**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI**

**PHÂN HIỆU TẠI THÀNH PHỐ HỒ CHÍ MINH**



**BÁO CÁO THỰC TẬP CHUYÊN MÔN**

**ĐỀ TÀI:**

**QUẢN LÝ WEB BÁN HÀNG**

Giảng viên hướng dẫn: NGUYỄN LÊ MINH

Sinh viên thực hiện: LÊ THỊ TƯỜNG VI

Lớp: CQ.59.CNTT

Khoá: 59

Tp. Hồ Chí Minh, năm 2021

**BỘ GIÁO DỤC VÀ ĐÀO TẠO**

**TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI**

**PHÂN HIỆU TẠI THÀNH PHỐ HỒ CHÍ MINH**



**BÁO CÁO THỰC TẬP CHUYÊN MÔN**

**ĐỀ TÀI:**

**QUẢN LÝ WEB BÁN HÀNG**

Giảng viên hướng dẫn: NGUYỄN LÊ MINH

Sinh viên thực hiện:LÊ THỊ TƯỜNG VI

Lớp: CQ.59.CNTT

Khoá: 59

Tp. Hồ Chí Minh, năm 2021

**TRƯỜNG ĐẠI HỌC GIAO THÔNG VẬN TẢI CỘNG HÒA XÃ HỘI CHỦ NGHIÃ VIỆT NAM**

**PHÂN HIỆU TẠI THÀNH PHỐ HỒ CHÍ MINH Độc lập – Tự do – Hạnh phúc**

**NHIỆM VỤ THIẾT KẾ ĐỀ TÀI**

BỘ MÔN: **CÔNG NGHỆ THÔNG TIN**

-------\*\*\*-------

**Mã sinh viên:** 5951071119 **Họ tên SV:** Lê Thị Tường Vi

**Khóa:** 59 **Lớp:** CQ.59.CNTT

1. **Tên đề tài**

QUẢN LÝ WEB BÁN HÀNG

1. **Mục đích, yêu cầu**
   1. **Mục đích:**

* Hiểu được phương pháp hoạt động của mô hình ASP. MVC dựa trên Visual Studio 2019.
* Xây dựng “Phần mềm quản lý web bán hàng” với mục đích chính là ứng dụng được các thuật toán vào phần mềm nhằm giải quyết các nhu cầu của người dùng trong việc quản lí điều hành một trang web bán hàng trực tuyến. Hệ thống được xây dựng trên nền tảng Web...
  1. **Yêu cầu:**
* **Yêu cầu công nghệ**
  + Sử dụng ngôn ngữ lập trình C#, HTML, CSS, JAVASCRIPT, BOOTSTRAP.
  + Sử dụng công cụ Visual Studio 2019 và ASP.NET MVC.
* **Yêu cầu chức năng**
  + Phần mềm cho phép người dùng: mua hàng trực tuyến một cách nhanh chống, quản lý thông tin các nhà cung cấp, thông tin sản phẩm, thông tin nhóm sản phẩm, cũng như thông tin các phiếu nhập, hóa đơn một cách thuận tiện nhất.
* **Yêu cầu phi chức năng**
  + Tốc độ: Tối ưu về tốc độ, làm việc với dữ liệu vừa phải, tốc độ hiển thị nhanh.
  + Giao diện: Thân thiện với người dùng và dễ dàng thao tác.
* **Yêu cầu về tổ chức code**
  + Tổ chức code ngay ngắn dễ đọc, lược bỏ các thành phần rườm rà.

1. **Nội dung và phạm vi đề tài**
   1. **Nội dung:**
      * Tổng quan về các công nghệ đang sử dụng.
      * Tổng quan bài toán.
      * Phân tích và thiết kế phần mềm.
      * Lập trình xây dựng “Phần mềm quản lý web bán hàng”.
      * Kiểm thử và chạy thực nghiệm.
      * Kết quả thu được
   2. **Phạm vi:**
      * Nghiên cứu ASP.NET MVC.
      * Nghiên cứu sử dụng công cụ Visual Studio và ngôn ngữ C# Web.
2. **Công nghệ, công cụ và ngôn ngữ lập trình**
   * + Công nghệ sử dụng: ASP.NET MVC.
     + Công cụ Visual Studio 2019.
     + Ngôn ngữ lập trình: C#, HTML, CSS, JAVASCRIPT, BOOTSTRAPT.
3. **Các kết quả chính dự kiến sẽ đạt được và ứng dụng**
   * + Quyển báo cáo đề tài thực tập.
     + Xây dựng được phần mềm quản lý web bán hàng với đầy đủ chức năng đã nêu.
     + Hiểu được nguyên lý hoạt động của mô hình MVC.
4. **Giáo viên và cán bộ hướng dẫn**

Họ tên: Nguyễn Lê Minh

Đơn vị công tác: Phân hiệu Trường Đại học Giao thông Vận tải.

Điện thoại: Email:

**Ngày ...... tháng ...... năm 2021**

**Giảng viên hướng dẫn**

**LỜI CẢM ƠN**

Để hoàn thành đề tài thực tậpnchuyeen môn này trước hết em xin gửi đến quý thầy, cô **Bộ môn Công nghệ thông tin – Phân hiệu Trường Đại học Giao thông Vận tải tại Thành phố Hồ Chí Minh** lời cảm ơn chân thành vì đã truyền đạt cho em những kiến thức không chỉ từ sách vở, mà còn những kinh nghiệm quý giá từ cuộc sống trong khoảng thời gian học tập tại trường. Đặc biệt em xin gửi đến thầy Nguyễn Lê Minh lời cảm ơn sâu sắc nhất, thầy đã tận tình hướng dẫn, chỉ bảo em trong suốt quá trình thực hiện đề tài. Tuy đề tài không được lớn lao nhưng nếu không được sự hướng dẫn chỉ bảo tận tình của thầy và cô thì đề tài thực tập chuyên môn này khó có thể hoàn thành được.

Vì thời gian làm đồ án thực tập chuyên môn có hạn cũng như hiểu biết cá nhân còn hạn chế, em cũng đã nỗ lực hết sức để hoàn thành bài báo cáo đồ án một cách tốt nhất, nhưng chắc chắn vẫn sẽ có những thiếu sót không thể tránh khỏi. Em kính mong nhận được sự thông cảm và những ý kiến đóng góp chân thành từ quý thầy cô.

