TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP

**KHOA ĐIỆN TỬ**

**Bộ môn: Công nghệ Thông tin**.

**BÀI TẬP KẾT THÚC MÔN HỌC**

MÔN HỌC

**KHOA HỌC DỮ LIỆU**

Sinh viên: Hoàng Đức Thuần

Lớp: K57KMT

Giáo viên giảng dạy: TS.NGUYỄN VĂN HUY

LinkGitHub: [KHOAHOCDULIEU](mailto:https://github.com/thuanhdday/KHOAHOCDULIEU)

**Thái Nguyên – 2025**

|  |  |
| --- | --- |
| **TRƯỜNG ĐHKTCN** | **CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM** |
| **KHOA ĐIỆN TỬ** | ***Độc lập - Tự do - Hạnh phúc*** |

**BÀI TẬP KẾT THÚC MÔN HỌC**

**MÔN HỌC: KHOA HỌC DỮ LIỆU**

BỘ MÔN : CÔNG NGHỆ THÔNG TIN

*Sinh viên:* Hoàng Đức Thuần

*Lớp*: K57KMT

*Ngành: Kỹ thuật máy tính*

*Giáo viên hướng dẫn: TS. Nguyễn Văn Huy*

*Ngày giao đề*  *Ngày hoàn thành*

*Tên đề tài : Xây dựng hệ thống phân tích comment để xác định sản phẩm được yêu thích*.

*Yêu cầu :*

* Biểu đồ sản phẩm theo số lượng comment tích cực/tiêu cực.

**Các tính năng:**

* Text processing, sentiment analysis
* Visualization số liệu tương tác

|  |
| --- |
| **GIÁO VIÊN HƯỚNG DẪN** |
| *(Ký và ghi rõ họ tên)* |

**NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN**

### Thái Nguyên, ngày....tháng....năm 20....

## GIÁO VIÊN HƯỚNG DẪN

*(Ký ghi rõ họ tên)*

**MỤC LỤC**

[TRƯỜNG ĐẠI HỌC KỸ THUẬT CÔNG NGHIỆP 1](#_Toc199363660)

[**KHOA ĐIỆN TỬ** 1](#_Toc199363661)

[GIÁO VIÊN HƯỚNG DẪN 3](#_Toc199363664)

[CHƯƠNG 1. GIỚI THIỆU ĐẦU BÀI 6](#_Toc199363665)

[1. Giới thiệu đề tài 6](#_Toc199363666)

[CHƯƠNG 2: CƠ SỞ LÝ THUYẾT 10](#_Toc199363668)

[2.Các thư viện cần thiết 10](#_Toc199363669)

[CHƯƠNG 3. THIẾT KẾ VÀ XÂY DỰNG CHƯƠNG TRÌNH 12](#_Toc199363670)

[3.1.Thiết kế chương trình 12](#_Toc199363671)

[3.2. Analyze\_sentiment(text) 19](#_Toc199363672)

[3.3. Wordcloud sinh từ khóa phổ biến 19](#_Toc199363673)

[3.4. Các phần trực quan hóa (không gói thành hàm riêng nhưng có thể chia nhỏ): 19](#_Toc199363674)

[CHƯƠNG 4: THỰC NGHIỆM VÀ KẾT QUẢ 21](#_Toc199363675)

[4.1. Thực nghiệm 21](#_Toc199363676)

[4.2 Tổng Kết: 26](#_Toc199363677)

[4.3. Kết luận 26](#_Toc199363678)

### Đặt vấn đề

Trong thời đại thương mại điện tử bùng nổ, người tiêu dùng ngày càng có xu hướng tham khảo các đánh giá (review) trực tuyến trước khi quyết định mua một sản phẩm. Đặc biệt trên các nền tảng lớn như Amazon, hàng triệu bình luận được người dùng để lại mỗi ngày, phản ánh trực tiếp mức độ hài lòng và cảm nhận cá nhân về sản phẩm. Tuy nhiên, với khối lượng dữ liệu lớn và đa dạng về nội dung, việc phân tích thủ công các bình luận này là không khả thi.

Việc xây dựng một hệ thống tự động có khả năng phân tích nội dung comment và xác định mức độ yêu thích của người tiêu dùng đối với sản phẩm là một bài toán có ý nghĩa thực tiễn rất cao. Không chỉ giúp doanh nghiệp nắm bắt nhu cầu khách hàng, cải tiến sản phẩm và dịch vụ, mà còn hỗ trợ người tiêu dùng đưa ra quyết định mua sắm thông minh hơn.

