

## **BÁO CÁO MÔN HỌC LẬP TRÌNH THIẾT BỊ DI ĐỘNG**

### **ĐỀ TÀI**

### **ỨNG DỤNG TMĐT CHO ĐIỆN THOẠI TÍCH HỢP PHÂN TÍCH CẢM XÚC THEO KHÍA CẠNH**

Giảng viên hướng dẫn: **Nguyễn Ánh**

Sinh viên thực hiện:

- 2280603696 - Nguyễn Quang Vinh**
- 2280603283 - Đặng Doanh Toại**
- 2280618597 - Trần Đình Ty**
- 2280603036 - Phan Thanh Thiên**
- 2280602828 - Trần Tấn Tài**

# Mục lục

<b>LỜI MỞ ĐẦU</b>	<b>1</b>
<b>LỜI CAM ĐOAN</b>	<b>2</b>
<b>Chương 1. TỔNG QUAN HỆ THỐNG</b>	<b>1</b>
1.1 Giới thiệu . . . . .	1
1.2 Yêu cầu chức năng . . . . .	1
1.3 Kiến trúc hệ thống . . . . .	1
1.4 Quy trình thanh toán MoMo . . . . .	2
<b>Chương 2. CƠ SỞ LÝ THUYẾT</b>	<b>3</b>
2.1 Công nghệ Backend . . . . .	3
2.2 Công nghệ Mobile . . . . .	3
2.3 Công nghệ tích hợp . . . . .	3
2.4 ABSA - Aspect-Based Sentiment Analysis . . . . .	3
<b>Chương 3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG</b>	<b>5</b>
3.1 Phân tích Use Case . . . . .	5
3.2 Đặc tả Use Case chính . . . . .	6
3.3 Luồng màn hình Customer App . . . . .	6
3.4 Quy trình đánh giá với ABSA . . . . .	6
<b>Chương 4. THIẾT KẾ CƠ SỞ DỮ LIỆU</b>	<b>7</b>
4.1 Collection Users . . . . .	7
4.2 Collection Products . . . . .	7
4.3 Collection Orders và Payments . . . . .	8
4.4 Collection Reviews . . . . .	8
4.5 Collection Carts và Coupons . . . . .	8
<b>Chương 5. MÔ HÌNH, THỰC NGHIỆM VÀ ĐÁNH GIÁ</b>	<b>9</b>
5.1 Mô hình . . . . .	9
5.1.1 Single-Task Learning (STL) . . . . .	9
5.1.2 Multi-Task Learning (MTL): . . . . .	9
5.2 Tập dữ liệu . . . . .	9
5.3 Đánh nhãn dữ liệu . . . . .	10
5.4 Tiền xử lý dữ liệu . . . . .	11
5.5 Kết quả và đánh giá . . . . .	12

5.5.1 Metrics . . . . .	12
5.5.2 Cài đặt thí nghiệm . . . . .	12
5.5.3 Baseline Models . . . . .	13
5.5.4 Kết Quả . . . . .	13
5.5.5 Phân tích kết quả . . . . .	14
5.6 Kết luận . . . . .	15
5.6.1 Kết quả đạt được . . . . .	15
5.6.2 Hạn chế cần khắc phục . . . . .	15
5.6.3 Hướng phát triển . . . . .	15
<b>TÀI LIỆU THAM KHẢO</b>	<b>16</b>
<b>ĐÁNH GIÁ THÀNH VIÊN</b>	<b>17</b>

**Danh sách hình vẽ**

1	Sơ đồ Kiến trúc Hệ thống VeritaShop . . . . .	2
2	Sơ đồ Use Case Tổng quát hệ thống VeritaShop . . . . .	5

## Danh sách bảng

1	Đặc tả các Use Case chính . . . . .	6
2	Cấu trúc Collection Users . . . . .	7
3	Cấu trúc Collection Products . . . . .	7
4	Cấu trúc Collection Reviews . . . . .	8
5	11 Aspects trong mô hình ABSA . . . . .	10
6	Tổng hợp Fleiss' Kappa qua 2 giai đoạn . . . . .	11
7	Fleiss' Kappa theo khía cạnh (cải thiện từ Phase 1 → Phase 2) . . .	11
8	Kết quả các mô hình BiLSTM . . . . .	13
9	Kết quả các mô hình PhoBERT . . . . .	13
10	Kết quả các mô hình VisoBERT . . . . .	13
11	Bảng so sánh F1-Score của các mô hình . . . . .	14
12	Bảng đánh giá công việc thực hiện và tỷ lệ đóng góp . . . . .	17

## LỜI MỞ ĐẦU

Trong bối cảnh thương mại điện tử phát triển vượt bậc trên toàn cầu, việc mua sắm trực tuyến đã trở thành xu hướng không thể đảo ngược. Đặc biệt tại Việt Nam, thị trường điện thoại di động luôn sôi động với nhu cầu tiêu dùng cao và sự cạnh tranh gay gắt giữa các thương hiệu. Người tiêu dùng ngày càng có xu hướng tìm kiếm, so sánh và mua sắm điện thoại thông qua các nền tảng di động nhờ sự tiện lợi, đa dạng lựa chọn và khả năng tiếp cận thông tin sản phẩm một cách toàn diện.