Sau cùng, em xin kính chúc Quý Thầy Cô trong **Bộ môn Công nghệ thông tin** đang công tác tại Bộ phận một cửa – Phòng Tổ chức hành chính luôn mạnh khoẻ, hạnh phúc và thành công hơn nữa trong công việc cũng như trong cuộc sống.

Em xin chân thành cảm ơn!

Tp. Hồ Chí Minh, ngày ….… tháng ….… năm 2021

Sinh viên thực hiện

**Lê Thị Tường Vi**

# **NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN**

|  |
| --- |
| ***Tp. Hồ Chí Minh, ngày ……. tháng ….… năm ….…***  **Giảng viên hướng dẫn** |

**MỤC LỤC**

**[NHIỆM VỤ THIẾT KẾ TỐT NGHIỆP](#_Toc11335052)** [I](#_Toc11335052)

**[LỜI CẢM ƠN](#_Toc11335053)** [IV](#_Toc11335053)

**[NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN](#_Toc11335054)** [V](#_Toc11335054)

**[DANH MỤC THUẬT NGỮ](#_Toc11335055)** [VIII](#_Toc11335055)

**[DANH MỤC BẢNG BIỂU](#_Toc11335056)** [X](#_Toc11335056)

**[DANH MỤC HÌNH ẢNH](#_Toc11335057)** [XI](#_Toc11335057)

**[CHƯƠNG 1: MỞ ĐẦU](#_Toc11335058)** [1](#_Toc11335058)

[1.1. Lý do chọn đề tài 1](#_Toc11335059)

[1.2. Mục đích nghiên cứu 2](#_Toc11335060)

[1.3. Đối tượng và phạm vi nghiên cứu 3](#_Toc11335061)

[1.4. Phương pháp nghiên cứu 3](#_Toc11335062)

[1.5. Cấu trúc báo cáo đồ án tốt nghiệp 3](#_Toc11335063)

[1.6. Kết luận chương 1 3](#_Toc11335064)

**[CHƯƠNG 2: CƠ SỞ LÝ THUYẾT](#_Toc11335065)** [4](#_Toc11335065)

[2.1. Tổng quan về Machine Learning 4](#_Toc11335066)

[2.1.1. Giới thiệu về Machine Learning 4](#_Toc11335067)

[2.1.2. Định nghĩa về Machine Learning 5](#_Toc11335068)

[2.1.3. Ứng dụng của Machine Learning 5](#_Toc11335069)

[2.2. Phân loại thuật toán Machine Learning 6](#_Toc11335070)

[2.2.1. Học có giám sát 6](#_Toc11335071)

[2.2.2. Học không giám sát 7](#_Toc11335072)

[2.2.3. Học bán giám sát 8](#_Toc11335073)

[2.2.4. Học củng cố 8](#_Toc11335074)

[2.3. Tổng quan về phân cụm dữ liệu 9](#_Toc11335075)

[2.3.1. Phân cụm dữ liệu là gì? 9](#_Toc11335076)

[2.3.2. Mục đích của phân cụm dữ liệu 9](#_Toc11335077)

[2.3.3. Quá trình phân cụm dữ liệu 10](#_Toc11335078)

[2.3.4. Những kỹ thuật tiếp cận trong phân cụm dữ liệu 15](#_Toc11335079)

[2.3.5. Phương pháp phân cụm phân hoạch 16](#_Toc11335080)

[2.3.6. Ứng dụng phân cụm dữ liệu 17](#_Toc11335081)

[2.4. Kết luận chương 2 18](#_Toc11335082)

**[CHƯƠNG 3: PHÂN TÍCH THUẬT TOÁN K-MEANS VÀ VẤN ĐỀ PHÂN CỤM DỮ LIỆU VĂN BẢN](#_Toc11335083)** [19](#_Toc11335083)

[3.1. Đặt vấn đề 19](#_Toc11335084)

[3.2. Thuật toán K-means 20](#_Toc11335085)

[3.2.1. Sơ lược về thuật toán K-means 20](#_Toc11335086)

[3.2.2. Ý tưởng của thuật toán K-means 20](#_Toc11335087)

[3.2.3. Đánh giá thuật toán K-means 27](#_Toc11335088)

[3.2.4. Những lưu ý khi sử dụng thuật toán K-means 28](#_Toc11335089)

[3.3. Vấn đề phân cụm dữ liệu văn bản 28](#_Toc11335090)

[3.3.1. Dữ liệu văn bản 28](#_Toc11335091)

[3.3.2. Một số vấn đề trong xử lý dữ liệu văn bản 29](#_Toc11335092)

[3.3.3. Mô hình biểu diễn dữ liệu văn bản 30](#_Toc11335093)

[3.3.4. Quy trình thực hiện phân cụm văn bản 31](#_Toc11335094)

[3.4. Kết luận chương 3 41](#_Toc11335095)

**[CHƯƠNG 4: THỬ NGHIỆM VÀ ĐÁNH GIÁ](#_Toc11335096)** [42](#_Toc11335096)

[4.1. Giới thiệu bài toán 42](#_Toc11335097)

[4.2. Công nghệ và công cụ thực hiện 42](#_Toc11335098)

[4.2.1. Tổng quan về .NET Framework 42](#_Toc11335099)

[4.2.2. Tổng quan về ngôn ngữ C# 44](#_Toc11335100)

[4.3. Cơ sở dữ liệu 44](#_Toc11335101)

[4.4. Giao diện của chương trình 46](#_Toc11335102)

[4.5. Kết quả và đánh giá 56](#_Toc11335103)

[4.6. Kết luận chương 4 57](#_Toc11335104)