Trong khuôn khổ đề tài này, chúng tôi sử dụng bộ dữ liệu **Consumer Reviews of Amazon Products**, bao gồm các bình luận của khách hàng về nhiều loại sản phẩm khác nhau. Thông qua các kỹ thuật xử lý ngôn ngữ tự nhiên (NLP) và học máy (Machine Learning), hệ thống sẽ phân tích nội dung comment để xác định sản phẩm nào đang được người tiêu dùng yêu thích, từ đó cung cấp một công cụ hỗ trợ ra quyết định hiệu quả cho các nhà quản lý, nhà bán hàng và khách hàng.

# CHƯƠNG 1. GIỚI THIỆU ĐẦU BÀI

## Giới thiệu đề tài

**Đầu bài:**  
Xây dựng hệ thống phân tích comment để xác định sản phẩm được yêu thích.

**Đầu vào:**

# Dữ liệu từ [Consumer Reviews of Amazon Products](https://www.kaggle.com/datasets/datafiniti/consumer-reviews-of-amazon-products)

**Đầu ra:**

* Biểu đồ sản phẩm theo số lượng comment tích cực/tiêu cực.

**Các tính năng:**

* Text processing, sentiment analysis
* Visualization số liệu tương tác

**Hướng dẫn:**

* Xử lý văn bản và phân tích cảm xúc.
* Visualization kết quả phân tích.

### Thách thức trong bài toán:

### ****Đa dạng và nhiễu trong ngôn ngữ tự nhiên****

Các bình luận từ người dùng không tuân theo chuẩn ngữ pháp chặt chẽ, chứa lỗi chính tả, từ viết tắt, biểu tượng cảm xúc, tiếng lóng, hoặc kết hợp nhiều ngôn ngữ.

Người dùng có thể sử dụng từ ngữ mang tính mỉa mai, đa nghĩa hoặc ngữ cảnh đặc biệt, gây khó khăn cho việc xác định chính xác cảm xúc.

**Độ dài và cấu trúc đánh giá không đồng nhất**

Một số bình luận chỉ gồm vài từ đơn giản (ví dụ: "Good!", "Terrible!"), trong khi một số khác là đoạn văn dài mô tả chi tiết.

Điều này ảnh hưởng đến khả năng đánh giá cảm xúc một cách chính xác, đặc biệt với các mô hình đơn giản như TextBlob vốn dựa trên phân tích câu và từ.

### ****Phân loại cảm xúc chủ quan****

Phân loại cảm xúc là một tác vụ có tính **chủ quan cao**, do cùng một bình luận có thể được diễn giải khác nhau tuỳ theo ngữ cảnh hoặc người đọc.

Ví dụ, bình luận "It works... for now." có thể mang sắc thái nghi ngờ, nhưng thuật toán đơn giản có thể hiểu là trung lập hoặc tích cực.

### ****Mối liên hệ phức tạp giữa điểm số và cảm xúc****

Một đánh giá 5 sao có thể đi kèm với một bình luận tiêu cực (do lỗi người dùng hoặc hệ thống), và ngược lại.

Do đó, cần có phân tích kết hợp cả văn bản và điểm số thay vì chỉ dựa vào một yếu tố duy nhất.

**Khối lượng dữ liệu lớn**

Bộ dữ liệu Amazon chứa hàng trăm nghìn đến hàng triệu đánh giá. Việc xử lý dữ liệu với khối lượng lớn yêu cầu hệ thống hiệu quả và khả năng mở rộng.

### ****Phân tích cảm xúc theo sản phẩm****

Không chỉ dừng lại ở từng đánh giá riêng lẻ, bài toán cần **tổng hợp cảm xúc ở cấp độ sản phẩm**, xác định sản phẩm được yêu thích nhất dựa trên tỷ lệ và số lượng comment tích cực/tiêu cực — điều này yêu cầu thêm bước xử lý nhóm dữ liệu và trực quan hóa.

### 1.2.Các kiến thức được vận dụng:

Trong quá trình thực hiện bài toán, các kiến thức sau đã được áp dụng:

**Xử lý dữ liệu với Pandas**: Đọc dữ liệu, loại bỏ dòng trống, thống kê và nhóm theo sản phẩm.

**Xử lý văn bản**: Làm sạch văn bản, chuyển đổi định dạng phù hợp để phân tích.

**Phân tích cảm xúc với TextBlob**: Xác định mức độ tích cực/trung lập/tiêu cực của mỗi đánh giá dựa trên độ phân cực (polarity).