Nắm bắt được nhu cầu thực tiễn đó, đề án này tập trung nghiên cứu và phát triển “VeritaShop” - một ứng dụng thương mại điện tử di động chuyên biệt cho lĩnh vực kinh doanh điện thoại. Hệ thống không chỉ cung cấp trải nghiệm mua sắm trực tuyến hoàn chỉnh với đầy đủ các tính năng từ tìm kiếm, lọc sản phẩm, giỏ hàng đến thanh toán trực tuyến, mà còn tích hợp các công nghệ tiên tiến như phân tích cảm xúc tự động (Aspect-Based Sentiment Analysis) cho đánh giá sản phẩm và Content Moderation để kiểm duyệt nội dung.

Với sự kết hợp giữa ứng dụng di động đa nền tảng xây dựng bằng Flutter và hệ thống quản trị mạnh mẽ, dự án mong muốn kiến tạo một giải pháp thương mại điện tử toàn diện, hiện đại và tối ưu hóa trải nghiệm cho cả người mua hàng lẫn người bán, góp phần thúc đẩy sự phát triển của ngành bán lẻ điện thoại trực tuyến tại Việt Nam.

## LỜI CAM ĐOAN

Nhóm chúng xin cam đoan rằng đề án **”Ứng dụng thương mại điện tử cho điện thoại tích hợp phân tích cảm xúc theo khía cạnh** là kết quả nghiên cứu và thực hiện của tất cả thành viên trong nhóm, dưới sự hướng dẫn khoa học của giảng viên.

Toàn bộ nội dung, số liệu và kết quả trình bày trong đề án này là trung thực, được chúng tôi tìm hiểu, phân tích và xây dựng dựa trên kiến thức chuyên môn và các tài liệu tham khảo uy tín. Các chức năng của hệ thống như tích hợp cổng thanh toán MoMo, phân tích cảm xúc đánh giá sản phẩm (ABSA - Aspect-Based Sentiment Analysis), và kiểm duyệt nội dung tự động (Content Moderation), Tìm kiếm bằng giọng nói (Speech to text) đều được nhóm tự nghiên cứu và phát triển. Nhóm cam kết không sao chép nguyên văn hoặc sử dụng trái phép nội dung từ bất kỳ nguồn nào mà không có trích dẫn rõ ràng.

Nếu phát hiện có bất kỳ hành vi đạo văn hoặc gian lận nào trong quá trình thực hiện đề án này, chúng tôi xin chịu hoàn toàn trách nhiệm trước Nhà trường và Hội đồng đánh giá.

Chúng tôi xin chân thành cảm ơn!

# Chương 1. TỔNG QUAN HỆ THỐNG

## 1.1. Giới thiệu

Trong bối cảnh thương mại điện tử phát triển mạnh mẽ, thị trường điện thoại thông minh tại Việt Nam luôn duy trì sự sôi động với doanh số hàng chục triệu máy mỗi năm. Các nền tảng lớn như Shopee, Lazada, Tiki tuy phổ biến nhưng thiếu khả năng phân tích cảm xúc tự động từ bình luận sản phẩm.

**VeritaShop** được xây dựng nhằm giải quyết các hạn chế trên:

- Ứng dụng TMĐT chuyên biệt cho điện thoại với thông số kỹ thuật chi tiết
- Phân tích cảm xúc theo khía cạnh (ABSA) giúp người mua hiểu rõ ưu/nhược điểm
- Kiểm duyệt nội dung tự động bảo vệ cộng đồng
- Hệ thống quản trị toàn diện cho người bán