**[KẾT LUẬN VÀ KIẾN NGHỊ](#_Toc11335105)** [58](#_Toc11335105)

**[PHỤ LỤC](#_Toc11335106)** [59](#_Toc11335106)

**[TÀI LIỆU THAM KHẢO](#_Toc11335107)** [60](#_Toc11335107)

# **DANH MỤC THUẬT NGỮ**

|  |  |  |  |
| --- | --- | --- | --- |
| Từ viết tắt | Mô tả | Ý nghĩa | Ghi chú |
| CNTT |  | Công nghệ thông tin | Thuật ngữ |
| CLI | Common Language Infrastructure | Hạ tầng ngôn ngữ chung |  |
| ECMA | European Computer Manufacturer's Association |  |  |
| CLR | Common Language Runtime | Khối thực thi ngôn ngữ chung |  |
| FCL | Framework Class Library | Thư viện lớp khung |  |
| ML | Machine Learning | Học máy |  |
| AI | Artificial Intelligence | Trí tuệ nhân tạo |  |
| DL | Deep Learning | Học sâu |  |
| DM | Data Mining | Khai phá dữ liệu |  |
| SL | Supervised Learning | Học có giám sát |  |
|  | Classification | Phân loại |  |
|  | Regression | Hồi quy |  |
| UL | Unsupervised Learning | Học không giám sát |  |
|  | Clustering | Phân nhóm |  |
|  | Association |  |  |
| SSL | Semi-Supervised Learning | Học bán giám sát |  |
| RL | Reinforcement Learning | Học củng cố |  |
| PM | Partitioning Methods | Phương pháp phân cụm phân hoạch | Phương pháp |
| CSDL | Database | Cơ sở dữ liệu |  |
| CSV | Comma-Separated Values |  | File |
| SaaS | Software-as-a-Service |  | Mô hình |
| EM | Elbow method | Phương pháp khuỷu tay | Phương pháp |
| VSM | Vector Space Model | Mô hình không gian vector | Mô hình |
| TF-IDF | Term Frequency – Inverse Document Frequency | Độ quan trọng của từ trong văn bản |  |
| WSS | Within-cluster Sum of Square | Tổng biến thiên bình phương khoảng cách trong cụm |  |
| DC | Data Clustering | Phân cụm dữ liệu |  |
| CS | Cosine Similarity | Độ tương tự Cosine | Độ đo |

# **DANH MỤC BẢNG BIỂU**

[Bảng 3. 1 - Bảng biểu thị dữ liệu cần phân cụm 22](#_Toc10820855)

[Bảng 3. 2 - Kết quả sau khi phân cụm bằng K-means 27](#_Toc10820856)

[Bảng 3. 3 - Thống kê các từ có tần suất xuất hiện cao(B. Croft, UMass) 30](#_Toc10820857)

[Bảng 3. 4 - Độ tương tự Cosine giữa 2 tài liệu d1 và d2 38](#_Toc10820858)

# **DANH MỤC HÌNH ẢNH**

[Hình 2. 1 - Giới thiệu về ML[2] 4](#_Toc28895559)

[Hình 2. 2 - Các thuật toán ML[3] 6](#_Toc28895560)

[Hình 2. 3 - Ví dụ về DC[8] 9](#_Toc28895561)

[Hình 2. 4 - Quá trình phân cụm dữ liệu[8] 10](#_Toc28895562)

[Hình 2. 5 - Tiêu chuẩn phân cụm[7] 13](#_Toc28895563)

[Hình 2. 6 - Bảng độ đo đánh giá ngoại[7] 14](#_Toc28895564)

[Hình 2. 7 - Bảng độ đo đánh giá nội[7] 15](#_Toc28895565)

[Hình 2. 8 - Gợi ý từ khoá tìm kiếm có liên quan của Google 18](#_Toc28895566)

[Hình 3. 1 - Mô tả phân bố các cụm[15] 19](#_Toc28895567)

[Hình 3. 2 - Sơ đồ khối thuật toán K-means 21](#_Toc28895568)

[Hình 3. 3 - Khởi tạo tâm cho 2 nhóm 22](#_Toc28895569)

[Hình 3. 4 - Gom các đối tượng vào nhóm gần nhất lần 1 23](#_Toc28895570)

[Hình 3. 5 - Tính lại toạ độ của tâm nhóm lần 1 24](#_Toc28895571)

[Hình 3. 6 - Gom các đối tượng vào nhóm gần nhất lần 2 25](#_Toc28895572)

[Hình 3. 7 - Tính lại toạ độ của tâm nhóm lần 2 26](#_Toc28895573)

[Hình 3. 8 - Quy trình thực hiện phân cụm 31](#_Toc28895574)

[Hình 3. 9 - Các độ đo tương tự thường dùng 34](#_Toc28895575)

[Hình 3. 10 - Thuật toán K-means trong phân cụm tài liệu tin tức 36](#_Toc28895576)

[Hình 3. 11 - Mức tăng về số lượt tìm kiếm trên Google(statisticbrain.com) 37](#_Toc28895577)

[Hình 3. 12 - Sơ đồ khối quy trình hệ thống 39](#_Toc28895578)

[Hình 3. 13 - Đồ thị đường khuỷu tay(EM) 41](#_Toc28895579)

[Hình 4. 1 - Kiến trúc .NET Framework 43](#_Toc11334495)

[Hình 4. 2 - File Dataset 45](#_Toc11334496)

[Hình 4. 3 - Giao diện chính của chương trình khi khởi động 46](#_Toc11334497)

[Hình 4. 4 - Giao diện của chức năng phân cụm văn bản 47](#_Toc11334498)

[Hình 4. 5 - Giao diện sau khi đọc file CSV và lựa chọn số cụm 47](#_Toc11334499)

[Hình 4. 6 - Giao diện khi hoàn thành phân cụm 48](#_Toc11334500)

[Hình 4. 7 - Thời gian khi phân cụm với 500 tài liệu 48](#_Toc11334501)

[Hình 4. 8 - Kết quả sau khi phân cụm tin tức 49](#_Toc11334502)

[Hình 4. 9 - Danh sách các từ đặc trưng của mỗi cụm 50](#_Toc11334503)

[Hình 4. 10 - Tập dữ liệu data\_textclustering.csv 51](#_Toc11334504)

[Hình 4. 11 - Kết quả phân cụm lần 1 52](#_Toc11334505)

[Hình 4. 12 - Kết quả phân cụm lần thứ 10 52](#_Toc11334506)

[Hình 4. 13 - Kết quả phân cụm lần 1 (30%) 53](#_Toc11334507)

[Hình 4. 14 - Kết quả phân cụm lần thứ 7 (30%) 53](#_Toc11334508)

[Hình 4. 15 - Kết quả phân cụm lần thứ 10 (30%) 54](#_Toc11334509)

[Hình 4. 16 - Kết quả khi phân cụm 10 lần với 510 tài liệu 55](#_Toc11334510)

[Hình 4. 17 - Giao diện phương pháp khuỷu tay 56](#_Toc11334511)

# **CHƯƠNG 1: MỞ ĐẦU**

* 1. **Lý do chọn đề tài**

## Trong những năm gần đây sự phát triển nhanh chóng của xã hội là do có một phần đóng góp không nhỏ trong sự phát triển của ngành CNTT. Đặc biệt bối cảnh thế giới đang hướng đến cuộc cách mạng công nghiệp lần thứ 4 với ngành CNTT đã gặt hái được những thành tựu ban đầu liên quan đến lĩnh vực trí tuệ nhân tạo, Machine Learning, Deep Learning ,Big Data…thuật ngữ Electronic Business (Thương mại điện tử) ra đời và ngày càng trở lên phổ biến.