**Trực quan hóa dữ liệu**:

Dùng **Seaborn, Matplotlib** để vẽ biểu đồ cảm xúc, điểm đánh giá và so sánh theo sản phẩm.

Dùng **WordCloud** để hiển thị từ khóa nổi bật trong các đánh giá.

**Lập trình Python**: Áp dụng các kỹ năng viết hàm, xử lý lỗi, kiểm tra điều kiện, và làm việc với dữ liệu lớn.

**Phân tích tổng hợp**: Kết hợp cảm xúc và sản phẩm để xác định sản phẩm được yêu thích dựa trên số lượng phản hồi tích cực/tiêu cực.

# CHƯƠNG 2: CƠ SỞ LÝ THUYẾT

## 2.Các thư viện cần thiết

### pandas

**chức năng**: hỗ trợ đọc, xử lý và phân tích dữ liệu dạng bảng (csv, excel, sql...).

**vai trò trong bài toán**: được dùng để đọc dữ liệu đánh giá từ file .csv, loại bỏ dữ liệu trống, nhóm theo sản phẩm, và xử lý cột chứa đánh giá văn bản.

### matplotlib

**chức năng**: thư viện vẽ biểu đồ cơ bản trong python.

**vai trò trong bài toán**: dùng để hiển thị các biểu đồ như biểu đồ cột, biểu đồ boxplot để phân tích điểm đánh giá và cảm xúc.

### seaborn

**chức năng**: thư viện trực quan hóa dữ liệu nâng cao, được xây dựng trên matplotlib.

**vai trò trong bài toán**: dùng để vẽ các biểu đồ phân phối cảm xúc, phân bố điểm số và so sánh cảm xúc giữa các sản phẩm với màu sắc trực quan và dễ hiểu hơn.

### textblob

**chức năng**: thư viện xử lý ngôn ngữ tự nhiên đơn giản, dễ dùng.

**vai trò trong bài toán**: phân tích văn bản đánh giá để xác định cảm xúc (positive/neutral/negative) dựa trên chỉ số **polarity**.

### wordcloud

**chức năng**: tạo đám mây từ khóa (word cloud) trực quan từ văn bản.**vai trò trong bài toán**: hiển thị các từ khóa thường xuất hiện trong các đánh giá tích cực hoặc tiêu cực, giúp nắm bắt nội dung nổi bật mà người dùng đề cập.

# CHƯƠNG 3. THIẾT KẾ VÀ XÂY DỰNG CHƯƠNG TRÌNH

## 3.Thiết kế chương trình

### 3.1 Thiết kế modul

#### ****Mô tả các module chính****

Chương trình gồm các module chính sau

### ****Module nhập và xử lý dữ liệu****

**Chức năng**:

Đọc dữ liệu đánh giá từ file .csv.

Loại bỏ dòng trống, chuẩn hóa văn bản.

**Thư viện sử dụng**: pandas

### ****Module phân tích cảm xúc****

**Chức năng**:

Áp dụng textblob để phân tích độ phân cực của nội dung đánh giá.

Gán nhãn cảm xúc: tích cực*,* trung lập*,* tiêu cực.

**Thư viện sử dụng**: textblob

### ****Module trực quan hóa****

**Chức năng**:

Vẽ biểu đồ phân phối cảm xúc và điểm đánh giá.

Vẽ biểu đồ cảm xúc theo sản phẩm (top 10).

**Thư viện sử dụng**: matplotlib, seaborn

### ****Module tạo wordcloud****

**Chức năng**:

Tạo đám mây từ khóa cho đánh giá tích cực và tiêu cực.

Làm nổi bật các từ khóa thường gặp trong từng nhóm cảm xúc.

**Thư viện sử dụng**: wordcloud

### ****Module tổng hợp và đánh giá sản phẩm****

**Chức năng**:

Tổng hợp số lượng comment theo cảm xúc cho từng sản phẩm.

Xác định sản phẩm được yêu thích nhất dựa trên số lượng đánh giá tích cực.