## 1.2. Yêu cầu chức năng

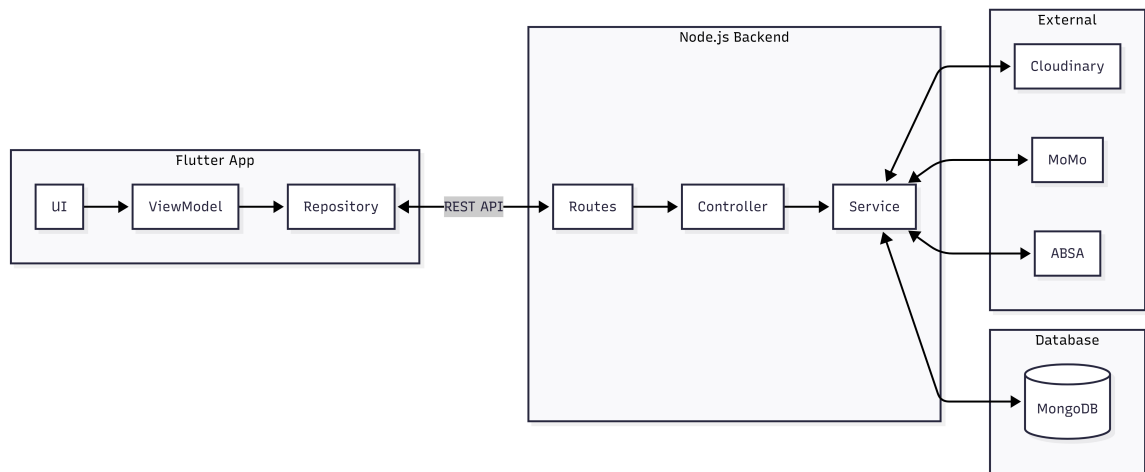
**Customer App (Mobile):**

- **Xác thực:** Đăng ký, đăng nhập JWT, quên mật khẩu, PIN Lock bảo mật
- **Sản phẩm:** Xem danh sách, tìm kiếm, lọc theo brand/giá/tình trạng
- **Mua sắm:** Giỏ hàng, wishlist, checkout, áp dụng coupon
- **Thanh toán:** COD và MoMo (Deep Link + IPN callback)
- **Đánh giá:** Viết review 1-5 sao, xem phân tích ABSA tự động

## 1.3. Kiến trúc hệ thống

Hệ thống VeritaShop được thiết kế theo mô hình Client-Server với kiến trúc RESTful API:

- **Mobile Client:** Flutter 3.8.1+ (Dart) - đa nền tảng Android, iOS, Web
- **Backend Server:** Node.js 18+ với Express.js framework
- **Database:** MongoDB Atlas (NoSQL) với Mongoose ODM
- **Tích hợp:** Cloudinary (ảnh), MoMo (thanh toán), VisoBERT (ABSA)



Hình 1. Sơ đồ Kiến trúc Hệ thống VeritaShop

## 1.4. Quy trình thanh toán MoMo

Hệ thống tích hợp MoMo Payment Gateway với Deep Link và IPN callback:

1. Người dùng checkout → Backend tạo Order với status "pending"
2. Backend gọi MoMo API với HMAC-SHA256 signature → Nhận deeplink
3. Flutter sử dụng url\_launcher mở MoMo app
4. Người dùng xác nhận thanh toán trong app MoMo
5. MoMo gửi IPN callback → Backend verify signature → Cập nhật paymentStatus

## Chương 2. CƠ SỞ LÝ THUYẾT

### 2.1. Công nghệ Backend

**Node.js** là nền tảng JavaScript phía máy chủ với kiến trúc non-blocking, event-driven, phù hợp cho hệ thống TMĐT có lưu lượng truy cập lớn.

**Express.js** là framework cho Node.js, cung cấp routing, middleware và cơ chế xử lý lỗi tập trung.

**MongoDB** là CSDL NoSQL dạng document với schema linh hoạt, hỗ trợ tìm kiếm văn bản và aggregation. **Mongoose ODM** cung cấp cơ chế định nghĩa schema và validation.

### 2.2. Công nghệ Mobile

**Flutter** là bộ công cụ UI của Google, phát triển ứng dụng đa nền tảng (Android, iOS, Web, Desktop) từ cùng một mã nguồn. Sử dụng ngôn ngữ Dart với Hot Reload, kiến trúc widget-based, biên dịch sang mã native.

Thư viện chính: **provider** (state management), **dio** (API calls), **cached\_network\_image** (image cache), **flutter\_secure\_storage** (secure token storage), **url\_launcher** và **app\_links** (MoMo Deep Link).

### 2.3. Công nghệ tích hợp

Hệ thống tích hợp các dịch vụ bên thứ ba nhằm đảm bảo khả năng thanh toán, lưu trữ, bảo mật và tương tác thông minh.

MoMo Payment Gateway được sử dụng cho thanh toán trực tuyến.

Cloudinary được tích hợp như một CDN để lưu trữ, tối ưu và phân phối hình ảnh toàn cầu.

Hệ thống bảo mật sử dụng JWT, bcrypt, HTTPS và *express-validator*, hỗ trợ mã PIN tùy chọn.

Chức năng tìm kiếm bằng giọng nói sử dụng mô hình *GPT-4o-mini-transcribe* để chuyển đổi âm thanh thành văn bản phục vụ tìm kiếm.

Cơ chế kiểm duyệt nội dung tự động phát hiện các đánh giá vi phạm và gắn cờ *isFlagged=true* để quản trị viên duyệt thủ công.

### 2.4. ABSA - Aspect-Based Sentiment Analysis

#### Giới thiệu bài toán

Sự phát triển của thương mại điện tử Việt Nam làm gia tăng nhu cầu phân tích cảm xúc từ bình luận khách hàng. Tuy nhiên, đặc thù tiếng Việt phi chuẩn (tiếng lóng, viết tắt, emoji) gây khó khăn cho các phương pháp truyền thống. Do

đó, VeritaShop được xây dựng nhằm tích hợp mô hình ngôn ngữ tiên tiến để tự động phân loại cảm xúc và hỗ trợ ra quyết định.

## Các nghiên cứu liên quan

**Phương pháp truyền thống** như Bag-of-Words, TF-IDF kết hợp SVM hoặc Naive Bayes còn hạn chế trong việc nắm bắt ngữ nghĩa sâu và phụ thuộc nhiều vào tiền xử lý thủ công, đặc biệt đối với tiếng Việt.

