liệu từ nhiều nguồn khác nhau, xử lý và tiền xử lý dữ liệu hiệu quả, đảm bảo chất lượng và tính nhất quán cho quá trình phân tích và dự đoán.

Đào tạo và đánh giá mô hình: Hệ thống đã xây dựng và huấn luyện các mô hình học máy tiên tiến, đạt được độ chính xác cao trong việc dự đoán.

Phát triển giao diện web: Hệ thống đã phát triển một giao diện web thân thiện với người dùng, cung cấp thông tin dự báo một cách trực quan, rõ ràng, và dễ hiểu.

Thực hiện và đánh giá hệ thống: Hệ thống đã được triển khai trong môi trường thực tế, thu thập dữ liệu và đánh giá hiệu quả của hệ thống dựa trên các chỉ số phù hợp.

Hệ thống dự đoán này có tiềm năng ứng dụng rộng rãi trong nhiều lĩnh vực, bao gồm:

Thúc đẩy lĩnh vực y học, cải thiện sức khỏe: Hệ thống cung cấp thông tin chính xác cho các lĩnh vực như y học hỗ trợ việc đưa ra các đánh giá sức khỏe hiệu quả.

Tuy nhiên, dự án vẫn còn một số hạn chế cần được khắc phục trong tương lai, như:

Mở rộng nguồn dữ liệu để tăng độ chính xác của mô hình dự đoán.

Tối ưu hóa hiệu suất của hệ thống để cải thiện thời gian xử lý dữ liệu và thời gian dự đoán.

Phát triển các tính năng mới cho giao diện web để nâng cao trải nghiệm người dùng.

Dự án BigData-Dự-đoán-chất-lượng-giấc-ngủ là một bước tiến quan trọng trong việc ứng dụng công nghệ Big Data vào lĩnh vực dự đoán chất lượng giấc ngủ. Hệ thống này hứa hẹn mang lại lợi ích thiết thực cho cộng đồng, hỗ trợ trong việc ứng phó với các nguy cơ mất ngủ, trầm cảm, ảnh hưởng đến sức khỏe và thúc đẩy phát triển các lĩnh vực y học.

# TÀI LIỆU THAM KHẢO

* 1. Tableau: <https://community.tableau.com/s/question/0D54T00000C5zUlSAJ/tableau-desktop-documentation>.
  2. Flask: <https://flask.palletsprojects.com/en/3.0.x/>.