### 3.2. Sơ đồ khối hệ thốngA diagram of a diagram Description automatically generated with medium confidence

### 3.3. Biểu đồ phân cấp chức năng

Dưới đây là biểu đồ phân cấp chức năng của hệ thống, mô tả các chức năng chính:

Biểu đồ phân cấp chức năng:

HỆ THỐNG

phân tích đánh giá sản phẩm

Đánh giá sản phẩm yêu thích

Trực quan hóa dữ liệu

Thống kê và tổng hợp

Tiền xử lý dữ liệu

Phân tích cảm xúc

Tính tỷ lệ đánh giá tích cực

Xác định sản phẩm yêu thích

-

Đọc dữ liệu CSV

Loại bỏ dòng trống

Chuẩn hóa văn bản

Biểu đồ phân phối cảm xúc

Biểu đồ điểm đánh giá

WordCloud theo cảm xúc

Nhóm theo sản phẩm

Đếm số lượng cảm xúc

Áp dụng TextBlob

Phân loại cảm xúc: Tích cực / Trung lập / Tiêu cực

### 3.4. Sơ đồ khối các thuật toán chính

#### Mô tả các khối thuật toán chính.Dưới đây là mô tả các khối thuật toán và quan hệ đầu vào/ra:

A diagram of a company

Description automatically generated

### ****A. Nhập dữ liệu đánh giá sản phẩm****

**Dữ liệu nguồn:** tệp csv từ tập “consumer reviews of amazon products”.

**Nội dung chính:** chứa thông tin như tên sản phẩm, đánh giá văn bản, điểm sao, ngày đánh giá,...

**Mục tiêu:** đưa vào chương trình để xử lý và phân tích.

### ****b. Tiền xử lý văn bản****

**Các bước chính:**

Loại bỏ các dòng có nan trong cột reviews.text.

Chuyển toàn bộ văn bản về chữ thường.

Loại bỏ ký tự đặc biệt, khoảng trắng thừa.

Có thể mở rộng thêm: bỏ stopwords, stemming nếu cần.

**Kết quả:** văn bản sạch sẵn sàng cho phân tích cảm xúc.

### ****c. Phân tích cảm xúc bằng textblob****

**Công cụ:** textblob – thư viện nlp đơn giản.

**Cách hoạt động:**

Với mỗi đánh giá, tính điểm **polarity** (độ cảm xúc).

Phân loại:

> 0.1: tích cực

< -0.1: tiêu cực

Ngược lại: trung lập

**Đầu ra:** cột mới cảm xúc chứa loại cảm xúc của từng đánh giá

.

### ****d. Tính toán thống kê cảm xúc theo sản phẩm****

**Xử lý:**

Đếm số lượng đánh giá theo từng loại cảm xúc (tích cực, tiêu cực, trung lập).

Gom nhóm theo name (tên sản phẩm).

**Mục tiêu:** xác định sản phẩm nào có nhiều cảm xúc tích cực nhất.

### ****e. Trực quan hóa dữ liệu****

**Mục đích:** biến kết quả phân tích thành biểu đồ dễ hiểu.

**Các loại biểu đồ:**

**F1:** biểu đồ cột đếm số cảm xúc toàn bộ.

**F2:** boxplot thể hiện mối liên hệ giữa cảm xúc và điểm đánh giá.

**F3:** biểu đồ cảm xúc theo sản phẩm (top 10).

**F4:** wordcloud để nhìn từ khóa xuất hiện nhiều trong cảm xúc tích cực/tiêu cực.

### ****g. Đánh giá sản phẩm yêu thích****

**Tiêu chí lựa chọn sản phẩm yêu thích:**

Tỷ lệ bình luận tích cực cao nhất.

Kết hợp với điểm số đánh giá trung bình.

**Ứng dụng:** giúp doanh nghiệp xác định sản phẩm nổi bật theo cảm xúc khách hàng.

### 3.5. Cấu trúc dữ liệu

#### Mô tả các bảng dữ liệu và các trường thông tin

Hệ thống sử dụng các bảng dữ liệu sau:

**1. Bảng dữ liệu chính:**

Đây là bảng chứa toàn bộ dữ liệu đánh giá sản phẩm được đọc từ tệp CSV.

| **Tên cột** | **Kiểu dữ liệu** | **Mô tả ngắn** |
| --- | --- | --- |
| name | string | Tên sản phẩm được đánh giá. |
| reviews.text | string | Nội dung văn bản đánh giá của người dùng. |
| reviews.rating | float (hoặc int) | Điểm số đánh giá sao từ 1 đến 5. |
| reviews.title | string | Tiêu đề ngắn của đánh giá (nếu có). |
| reviews.date | datetime | Ngày người dùng đăng đánh giá. |
| reviews.username | string | Tên người dùng để lại đánh giá (có thể bị ẩn danh). |