**Học sâu** với CNN, LSTM và Transformer đã trở thành nền tảng cho các mô hình ngôn ngữ tiên huấn luyện. Các mô hình chuyên biệt cho tiếng Việt như PhoBERT và VisoBERT cho hiệu quả vượt trội so với mBERT đa ngôn ngữ.

**Hạn chế hiện tại** gồm khả năng xử lý bình luận ngắn và tiếng lóng còn hạn chế, mất cân bằng dữ liệu và thiếu nhãn chuyên biệt cho ngữ cảnh TMDT, đồng thời việc tích hợp thời gian thực còn gặp nhiều thách thức.

**Nghiên cứu này** bổ sung bộ dữ liệu mới, lựa chọn mô hình VisoBERT-STL, áp dụng Focal Loss và thiết kế kiến trúc tối ưu nhằm khắc phục các hạn chế trên.

## Focal Loss - Xử lý mất cân bằng dữ liệu

**Vấn đề:** Trong dữ liệu bình luận, số lượng đánh giá tích cực thường áp đảo tiêu cực/trung lập (70–80%). Cross-Entropy truyền thống chỉ học tốt mẫu “dễ” (lớp đa số), bỏ qua mẫu “khó” (lớp thiểu số).

**Giải pháp:** Focal Loss (Lin et al.) điều chỉnh trọng số dựa trên độ khó mẫu, giảm ảnh hưởng mẫu dễ, tăng cường học từ mẫu khó.

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t)$$

Trong đó:

$$-\alpha_t = \begin{cases} p & \text{nếu } y = 1 \\ 1 - p & \text{nếu } y = 0 \end{cases}$$

- $\alpha_t$ : hệ số cân bằng trọng số lớp
- $\gamma$ : focusing parameter, khi  $\gamma = 0$  trở về Cross-Entropy

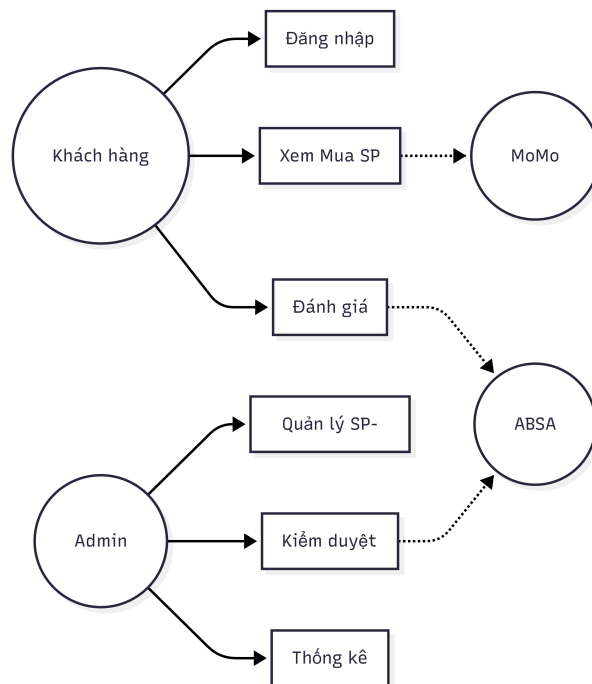
**Áp dụng:**  $\gamma = 2$ ,  $\alpha$  điều chỉnh theo tỷ lệ phân bố lớp. Cải thiện F1-score cho lớp thiểu số (Negative).

## Chương 3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

### 3.1. Phân tích Use Case

Các tác nhân chính:

- **Khách hàng:** Đăng ký/nhập, xem sản phẩm, tìm kiếm, giỏ hàng, thanh toán, đánh giá
- **Quản trị viên:** Quản lý sản phẩm, đơn hàng, người dùng, coupon, kiểm duyệt đánh giá
- **Hệ thống MoMo:** Xử lý thanh toán, gửi kết quả IPN
- **Hệ thống AI:** Phân tích ABSA, Content Moderation



Hình 2. Sơ đồ Use Case Tổng quát hệ thống VeritaShop

### 3.2. Đặc tả Use Case chính

Bảng 1. Đặc tả các Use Case chính

Use Case	Tác nhân	Mô tả
Tìm kiếm sản phẩm	Khách hàng	Lọc brand/giá/tình trạng, sắp xếp giá/rating/mới
Thanh toán đơn hàng	Khách hàng	Checkout 5 bước → COD hoặc MoMo (Deep Link + IPN)
Đánh giá sản phẩm	Khách hàng	Viết review → Content Moderation → ABSA → Lưu
Quản lý đơn hàng	Admin	Xem/cập nhật trạng thái (pending→delivered/cancelled)
Kiểm duyệt đánh giá	Admin	Xem review flag + ABSA, Approve/Reject/Delete

### 3.3. Luồng màn hình Customer App

1. **Khởi động:** Splash → Login/Register → Home
2. **Mua sắm:** Home → Product List → Product Detail → Add to Cart / Buy Now
3. **Thanh toán:** Cart → Checkout → Address → Payment → Order Success
4. **Đơn hàng:** Profile → Order History → Order Detail
5. **Đánh giá:** Product Detail → Reviews → Write Review → ABSA Result

### 3.4. Quy trình đánh giá với ABSA

#### Luồng xử lý chính:

Input (1-5 sao, tiêu đề, nội dung, ảnh) → Content Moderation → Flagged → isFlagged=true → Admin duyệt

hoặc Pass → ABSA Service (Aspect Detection → Sentiment Classification → Lưu Review → Cập nhật Rating)

#### Hiển thị kết quả ABSA:

- Biểu đồ thanh ngang với màu sắc (xanh/đỏ/xám)
- Overall sentiment badge (Positive/Negative/Neutral/Mixed)

## Chương 4. THIẾT KẾ CƠ SỞ DỮ LIỆU

Hệ thống VeritaShop sử dụng MongoDB Atlas - hệ quản trị CSDL NoSQL với 7 collections chính liên kết qua ObjectId.