## Trong bất cứ lĩnh vực nào, thương mại điện tử cũng có lợi thế. Nó không những làm giảm thời gian mà còn làm giảm chi phí cho mọi vấn đề giao dịch của con người. Việc kinh doanh qua mạng không chỉ đáp ứng nhu cầu của con người mà còn giúp cho người sử dụng so sánh giá ở nhiều nơi để mua những mặt hàng tốt, giá hợp lý. Với những kiến thức đã học và thời gian thực tế, em xin lựa chọn đề tài : “Xây dựng trang web bán hàng trực tuyến sử dụng ngôn ngữ lập trình ASP.NET MVC và cơ sở dữ liệu SQL Server”.

## **Mục đích nghiên cứu**

* Nghiên cứu các vấn đề cơ bản của thương mại điện tử.
* Áp dụng đề tài giúp việc kinh doanh qua mạng không chỉ đáp ứng nhu cầu của con người mà còn giúp cho người sử dụng so sánh giá ở nhiều nơi để mua những mặt hàng tốt, giá hợp lý.

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

* Tìm hiểu về một phần nhỏ của lĩnh vực thương mại điện tử và ASP.NET MVC, cách thức hoạt động của nó.
* Cài đặt và thử nghiệm với trang web bán hàng.

## **Cấu trúc báo cáo đồ án thực tập chuyên môn**

Cấu trúc đồ án được chia thành các chương như sau:

1. Chương 1: Mở đầu – Giới thiệu tổng quan về đề tài đồ án
2. Chương 2: Cơ sở lý thuyết – Trình bày cơ sở lý thuyết sẽ sử dụng trong đề tài đồ án
3. Chương 3: Phân tích – Trong chương này sẽ tiến hành phân tích thuật toán để ra kết quả và nhật xét về thuật toán và vấn đề phân cụm văn bản
4. Chương 4: Thử nghiệm và đánh giá, mô tả giao diện – cài đặt thuật toán và đưa ra đề xuất số cụm tối ưu với bài toán phân cụm văn bản.
5. Kết luận, kiến nghị và hướng phát triển - Đưa ra kết quả đạt được, những thứ còn thiếu sót và hướng phát triển về thuật toán cho ứng dụng
6. Tài liệu tham khảo

# **CHƯƠNG 2: CƠ SỞ LÝ THUYẾT**

## **Tổng quan về Thương mại điện tử**

### **Giới thiệu về Thương mại điện tử**

**Thương mại điện tử**, hay còn gọi là **e-commerce**, **e-comm** hay **EC**, là sự mua bán sản phẩm hay dịch vụ trên các hệ thống điện tử như [Internet](https://vi.wikipedia.org/wiki/Internet" \o ") và các [mạng máy tính](https://vi.wikipedia.org/wiki/M%E1%BA%A1ng_m%C3%A1y_t%C3%ADnh" \o "Mạng máy tính). Thương mại điện tử dựa trên một số công nghệ như [chuyển tiền điện tử](https://vi.wikipedia.org/wiki/Chuy%E1%BB%83n_ti%E1%BB%81n_%C4%91i%E1%BB%87n_t%E1%BB%AD" \o "Chuyển tiền điện tử), [quản lý chuỗi dây chuyền cung ứng](https://vi.wikipedia.org/w/index.php?title=Qu%E1%BA%A3n_l%C3%BD_chu%E1%BB%97i_d%C3%A2y_chuy%E1%BB%81n_cung_%E1%BB%A9ng&action=edit&redlink=1" \o "Quản lý chuỗi dây chuyền cung ứng (trang chưa được viết)), [tiếp thị Internet](https://vi.wikipedia.org/wiki/E-marketing" \o "Kinh doanh điện tử), [quá trình giao dịch trực tuyến](https://vi.wikipedia.org/w/index.php?title=Qu%C3%A1_tr%C3%ACnh_giao_d%E1%BB%8Bch_tr%E1%BB%B1c_tuy%E1%BA%BFn&action=edit&redlink=1" \o "Quá trình giao dịch trực tuyến (trang chưa được viết)), [trao đổi dữ liệu điện tử](https://vi.wikipedia.org/wiki/Trao_%C4%91%E1%BB%95i_d%E1%BB%AF_li%E1%BB%87u_%C4%91i%E1%BB%87n_t%E1%BB%AD" \o "Trao đổi dữ liệu điện tử) ([EDI](https://vi.wikipedia.org/w/index.php?title=EDI&action=edit&redlink=1" \o "EDI (trang chưa được viết))), các [hệ thống quản lý hàng tồn kho](https://vi.wikipedia.org/w/index.php?title=H%E1%BB%87_th%E1%BB%91ng_qu%E1%BA%A3n_l%C3%BD_h%C3%A0ng_t%E1%BB%93n_kho&action=edit&redlink=1" \o "Hệ thống quản lý hàng tồn kho (trang chưa được viết)), và các hệ thống tự động thu thập dữ liệu. Thương mại điện tử hiện đại thường sử dụng mạng [World Wide Web](https://vi.wikipedia.org/wiki/World_Wide_Web" \o "World Wide Web) là một điểm ít nhất phải có trong chu trình giao dịch, mặc dù nó có thể bao gồm một phạm vi lớn hơn về mặt công nghệ như [email](https://vi.wikipedia.org/wiki/Th%C6%B0_%C4%91i%E1%BB%87n_t%E1%BB%AD" \o "Thư điện tử), các thiết bị [di động](https://vi.wikipedia.org/wiki/%C4%90i%E1%BB%87n_tho%E1%BA%A1i_di_%C4%91%E1%BB%99ng" \o "Điện thoại di động) như là [điện thoại](https://vi.wikipedia.org/wiki/%C4%90i%E1%BB%87n_tho%E1%BA%A1i" \o ")

Thuật ngữ ICT (viết tắt của từ tiếng Anh information commercial technology) cũng có nghĩa là thương mại điện tử, nhưng ICT được hiểu theo khía cạnh công việc của các chuyên viên công nghệ. Theo nghĩa hẹp, thương mại điện tử chỉ đơn thuần bó hẹp thương mại điện tử trong việc mua bán hàng hóa và dịch vụ thông qua các phương tiện điện tử, nhất là qua Internet và các mạng liên thông khác. Theo Tổ chức Thương mại thế giới (WTO), "Thương mại điện tử bao gồm việc sản xuất, quảng cáo, bán hàng và phân phối sản phẩm được mua bán và thanh toán trên mạng Internet, nhưng được giao nhận một cách hữu hình, cả các sản phẩm giao nhận cũng như những thông tin số hoá thông qua mạng Internet". Theo Uỷ ban Thương mại điện tử của Tổ chức hợp tác kinh tế châu Á-Thái Bình Dương (APEC), "Thương mại điện tử là công việc kinh doanh được tiến hành thông qua truyền thông số liệu và công nghệ tin học kỹ thuật số".