**2. Cột tạo thêm sau xử lý (bổ sung trong chương trình)**

| **Cột mới** | **Nguồn sinh ra** | **Mô tả** |
| --- | --- | --- |
| Cảm xúc | Tính từ reviews.text | Phân loại cảm xúc: 'Tích cực', 'Trung lập', 'Tiêu cực' |
| review\_length | Tính từ độ dài reviews.text | (Tùy chọn) Độ dài chuỗi văn bản để phân tích thêm. |

**3.Ví dụ mẫu dữ liệu:**

| **name** | **reviews.text** | **reviews.rating** | **Cảm xúc** |
| --- | --- | --- | --- |
| Echo Dot | Great sound for the price! | 5 | Tích cực |
| Kindle Paper | Poor battery after update. | 2 | Tiêu cực |
| Fire Tablet | It's ok, but not amazing. | 3 | Trung lập |

### 3.6. Chương Trình

## ****1.**** Analyze\_sentiment(text)

### Chức năng:

Phân tích cảm xúc của một chuỗi văn bản đầu vào (review) bằng thư viện textblob.

### Mô tả logic:

Tính polarity (mức độ tích cực/tính tiêu cực) của văn bản.

Trả về một trong ba loại:

* 'Tích cực' nếu polarity > 0.1
* 'Tiêu cực' nếu polarity < -0.1
* 'Trung lập' nếu nằm giữa

### Đầu Vào – text:

Một câu hoặc đoạn đánh giá sản phẩm.

### Đầu ra:

Một nhãn cảm xúc: "tích cực", "trung lập" hoặc "tiêu cực"

## ****2.**** Wordcloud ****sinh từ khóa phổ biến****

Tạo và hiển thị wordcloud từ nội dung đánh giá:

### Gồm 2 phần:

* Với đánh giá "tích cực" → positive\_reviews
* Với đánh giá "tiêu cực" → negative\_reviews

## ****3. Các phần trực quan hóa (không gói thành hàm riêng nhưng có thể chia nhỏ):****

### Draw\_sentiment\_distribution(df)

* Biểu đồ cột phân phối cảm xúc (countplot).

### Draw\_rating\_distribution(df)

* Biểu đồ số sao người dùng đánh giá (countplot).

### Draw\_sentiment\_vs\_rating(df)

* Biểu đồ boxplot: điểm đánh giá (sao) theo từng loại cảm xúc.

### Draw\_top\_products\_sentiment(df)

* Biểu đồ các sản phẩm có nhiều comment tích cực/tiêu cực nhất (stacked bar chart).

*Tóm tắt chương :*

*Chương 3 mô tả các chuyên môn sử dụng, các hàm, thuật toán xử dụng,các modul, biểu đồ phân cấp chức năng, biểu đồ quan hệ thuật toán, mô tả bảng dữ liệu.*

# CHƯƠNG 4: THỰC NGHIỆM VÀ KẾT QUẢ

## 4.1.1 Thực nghiệm A graph with different colored bars Description automatically generatedA green and grey rectangular object with white background Description automatically generated

## A diagram of different colors Description automatically generated

A graph with blue and red bars

Description automatically generated A close up of words

Description automatically generated

A close up of words

Description automatically generated

### ****4.1.2 Ưu điểm:****

**Ứng dụng thực tiễn cao**:  
Giải quyết một bài toán phổ biến trong thương mại điện tử – phân tích phản hồi khách hàng để hiểu mức độ yêu thích sản phẩm.

**Sử dụng thư viện mạnh mẽ và phổ biến**:

TextBlob để phân tích cảm xúc (dễ triển khai).

Seaborn, Matplotlib để trực quan hóa dữ liệu.

WordCloud giúp minh họa trực quan từ khóa nổi bật.

**Dễ hiểu, dễ triển khai**:

Các bước xử lý được chia theo từng khối logic (đọc dữ liệu, phân tích cảm xúc, trực quan hóa).

Có thể mở rộng thêm các mô hình học máy nếu cần.

**Biểu đồ đa dạng và hữu ích**:

Phân phối cảm xúc, phân phối điểm sao, mối quan hệ giữa cảm xúc và điểm số, đánh giá theo từng sản phẩm.

### 4.1.3 ****Hạn chế và điểm cần cải thiện:****

**Mô hình phân tích cảm xúc còn đơn giản**:  
textblob là mô hình từ điển (rule-based), không xử lý được ngữ cảnh phức tạp.  
 Có thể thay bằng mô hình học sâu như bert, vader, hoặc mô hình huấn luyện riêng.