### 4.1. Collection Users

Lưu thông tin tài khoản người dùng và địa chỉ giao hàng.

Bảng 2. Cấu trúc Collection Users

Trường	Kiểu dữ liệu	Mô tả
name	String	Tên người dùng (bắt buộc)
email	String (unique)	Email đăng nhập
password	String	Mật khẩu mã hóa bcrypt
addresses[]	Array	Danh sách địa chỉ giao hàng
role	String	"customer" hoặc "admin"
pinHash	String	Mã PIN bảo mật (tùy chọn)
isActive	Boolean	Trạng thái tài khoản

### 4.2. Collection Products

Lưu thông tin sản phẩm điện thoại với thông số kỹ thuật chi tiết.

Bảng 3. Cấu trúc Collection Products

Trường	Kiểu dữ liệu	Mô tả
name, brand	String	Tên và hãng sản phẩm
price, originalPrice	Number	Giá bán và giá gốc (VNĐ)
images[]	Array	URLs hình ảnh (Cloudinary)
specs	Object	{ram, rom, chip, battery, screen, camera}
colors[]	Array	{name, code, image} cho mỗi màu
condition	String	"new", "likenew", "used"
stock, rating	Number	Tồn kho và điểm đánh giá

### 4.3. Collection Orders và Payments

**Orders** lưu thông tin đơn hàng với workflow trạng thái: pending → confirmed → processing → shipping → delivered (hoặc cancelled).

**Payments** lưu giao dịch MoMo với các trường: order (tham chiếu), method, amount, requestId, momoOrderId, transId, status (pending/success/failed), resultCode.

### 4.4. Collection Reviews

Tích hợp kết quả ABSA và Content Moderation.

Bảng 4. Cấu trúc Collection Reviews

Trường	Kiểu dữ liệu	Mô tả
user, product	ObjectId	Tham chiếu người dùng và sản phẩm
rating	Number (1-5)	Điểm đánh giá
text, images[]	String, Array	Nội dung và hình ảnh
sentimentAnalysis[]	Array	Kết quả ABSA: {aspect, sentiment, confidence}
overallSentiment	String	positive/negative/neutral/mixed
isFlagged	Boolean	Bị flag bởi Content Moderation
moderationStatus	String	pending/approved/rejected

### 4.5. Collection Carts và Coupons

**Carts:** Lưu giỏ hàng tạm với user, product, quantity, color.

**Coupons:** Mã giảm giá với code, discountType (percentage/fixed), discountValue, minOrderValue, usageLimit, validUntil, isActive.

## Chương 5. MÔ HÌNH, THỰC NGHIỆM VÀ ĐÁNH GIÁ

### 5.1. Mô hình

Nghiên cứu này triển khai và so sánh hai phương pháp học cho bài toán Aspect-Based Sentiment Analysis (ABSA): Multi-Task Learning (MTL) và Single-Task Learning (STL).

Lưu ý: Packaging và Shipping là khía cạnh quan trọng trong TMĐT nhưng thường thiếu trong bộ dữ liệu ABSA truyền thống.

#### 5.1.1. Single-Task Learning (STL)

**Kiến trúc mô hình AD (Aspect Detection):**

Input → Tokenizer → VisoBERT → [CLS] Tokens → Dense + ReLU + Dropout(0.3)  
→ Classifier Sigmoid → Output (11 Aspects)

**Kiến trúc mô hình SC (Sentiment Classification):**

Input → Tokenizer → VisoBERT → [CLS] Tokens → Dense + ReLU + Dropout(0.3)  
→ Classifier Softmax → Output (Positive/Negative/Neutral)

#### 5.1.2. Multi-Task Learning (MTL):

Trong phương pháp MTL, cả hai tác vụ Aspect Detection (AD) và Sentiment Classification (SC) được huấn luyện đồng thời với backbone chia sẻ. Vector biểu diễn [CLS] từ lớp cuối cùng của mô hình được sử dụng làm đặc trưng đầu vào cho hai nhánh phân loại riêng biệt

Input → Tokenizer → VisoBERT → [CLS] Tokens

→ Shared Dense + ReLU + Dropout(0.3)

→  $\begin{cases} \text{AD Head} & \rightarrow \text{Dense + ReLU + Dropout} \rightarrow \text{Sigmoid} \rightarrow \text{Output} \\ \text{SC Head} & \rightarrow \text{Dense + ReLU + Dropout} \rightarrow \text{Softmax} \rightarrow \text{Output} \end{cases}$

### 5.2. Tập dữ liệu

**Nguồn dữ liệu:** 14.912 bình luận điện thoại di động từ Shopee, Lazada, Tiki. 11 khía cạnh bao gồm Packaging và Shipping - thường thiếu trong bộ dữ liệu hiện có.