Thương mại điện tử hiểu theo nghĩa rộng là các giao dịch tài chính và thương mại bằng phương tiện điện tử như: trao đổi dữ liệu điện tử, chuyển tiền điện tử và các hoạt động như gửi/rút tiền bằng thẻ tín dụng.

Các điểm đặc biệt của thương mại điện tử so với các kênh phân phối truyền thống là tính linh hoạt cao độ về mặt cung ứng và giảm thiểu lớn phí tổn vận tải với các đối tác kinh doanh. Các phí tổn khác thí dụ như phí tổn điện thoại và đi lại để thu nhập khác hàng hay phí tổn trình bày giới thiệu cũng được giảm xuống. Mặc dầu vậy, tại các dịch vụ vật chất cụ thể, khoảng cách không gian vẫn còn phải được khắc phục và vì thế đòi hỏi một khả năng tiếp vận phù hợp nhất định.



### **Định nghĩa về Thương mại điện tử**

### Khi nói về khái niệm thương mại điện tử (E-Commerce), nhiều người nhầm lẫn với khái niệm của [Kinh doanh điện tử](https://vi.wikipedia.org/wiki/Kinh_doanh_%C4%91i%E1%BB%87n_t%E1%BB%AD" \o "Kinh doanh điện tử) (E-Business). Tuy nhiên, thương mại điện tử đôi khi được xem là [tập con](https://vi.wikipedia.org/wiki/T%E1%BA%ADp_h%E1%BB%A3p_con" \o "Tập hợp con) của [kinh doanh điện tử](https://vi.wikipedia.org/wiki/Kinh_doanh_%C4%91i%E1%BB%87n_t%E1%BB%AD" \o "Kinh doanh điện tử). Thương mại điện tử chú trọng đến việc mua bán trực tuyến (tập trung bên ngoài), trong khi đó [kinh doanh điện tử](https://vi.wikipedia.org/wiki/Kinh_doanh_%C4%91i%E1%BB%87n_t%E1%BB%AD" \o "Kinh doanh điện tử) là việc sử dụng [Internet](https://vi.wikipedia.org/wiki/Internet" \o "Internet) và các công nghệ trực tuyến tạo ra quá trình hoạt động kinh doanh hiệu quả dù có hay không có lợi nhuận, vì vậy tăng lợi ích với [khách hàng](https://vi.wikipedia.org/wiki/Kh%C3%A1ch_h%C3%A0ng" \o "Khách hàng) (tập trung bên trong).

## **Công cụ Visual Studio 2019**

* **Microsoft Visual Studio** là một [môi trường phát triển tích hợp](https://vi.wikipedia.org/wiki/M%C3%B4i_tr%C6%B0%E1%BB%9Dng_ph%C3%A1t_tri%E1%BB%83n_t%C3%ADch_h%E1%BB%A3p" \o "Môi trường phát triển tích hợp) (IDE) từ [Microsoft](https://vi.wikipedia.org/wiki/Microsoft" \o "Microsoft). Nó được sử dụng để phát triển [chương trình máy tính](https://vi.wikipedia.org/wiki/Ch%C6%B0%C6%A1ng_tr%C3%ACnh_m%C3%A1y_t%C3%ADnh" \o "Chương trình máy tính) cho Microsoft Windows, cũng như các trang web, các ứng dụng web và các dịch vụ web. Visual Studio sử dụng nền tảng phát triển phần mềm của Microsoft như [Windows API](https://vi.wikipedia.org/wiki/Windows_API" \o "Windows API), [Windows Forms](https://vi.wikipedia.org/wiki/Windows_Forms" \o "Windows Forms), [Windows Presentation Foundation](https://vi.wikipedia.org/wiki/Windows_Presentation_Foundation" \o "Windows Presentation Foundation), [Windows Store](https://vi.wikipedia.org/wiki/Windows_Store" \o "Windows Store) và [Microsoft Silverlight](https://vi.wikipedia.org/wiki/Microsoft_Silverlight" \o "). Nó có thể sản xuất cả hai [ngôn ngữ máy](https://vi.wikipedia.org/wiki/Ng%C3%B4n_ng%E1%BB%AF_m%C3%A1y" \o "Ngôn ngữ máy) và [mã số quản lý](https://vi.wikipedia.org/w/index.php?title=M%C3%A3_s%E1%BB%91_qu%E1%BA%A3n_l%C3%BD&action=edit&redlink=1" \o "Mã số quản lý (trang chưa được viết)).
* **Visual Studio** hỗ trợ nhiều [ngôn ngữ lập trình](https://vi.wikipedia.org/wiki/Ng%C3%B4n_ng%E1%BB%AF_l%E1%BA%ADp_tr%C3%ACnh" \o "Cải tiến mã nguồn) khác nhau và cho phép trình biên tập mã và gỡ lỗi để hỗ trợ (mức độ khác nhau) hầu như mọi ngôn ngữ lập trình. Các ngôn ngữ tích hợp gồm có [C](https://vi.wikipedia.org/wiki/C_(ng%C3%B4n_ng%E1%BB%AF_l%E1%BA%ADp_tr%C3%ACnh)" \o "C (ngôn ngữ lập trình)), [C++](https://vi.wikipedia.org/wiki/C++" \o ") và [C++/CLI](https://vi.wikipedia.org/w/index.php?title=C++/CLI&action=edit&redlink=1" \o "C++/CLI (trang chưa được viết)) (thông qua [Visual C++](https://vi.wikipedia.org/wiki/Visual_C++" \o "Visual C++)), [VB.NET](https://vi.wikipedia.org/wiki/VB.NET" \o "VB.NET) (thông qua [Visual Basic.NET](https://vi.wikipedia.org/wiki/Visual_Basic.NET" \o "Visual Basic.NET)), [C#](https://vi.wikipedia.org/wiki/C_th%C4%83ng" \o "C thăng) (thông qua [Visual C#](https://vi.wikipedia.org/w/index.php?title=Visual_C&action=edit&redlink=1" \o "Visual C (trang chưa được viết))) và [F#](https://vi.wikipedia.org/w/index.php?title=F_th%C4%83ng&action=edit&redlink=1" \o "F thăng (trang chưa được viết)) (như của [Visual Studio 2010](https://vi.wikipedia.org/w/index.php?title=Visual_Studio_2010&action=edit&redlink=1" \o "Visual Studio 2010 (trang chưa được viết))). Hỗ trợ cho các ngôn ngữ khác như [J++](https://vi.wikipedia.org/w/index.php?title=J++&action=edit&redlink=1" \o "J++ (trang chưa được viết))/[J#](https://vi.wikipedia.org/w/index.php?title=J_th%C4%83ng&action=edit&redlink=1" \o "J thăng (trang chưa được viết)), [Python](https://vi.wikipedia.org/wiki/Python" \o "Python) và [Ruby](https://vi.wikipedia.org/wiki/Ruby" \o "Ruby) thông qua dịch vụ cài đặt riêng rẽ. Nó cũng hỗ trợ [XML](https://vi.wikipedia.org/wiki/XML" \o "XML)/[XSLT](https://vi.wikipedia.org/wiki/XSLT" \o "Quy trình phát triển phần mềm), [HTML](https://vi.wikipedia.org/wiki/HTML" \o "HTML)/[XHTML](https://vi.wikipedia.org/wiki/XHTML" \o "XHTML), [JavaScript](https://vi.wikipedia.org/wiki/JavaScript" \o ") và [CSS](https://vi.wikipedia.org/wiki/CSS" \o "CSS).

