**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TPHCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

------------------

A logo with a book and a hat

Description automatically generated

**BÁO CÁO ĐỒ ÁN**

**MÔN HỌC : KHAI PHÁ DỮ LIỆU**

**ĐỀ TÀI : ÁP DỤNG KỸ THUẬT GIẢM CHIỀU DỮ LIỆU T-SNE TRONG PHÂN TÍCH DỮ LIỆU KHÁCH HÀNG**

**Giảng viên hướng dẫn: Đinh Nguyễn Trọng Nghĩa**

**Sinh viên thực hiện (nhóm 4) :**

1. **Nguyễn Mai Linh – 2001222369**
2. **Nguyễn Bảo Trinh - 2001225583**
3. **Trần Nguyễn Như Huỳnh - 2001221771**

*TP HỒ CHÍ MINH, tháng 12 năm 2024*

**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TPHCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

------------------

A logo with a book and a hat

Description automatically generated

**BÁO CÁO ĐỒ ÁN**

**MÔN HỌC : KHAI PHÁ DỮ LIỆU**

**ĐỀ TÀI : ÁP DỤNG KỸ THUẬT GIẢM CHIỀU DỮ LIỆU T-SNE TRONG PHÂN TÍCH DỮ LIỆU KHÁCH HÀNG**

**Giảng viên hướng dẫn: Đinh Nguyễn Trọng Nghĩa**

**Sinh viên thực hiện:**

1. **Nguyễn Mai Linh – 2001222369**
2. **Nguyễn Bảo Trinh - 2001225583**
3. **Trần Nguyễn Như Huỳnh - 2001221771**

*TP HỒ CHÍ MINH, tháng 12 năm 2024*

**[MỤC LỤC]**

**I. Mở đầu**

1. **Giới thiệu đề tài:**

* Giới thiệu về bài toán giảm chiều dữ liệu và tầm quan trọng trong phân tích dữ liệu.*

* Tổng quan về kỹ thuật t-SNE.*

1. **Mục tiêu báo cáo :**

**II. Cơ sở lý thuyết**

1. **Nguyên lý hoạt động của t-SNE :**
2. **Công thức toán học mô tả t-SNE** *(giải thích ý nghĩa các thuật ngữ)* **:**
3. **Ưu điểm và hạn chế của t-SNE:**
4. **So sánh t-SNE với PCA:**

* Khác biệt trong mục tiêu giảm chiều.*

* Trường hợp áp dụng của từng kỹ thuật.*

**III. Quy trình thực hiện**

1. **Chuẩn bị dữ liệu:**
   1. **Nguồn gốc dữ liệu khách hàng:**
   2. **Xử lý dữ liệu** *(làm sạch,chuẩn hóa và lựa chọn đặc trưng)* **:**
2. **Áp dụng t-SNE:**
   1. **Giải thích các tham số chính** *(perplexity,learning rate)* **:**
   2. **Quy trình triển khai bằng Python:**
3. **Trực quan hóa và phân tích:**
   1. **Các bước vẽ biểu đồ:**
   2. **Ý nghĩa các nhóm dữ liệu biểu diễn:**
4. **Công cụ và thư viện lập trình sử dụng :** *(ví dụ: Python, thư viện scikit-learn, matplotlib...).*

**IV. Kết quả**

1. **Biểu diễn trực quan:**

* *Kết quả t-SNE, các biểu đồ,cụm dữ liệu*

1. **Phân tích kết quả:**

* *Đánh giá và giải thích xu hướng hoặc sự phân nhóm trong dữ liệu*

**V. Kết luận:**

1. **Tóm tắt nội dung:**

* *Những gì đạt được qua quá trình thực hiện*
* *Các nhận xét về hiệu quả của t-SNE trong phân tích dữ liệu.*

1. **Hạn chế và hướng phát triển:**

**VI. Tài liệu tham khảo :**