**Phân chia:** 80% train (11.930), 10% validation (1.491), 10% test (1.491). Kappa = 0.8208 (rất tốt).

## Các khía cạnh được phân tích

Mô hình nhận diện 11 khía cạnh ngành điện thoại:

Bảng 5. 11 Aspects trong mô hình ABSA

STT	Aspect	Mô tả
1	Battery	Thời lượng pin, tốc độ sạc
2	Camera	Chất lượng ảnh, video, selfie
3	Performance	Tốc độ, đa nhiệm, gaming, chip
4	Display	Độ sáng, màu sắc, tần số quét
5	Design	Ngoại hình, chất liệu, trọng lượng
6	Packaging	Hộp, phụ kiện, seal
7	Price	Giá cả, khuyến mãi
8	Shop_Service	Tư vấn, hỗ trợ
9	Shipping	Tốc độ giao hàng
10	General	Nhận xét tổng quan
11	Others	Các vấn đề khác

## Nhãn cảm xúc

Mỗi khía aspect (trừ Others) được gán nhãn: **Positive**, **Negative**, hoặc **Neutral**.

## 5.3. Đánh nhãn dữ liệu

**Quy trình:** 9 annotators sử dụng Label Studio. Phase 1 gồm có 1,000 mẫu xây dựng guideline. Phase 2 gồm 13,912 mẫu còn lại.

**Kích thước mẫu đánh giá:** Công thức Slovin với  $N = 13,912$ ,  $e = 2.5\%$ :

$$n = \frac{N}{1 + N \cdot e^2} = \frac{13,912}{1 + 13,912 \times 0.025^2} \approx 1,435$$

Chọn **1,546 mẫu** để đánh giá độ đồng thuận.

**Fleiss' Kappa:**

$$\kappa = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

Trong đó:

- $\bar{P}$  là tỷ lệ đồng thuận quan sát được (observed agreement)
- $\bar{P}_e$  là tỷ lệ đồng thuận kỳ vọng do ngẫu nhiên (expected agreement)

Với  $n$  annotator và  $N$  mẫu,  $k$  categories:

$$\bar{P} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n(n-1)} \sum_{j=1}^k n_{ij}(n_{ij} - 1) \quad (1)$$

$$\bar{P}_e = \sum_{j=1}^k p_j^2, \quad \text{với } p_j = \frac{1}{Nn} \sum_{i=1}^N n_{ij} \quad (2)$$

Bảng 6. Tổng hợp Fleiss' Kappa qua 2 giai đoạn

Phase	Mẫu	Annotators	Kappa TB	Kappa T. vị	Đánh giá
Phase 1	1,000	8	0.7544	0.8469	Tốt
Phase 2	1,546	9	<b>0.8208</b>	0.8382	Rất tốt

Bảng 7. Fleiss' Kappa theo khía cạnh (cải thiện từ Phase 1 → Phase 2)

Aspect	Phase 1 Kappa	Phase 2 Kappa
Battery	0.8747	0.8834
Camera	0.8861	<b>0.9166</b>
Performance	0.7220	0.7842
Display	0.8469	0.7915
Design	0.5794	0.7105
Packaging	0.8528	0.8703
Price	0.6610	0.8770
Shop_Service	0.8548	0.8301
Shipping	0.8646	0.8917
General	0.5277	0.6352
Others	0.6287	0.8382

Theo thang Landis & Koch: <0.20 (kém), 0.21-0.40 (yếu), 0.41-0.60 (trung bình), 0.61-0.80 (tốt), 0.81-1.00 (rất tốt). Kappa tăng 8.8% (0.7544 → 0.8208).

## 5.4. Tiền xử lý dữ liệu

Trước khi huấn luyện mô hình, dữ liệu thô được tiền xử lý qua các bước sau:

1. **Chuẩn hóa chính tả:** Loại bỏ và sửa các lỗi chính tả phổ biến trong bình luận tiếng Việt (ví dụ: "đẹp" → "đẹp", "tooott" → "tốt").
2. **Chuyển đổi từ tiếng Anh:** Chuyển các từ tiếng Anh thông dụng sang tiếng Việt tương đương (ví dụ: "good" → "tốt", "nice" → "đẹp", "bad" → "tệ", "shop" → "cửa hàng").

3. **Chuẩn hóa viết tắt:** Chuyển các từ viết tắt, teencode sang dạng chuẩn (ví dụ: "k" → "không", "đc" → "được", "nc" → "nói chuyện").
4. **Loại bỏ nhiễu:** Xóa các ký tự đặc biệt không cần thiết, URL, số điện thoại, nhưng giữ lại emoji quan trọng biểu thị cảm xúc.
5. **Làm sạch nhãn sai:** Rà soát và loại bỏ các mẫu có nhãn không nhất quán hoặc gán nhãn sai so với nội dung bình luận.