## **Ngôn ngữ ASP.NET MVC**

* **MVC** là từ viết tắt của Model-View-Controller. Đây là một khái niệm lập trình ban đầu được giới thiệu vào những năm 1970 bởi Trygve Reenskaug sử dụng Smalltalk-76. Nó cũng rõ ràng trong một số khung JavaScript như AngularJS, Aurelia, Ember và Meteor.
* **ASP.NET** là sự kế thừa cho công nghệ Classic ASP (xem trước) do Microsoft phát triển và được chia thành hai loại phát triển web: Web Forms (năm 2002) và MVC (năm 2008). Nếu bạn muốn tạo các trang web ASP.NET, bạn nên sử dụng [Visual Studio](https://www.visualstudio.com/" \o "Truy cập Visual Studio trên Microsoft.com" \t "https://helpex.vn/article/_blank) của Microsoft để tối ưu hóa việc phát triển trang web của mình.
* **ASP.NET MVC** là tầm nhìn của Microsoft về khái niệm Model-View-Controller (MVC) được tích hợp vào công nghệ ASP.NET của họ để xây dựng các trang web hiện đại.

## **Hệ quản trị cơ sở dữ liệu SQL Server**

* SQL (Structured Query Language) là công cụ sử dụng để tổ chức, quản lý và truy xuất dữ liệu được lưu trữ trong các cơ sở dữ liệu.
* SQL là một hệ thống ngôn ngữ bao gồm tập các câu lệnh sử dụng để tƣơng tác với cơ sở dữ liệu quan hệ.
* SQL được sử dụng để điều khiển tất cả các chức năng mà một hệ quản trị cơ sở dữ liệu thƣơng mại cung cấp cho ngƣời dùng bao gồm: 
* Định nghĩa dữ liệu: SQL cung cấp khả năng định nghĩa các cơ sở dữ liệu, các cấu trúc lưu trữ và tổ chức dữ liệu cũng như mối quan hệ giữa các thành phần dữ liệu. 
* Truy xuất và thao tác dữ liệu: Với SQL ngƣời dùng có thể dễ dàng thực hiện các thao tác truy xuất, bổ sung, cập nhật và loại bỏ dữ liệu trong các cơ sở dữ liệu. 
* Điều khiển truy cập: SQL có thể đƣợc sử dụng để cấp phát và kiểm soát các thao tác của người sử dụng trên dữ liệu, đảm bảo sự an toàn cho cơ sở dữ liệu.
* Đảm bảo toàn vẹn dữ liệu: SQL định nghĩa các ràng buộc toàn vẹn trong cơ sở dữ liệu nhờ đó đảm bảo tính hợp lệ và chính xác của dữ liệu trước các thao tác cập nhật cũng như các lỗi của hệ thống.
* Như vậy, có thể nói rằng SQL là một ngôn ngữ hoàn thiện được sử dụng trong các hệ thống cơ sở dữ liệu và là một thành phần không thể thiếu trong các hệ quản trị cơ sở dữ liệu. Mặc dù SQL không phải là một ngôn ngữ lập trình như C, C++, Java,... song các câu lệnh mà SQL cung cấp có thể được nhúng vào trong các ngôn ngữ lập trình nhằm xây dựng các ứng dụng tương tác với cơ sở dữ liệu.

## **Một số cơ sở nghiên cứu khác**

* **JavaScript**, theo phiên bản hiện hành, là một ngôn ngữ lập trình kịch bản dựa trên đối tượng được phát triển từ các ý niệm nguyên mẫu. Ngôn ngữ này được dùng rộng rãi cho các trang web, nhưng cũng được dùng để tạo khả năng viết script sử dụng các đối tượng nằm sẵn trong các ứng dụng. Nó vốn được phát triển bởi Brendan Eich tại Hãng truyền thông Netscape với cái tên đầu tiên Mocha, rồi sau đó đổi tên thành LiveScript, và cuối cùng thành JavaScript. Giống Java, JavaScript có cú pháp tương tự C, nhưng nó gần với Self hơn Java. .js là phần mở rộng thường được dùng cho tập tin mã nguồn JavaScript.
* **CSS**: Trong tin học, các tập tin định kiểu theo tầng – dịch từ tiếng Anh là Cascading Style Sheets (CSS) – được dùng để miêu tả cách trình bày các tài liệu viết bằng ngôn ngữ HTML và XHTML. Ngoài ra ngôn ngữ định kiểu theo tầng cũng có thể dùng cho XML, SVG, XUL v.v... Các đặc điểm kỹ thuật của CSS được duy trì bởi World Wide Web Consortium (W3C). Thay vì đặt các thẻ qui định kiểu dáng cho văn bản HTML (hoặc XHTML) ngay trong nội dung của nó, bạn nên sử dụng CSS.