**Thiếu xử lý tiếng ồn văn bản**:  
Chưa xử lý các bước như:

* + Chuyển chữ thường
  + Loại bỏ stopwords, ký tự đặc biệt, từ không mang ý nghĩa

**Chưa có giao diện tương tác**:  
hiện chương trình chỉ chạy dạng script.  
 Có thể kết hợp streamlit để tạo giao diện web thân thiện hơn.

**Chưa lưu kết quả**:  
 Chưa có chức năng lưu dữ liệu phân tích (csv, excel, database,…).

## 4.1.4 ****Tổng Kết:****

Chương trình là một **hệ thống phân tích cảm xúc cơ bản nhưng hiệu quả**, phù hợp với các bài toán nhập môn khai phá dữ liệu và xử lý ngôn ngữ tự nhiên. Dễ triển khai, có thể nâng cấp dần theo yêu cầu thực tế.

## 4.2. Kết luận

#### 1.Sản phẩm đã làm được những gì

Xây dựng thành công hệ thống phân tích cảm xúc từ các đánh giá sản phẩm trên Amazon.

Áp dụng **TextBlob** để phân loại cảm xúc thành ba nhóm: **Tích cực**, **Trung lập**, **Tiêu cực**.

Trực quan hóa dữ liệu bằng các biểu đồ: phân phối cảm xúc, điểm số đánh giá, mối quan hệ giữa cảm xúc và điểm số, cũng như các sản phẩm được yêu thích nhất.

Tạo WordCloud thể hiện các từ khóa phổ biến trong các bình luận tích cực và tiêu cực.

### 2. ****Kiến thức và kỹ năng đã học được:****

Hiểu quy trình **xử lý văn bản (Text Processing)** và áp dụng trong ngữ cảnh thực tế.

Áp dụng **phân tích cảm xúc cơ bản** trong khai phá dữ liệu khách hàng.

Thực hành trực quan hóa dữ liệu bằng các thư viện như **Matplotlib**, **Seaborn**, và **WordCloud**.

Nắm vững cách làm việc với **dữ liệu thực tế** từ tệp CSV lớn và không sạch.

Tư duy phân tích và **chuyển đổi yêu cầu nghiệp vụ thành giải pháp lập trình**.

### 3.****Hướng cải tiến trong tương lai:****

1. **Nâng cấp mô hình phân tích cảm xúc**:
   * Sử dụng mô hình học sâu như **BERT**, **VADER** hoặc mô hình fine-tune từ dữ liệu thực tế.
   * Cải thiện độ chính xác khi xử lý câu có nhiều lớp nghĩa.
2. **Tiền xử lý dữ liệu nâng cao hơn**:
   * Loại bỏ stopwords, từ vô nghĩa, từ lặp.
   * Lemmatization/stemming.
3. **Thêm tính năng đánh giá sản phẩm tổng quan**:
   * Tạo bảng xếp hạng sản phẩm theo mức độ hài lòng.
4. **Xây dựng giao diện người dùng**:
   * Dùng **Streamlit hoặc Flask** để triển khai web đơn giản, giúp người dùng upload tệp và xem kết quả trực tiếp.
5. **Triển khai mô hình thực tế**:
   * Tích hợp vào hệ thống bán hàng để tự động theo dõi phản hồi khách hàng.

*Tóm tắt:*

*Chương 4 tổng kết,giới thiệu qua phần kết quả của hệ thống, nhận xét, đánh giá mô hình.*

## ****Lời cảm ơn****

Em xin chân thành cảm ơn **Thầy TS.Nguyễn Văn Huy** – giảng viên hướng dẫn, người đã tận tình chỉ bảo, định hướng và hỗ trợ em trong suốt quá trình thực hiện bài tập này. Những góp ý quý báu của cô là nguồn động lực và kim chỉ nam giúp em hoàn thiện sản phẩm một cách tốt nhất.

Em cũng xin gửi lời cảm ơn đến các thầy cô trong bộ môn đã tạo điều kiện thuận lợi về kiến thức và môi trường học tập để em có thể áp dụng và phát triển kỹ năng phân tích dữ liệu thực tế.

Cuối cùng, em xin cảm ơn bạn bè và gia đình đã luôn đồng hành và động viên em trong suốt quá trình học tập và hoàn thiện bài làm.

**Em xin chân thành cảm ơn!**