## 5.5. Kết quả và đánh giá

### 5.5.1. Metrics

Để đánh giá hiệu năng mô hình, nghiên cứu sử dụng các chỉ số sau:

- **Precision:** Tỷ lệ dự đoán đúng trên tổng số dự đoán positive.
- **Recall:** Tỷ lệ dự đoán đúng trên tổng số positive thực tế.
- **F1-Score:** Trung bình điều hòa của Precision và Recall.

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

### 5.5.2. Cài đặt thí nghiệm

- **Hardware:** GPU NVIDIA RTX 3070 (8GB VRAM)
- **Framework:** PyTorch 2.0, Transformers 4.30
- **Optimizer:** AdamW với learning rate  $2 \times 10^{-5}$ , weight decay = 0.01, epsilon =  $1.0 \times 10^{-8}$
- **Batch size:** 16 cho BERT models, 32 cho BiLSTM models
- **Epochs:** 12 với early stopping, patience = 5
- **Dropout:** 0.3
- **Max gradient norm:** 1.0
- **Focal Loss parameters:**  $\gamma = 2.0$ ,  $\alpha = \text{auto}$

### 5.5.3. Baseline Models

Nghiên cứu so sánh 8 mô hình chia thành 3 nhóm:

#### 1. BiLSTM-based Models (Traditional Deep Learning):

Bảng 8. Kết quả các mô hình BiLSTM

Mô hình	Kiến trúc	AD F1	SC F1
BILSTM-MTL	BiLSTM + Conv1D + MTL	84.09%	33.48%
BILSTM-MTL-NoCon	BiLSTM + MTL (không Conv1D)	82.85%	34.28%
BILSTM-STL	BiLSTM + Conv1D + STL	86.23%	36.87%
BILSTM-STL-NoCon	BiLSTM + STL	85.69%	39.83%

#### 2. PhoBERT-based Models:

Bảng 9. Kết quả các mô hình PhoBERT

Mô hình	Kiến trúc	AD F1	SC F1
PhoBERT-MTL	vinai/phobert-base + MTL	66.28%	92.93%
PhoBERT-STL	vinai/phobert-base + STL	88.84%	92.06%

#### 3. VisoBERT-based Models (Best Performance):

Bảng 10. Kết quả các mô hình VisoBERT

Mô hình	Kiến trúc	AD F1	SC F1
VisoBERT-MTL	5CD-AI/visobert + MTL	82.68%	93.63%
<b>VisoBERT-STL</b>	5CD-AI/visobert + STL	<b>89.39%</b>	<b>96.37%</b>

### 5.5.4. Kết Quả

Bảng dưới đây tổng hợp kết quả F1-Score của tất cả các mô hình đã thực nghiệm trên hai nhiệm vụ: Phát hiện khía cạnh (AD) và Phân loại cảm xúc (SC).

Bảng 11. Bảng so sánh F1-Score của các mô hình

Metric	STL		MTL	
Model	F1_ad	F1_sc	F1_ad	F1_sc
BiLSTM	85.69%	39.83%	82.85%	34.28%
BiLSTM + Conv1D	86.23%	36.87%	<b>84.09%</b>	33.48%
PhoBERT	88.84%	92.06%	66.28%	92.93%
VisoBERT	<b>89.39%</b>	<b>96.37%</b>	82.68%	<b>93.63%</b>

Từ kết quả trên, có thể thấy:

- Các mô hình BiLSTM đạt kết quả tốt ở nhiệm vụ Phát hiện khía cạnh (84-86%) nhưng kém ở Phân loại cảm xúc (33-40%).
- Các mô hình Transformer (PhoBERT, VisoBERT) đạt kết quả xuất sắc ở Phân loại cảm xúc (92-96%).
- **VisoBERT-STL** đạt kết quả cao nhất ở cả hai nhiệm vụ với AD F1 = 89.39% và SC F1 = 96.37%.

#### 5.5.5. Phân tích kết quả

##### 1. Học đơn nhiệm vụ tốt hơn học đa nhiệm vụ (STL > MTL):

- PhoBERT-STL: Phát hiện khía cạnh tăng +22.56% so với PhoBERT-MTL
- VisoBERT-STL: Phát hiện khía cạnh tăng +6.71%, Phân loại cảm xúc tăng +2.74% so với VisoBERT-MTL
- Huấn luyện riêng biệt giúp mỗi nhiệm vụ được tối ưu độc lập, tránh xung đột giữa các nhiệm vụ

##### 2. Mô hình Transformer vượt trội hơn BiLSTM:

- Phân loại cảm xúc: 96.37% (VisoBERT) so với 39.83% (BiLSTM) = cải thiện +56.54%
- Các mô hình Transformer được huấn luyện trước có khả năng biểu diễn ngữ cảnh mạnh hơn

##### 3. VisoBERT cho kết quả tốt hơn PhoBERT:

- VisoBERT-STL: Phát hiện khía cạnh 89.39%, Phân loại cảm xúc 96.37%

- PhoBERT-STL: Phát hiện khía cạnh 88.84%, Phân loại cảm xúc 92.06%
- VisoBERT được huấn luyện trên tập dữ liệu 14GB tiếng Việt và có khả năng hiểu được emoji/icon nên cho kết quả tốt hơn