# **CHƯƠNG 3: PHÂN TÍCH HỆ THỐNG**

## **Mô tả bài toán**

Tên bài toán : “Xây dựng trang web bán hàng trực tuyến sử dụng ngôn ngữ lập trình ASP.NET MVC và cơ sở dữ liệu SQL Server”. Đây là một loại hình kinh doanh ảo qua mạng bằng cách xây dựng một trang web bán hàng để mọi người có thể vào đó xem hàng, lựa chọn hàng để mua hàng. Với trang web bán hàng trực tuyến, nhà kinh doanh sẽ đưa thông tin của các loại điện thoại lên trang web của mình. Việc đưa thông tin cũng phải theo một trật tự hợp lý. VD như đưa thông tin theo các danh mục sản phẩm như Cafe, Trà, Ăn sáng, ….Thông tin được đưa lên bao gồm: tên sản phăm, hình ảnh, giá cả, …Người quản trị trang web cũng có quyền thêm mới (khi có mặt hàng mới về), sửa đổi thông số của mặt hàng, và xóa đi những mặt hàng không còn hoặc đã lỗi. Mọi sự thay đổi đều được ghi nhận trong cơ sở dữ liệu để đến với người sử dụng. Với khách vãng lai (guest) thì được xem trang web, xem thông tin về các loại mặt hàng, nhưng không có quyền mua hàng. Để mua hàng, bạn phải đăng ký là thành viên của trang web. Khi đã là thành viên của website, bạn đăng nhập vào website theo tên truy nhập và mật khẩu của mình. Khi đó, bạn có quyền mua hàng. Với các mặt hàng, bạn có quyền chọn mua hàng, đưa nó vào trong giỏ hàng của mình với số lượng các mặt hàng. Bạn cũng có thể xóa đi loại sản phẩm mà mình đã cho vào giỏ hàng của mình. Ở trang này, bạn phải thanh toán bằng tiền mặt hoặc chuyển khoản trực tiếp cho shop thông qua số tài khoản. Ngoài ra bạn cũng phải đưa các thông tin về địa chỉ giao hàng, số điện thoại của mình để tiện cho việc giao hàng.

Khi có một đơn đặt hàng, người quản lý sẽ duyệt đơn đặt hàng và gửi mail thông báo cho khách hàng biết đơn hàng đã được duyệt.

* + 1. **Yêu cầu bài toán**

**Đối với khách vãng lai (guest):**

* Được xem thông tin về các mặt hàng sản phâm, các thông tin liên quan đến cửa hàng, liên hệ với trang web.
* Không được quyền mua hàng.

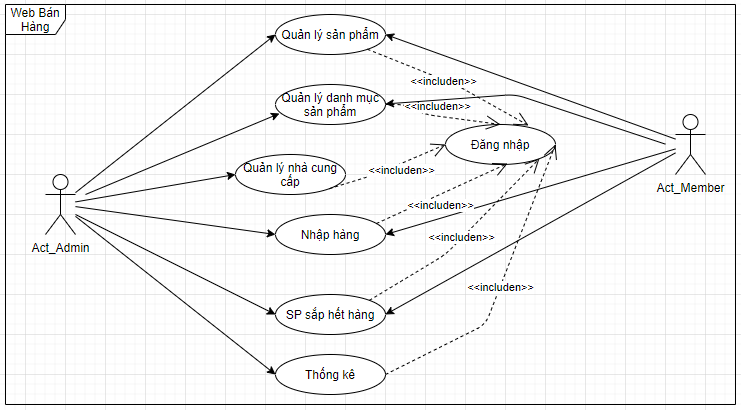
**Đối với thành viên trang web:**

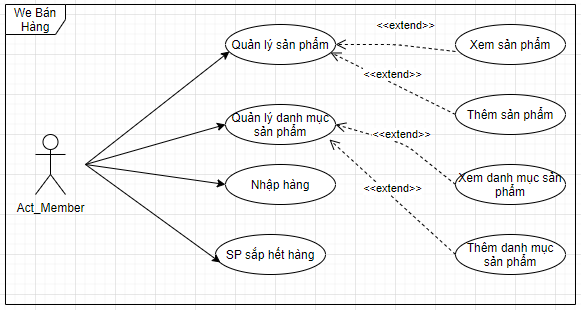
* Được xem thông tin về các mặt hàng, các thông tin liên quan đến cửa hàng, liên hệ tới trang web.
* Được quyền mua hàng khi mặt hàng đó còn trong kho hàng. Mọi sản phẩm thành viên mua sẽ được cho vào giỏ hàng của họ.
* Được quyền loại bỏ đi,chỉnh sửa số lượng các mặt hàng sau khi mình cho vào trong giỏ hàng của mình mà chưa đặt hàng.

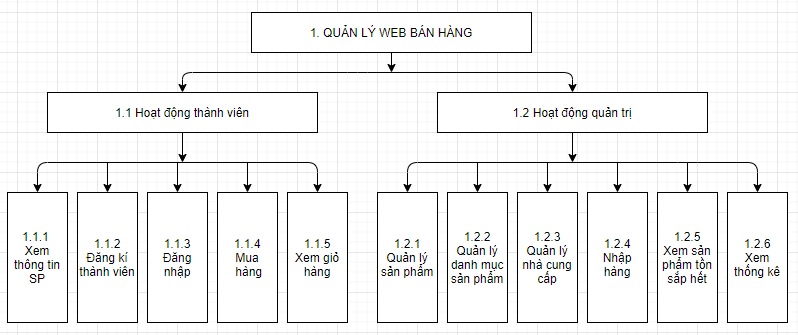
**Đối với người quản trị trang web (admin):**

* Được thêm mới thông tin về các loại sản phẩm: Khi có hàng mới về, admin sẽ cập nhật thông tin về sản phẩm vào trong cơ sở nhập hàng về.
* Được quyền sửa đổi các thông tin sao cho phù hợp với mặt hàng
* Được quyền xóa bỏ đi các loại sản phẩm khi sản phẩm đó đã hết hoặc cửa hàng không còn kinh doanh các mặt hàng đó nữa.
* Được thêm mới, chỉnh sửa, xóa bỏ nhà cung cấp.
* Được nhập hàng từ nhà cung cấp vào cửa hàng.
* Duyệt đơn hàng và gửi mail thông báo.
* Được xem thống kê.
* Phân quyền cho thành viên tham gia vào trang web.