## 5.6. Kết luận

### 5.6.1. Kết quả đạt được

1. **Xây dựng bộ dữ liệu:** Thu thập và đánh nhãn 14,912 bình luận từ các sàn thương mại điện tử Shopee, Lazada, Tiki với 11 khía cạnh và 3 mức cảm xúc. Độ đồng thuận Fleiss' Kappa đạt 0.8208 (mức rất tốt).
2. **Thực nghiệm đa mô hình:** So sánh 6 mô hình thuộc 3 nhóm kiến trúc (BiLSTM, PhoBERT, VisoBERT) với 2 chiến lược huấn luyện (MTL và STL).
3. **Kết quả tốt nhất:** Mô hình VisoBERT-STL đạt  $F1 = 89.39\%$  cho Phát hiện khía cạnh và  $F1 = 96.37\%$  cho Phân loại cảm xúc.
4. **Tích hợp hệ thống:** Triển khai mô hình vào hệ thống thương mại điện tử thực tế với API phân tích cảm xúc theo thời gian thực.

### 5.6.2. Hạn chế cần khắc phục

- Bộ dữ liệu tập trung vào lĩnh vực điện thoại di động, cần mở rộng sang các lĩnh vực khác.
- Chưa xử lý được các bình luận có nhiều khía cạnh với cảm xúc trái ngược phức tạp.
- Thời gian suy luận của mô hình Transformer còn chậm so với yêu cầu thời gian thực.

### 5.6.3. Hướng phát triển

- Mở rộng bộ dữ liệu sang các lĩnh vực khác như thời trang, mỹ phẩm, thực phẩm.
- Nghiên cứu kỹ thuật nén mô hình (Knowledge Distillation) để giảm thời gian suy luận.
- Tích hợp thêm phân tích hình ảnh sản phẩm để có đánh giá toàn diện hơn.
- Phát triển dashboard trực quan hóa xu hướng cảm xúc theo thời gian.

# TÀI LIỆU THAM KHẢO

## Tài liệu

- [1] Google Developers, "Flutter Documentation." [Online].  
Available: <https://flutter.dev/docs>
- [2] OpenJS Foundation, "Node.js Documentation." [Online].  
Available: <https://nodejs.org/en/docs/>
- [3] MongoDB, Inc., "MongoDB Manual." [Online].  
Available: <https://www.mongodb.com/docs/manual/>
- [4] MoMo for Business, "MoMo Payment API Integration." [Online].  
Available: <https://developers.momo.vn/>
- [5] Nguyen et al., "VisoBERT: A Pre-trained Language Model for Vietnamese Social Media." [Online].  
Available: <https://huggingface.co/uitnlp/visobert>
- [6] Nguyen D.Q., Nguyen A.T., "PhoBERT: Pre-trained language models for Vietnamese." EMNLP 2020.  
Available: <https://github.com/VinAIRResearch/PhoBERT>
- [7] Tran Q.-L., Le P.T.D., Do T.-H., "Aspect-based Sentiment Analysis for Vietnamese Reviews about Beauty Product on E-commerce Websites." *Proceedings of the 36th Pacific Asia Conference on Language, Information and Computation (PACLIC 36)*, 2022.  
Available: <https://aclanthology.org/2022.paclic-1.84.pdf>
- [8] OpenAI, "Moderation API." [Online].  
Available: <https://platform.openai.com/docs/guides/moderation>
- [9] OpenAI, "Text-to-Speech API." [Online].  
Available: <https://platform.openai.com/docs/guides/text-to-speech>
- [10] Phan L., Huynh P.P., Nguyen K.T.T., Huynh K.S., Nguyen T.T., Nguyen T.L., Huynh T.V., Nguyen K.V., "SA2SL: From Aspect-Based Sentiment Analysis to Social Listening System for Business Intelligence." *International Conference on Knowledge Science, Engineering and Management (KSEM)*, Springer, 2021.  
Available: [https://doi.org/10.1007/978-3-030-82147-0\\_53](https://doi.org/10.1007/978-3-030-82147-0_53)

## ĐÁNH GIÁ THÀNH VIÊN

Bảng 12. Bảng đánh giá công việc thực hiện và tỷ lệ đóng góp

STT	Mã SV	Họ và Tên	Nội dung công việc	Tỷ lệ (%)
1	2280603696	Nguyễn Quang Vinh	Thiết kế Database MongoDB, Cloudinary CDN, Review Models	20%
2	2280603283	Đặng Doanh Toại	State Management (Provider/ViewModels), Voice Search UI	20%
3	2280618597	Trần Đình Ty	Xây dựng Backend API (Auth, Product, Cart, Order), Coupon	20%
4	2280603036	Phan Thanh Thiên	Tích hợp thanh toán MoMo, Giao diện Customer App, Wishlist	20%
5	2280602828	Trần Tấn Tài	Tích hợp AI ABSA, API OpenAI, Speech to Text, Content Moderation	20%
<b>Tổng</b>				<b>100%</b>