### **Usecase**







# **CHƯƠNG 4: THỬ NGHIỆM VÀ ĐÁNH GIÁ**

## **Giới thiệu bài toán**

# **KẾT LUẬN VÀ KIẾN NGHỊ**

**Kết quả đạt được**

Trong suốt khoảng thời gian nghiên cứu và thực hiện đề tài, em đã cố gắng làm hết khả năng của mình và vì thời gian làm đồ án có hạn nên kết quả đạt được vẫn còn nhiều hạn chế nhưng em cũng học hỏi được khá nhiều kiến thức mới mẻ cũng như chuyên sâu về thuật toán, lập trình... và những việc em đã đạt được như sau:

* Trình bày các kiến thức cơ bản về máy học, về công nghệ, ngôn ngữ
* Giới thiệu tổng quan các kiến thức về phân cụm dữ liệu, phương pháp phân cụm phân hoạch, phương pháp Elbow Method.
* Trình bày thuật toán phân cụm K-means, ví dụ minh hoạ và bài toán phân cụm văn bản.
* Xây dựng được ứng dụng phân cụm văn bản tin tức.

**Tồn tại**

Trong quá trình làm đồ án, hiển nhiên sẽ có những lỗi chưa khắc phục hoàn toàn được cùng với thời gian có hạn nên sẽ có những chức năng chưa hoàn thiện:

* Giao diện chương trình còn sơ xài
* Các chức năng còn chưa thực sự hoàn thiện
* Số lượng chức năng còn ít
* Tập dữ liệu chạy thực nghiệm còn hạn chế về số lượng

**Hướng phát triển**

- Cần tối ưu xử lý dữ liệu đầu vào áp dụng triệt để việc giảm độ lớn ma trận để phân cụm được số lượng lớn dữ liệu với thời gian thực hiện thấp, độ chính xác đối với việc đề xuất số cụm được cao hơn.

- Tiếp tục phát triển hoàn thiện hệ thống giúp cho người dùng thực hiện phân cụm với số lượng tài liệu lớn hơn và mở rộng thực hiện qua phân cụm văn bản tiếng việt, ứng dụng kiểm tra độ trùng nhau giữa các văn bản, áp dụng thêm phương pháp tính toán điểm bóng để tối ưu việc lựa chọn số cụm hợp lý.

# **PHỤ LỤC**

**Phụ lục 1: Hướng dẫn sử dụng**

Ứng dụng gồm các chức năng

* Đọc dữ liệu đầu vào từ file csv
  + Chọn file csv làm tập dữ liệu tin tức để tiến hành phân cụm
* Nhập dữ liệu từ tay, các loại văn bản tuỳ ý người dùng
  + Chọn thêm dữ liệu từ các ô làm tập dữ liệu để tiến hành phân cụm
* Chuẩn hoá dữ liệu
  + Làm sạch dữ liệu đầu vào
  + Tạo danh sách chứa thuật ngữ
* Phân cụm bằng K-means
  + Phân cụm với dữ liệu vừa được
  + Chọn danh sách phân cụm hiển thị kết quả và đặc trưng phân cụm
* Đề xuất số cụm tối ưu
  + Chọn file csv làm tập dữ liệu và khoảng cụm k để tiến hành chạy chọn số cụm tốt nhất
  + Vẽ đồ thị đường khuỷu tay từ các giá trị WSS
* Ứng dụng chương trình
  + Chọn file csv cần thực hiện và chọn số cụm tối ưu vừa được đề xuất.
  + Xuất kết quả vừa được phân cụm với các đặc trưng để giúp sắp xếp các nhóm tin tức theo mục đích người dùng.

# **TÀI LIỆU THAM KHẢO**

1. Nguyễn Hoàng Hà, *Giáo trình C# và ứng dụng*, 2010
2. Hoàng Xuân Huấn, *Giáo trình Học Máy*, Đại học Quốc gia Hà Nội, 2015
3. Machine Learning là gì, https://nguyenvanhieu.vn/machine-learning-la-gi/, Truy cập ngày 26/5/2019 10:32 AM
4. Machine Learning cơ bản, https://machinelearningcoban.com, Truy cập ngày 26/5/2019 9:46 PM
5. Nguyễn Hoàng Tú Anh, *Giáo trình Khai thác dữ liệu và ứng dụng*, Đại học KHTN Tp Hồ Chí Minh, 2009
6. Osmar R.Zaiane, *Principles of knowledge discovery in databases,* University of Alberta, 2007
7. Võ Thị Ngọc Châu, *Giáo trình điện tử Cao Học Ngành Khoa Học Máy Tính,* Khoa Khoa Học & Kỹ Thuật Máy Tính Trường Đại Học Bách Khoa Tp. Hồ Chí Minh
8. Vi Văn Sơn, *Phân cụm thô của dữ liệu tuần tự*, Trường Đại học Công nghệ, ĐHQG Hà Nội
9. Tất tần tật về machine learning, https://techtalk.vn/tat-tan-tat-moi-kien-thuc-co-ban-ve-machine-learning.html, Truy cập ngày 29/5/2019 3:20 PM
10. .NET Framework, https://vi.wikipedia.org/wiki/.NET\_Framework, Truy cập ngày 25/5/2019 8:30 AM
11. Hoàng Văn Dũng, *Khai phá bằng kỹ thuật phân cụm*, Đại học SP Hà Nội, 2010
12. TF-IDF và Cosine Similarity, http://thanhvie.com/tf-idf-va-cosine-similarity/, Truy cập ngày 2/6/2019 6:15 PM
13. Wenyi Ni, *A Survey of Web Document Clustering*, Southern Methodist University, 2006
14. Oren Zamir and Oren Etzioni, *Web document Clustering: A Feasibility Demonstration*, University of Washington, USA, ACM, 2005
15. Phân cụm K-means, https://dominhhai.github.io/vi/2018/02/ml-kmeans/, Truy cập ngày 1/6/2019 5:22 PM
16. Khoo Khyou Bun, *Topic Trend Detection and Mining in World Wide Web*, A thesis for the degree of PhD, Japan, 2004
17. Thuật toán K-means, http://ungdung.khoahnvd.com/Hoc\_thuat/KMeans.html, Truy cập ngày 30/5/2019 2:54 PM
18. Vấn đề xác định số cụm trong phân cụm, http://bis.net.vn/forums/t/1831.aspx, Truy cập ngày 30/5/2019 4:35 PM
19. Bing Liu, Web mining, Springer, 2007