

**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THỰC PHẨM TP.HCM**

**KHOA CÔNG NGHỆ THÔNG TIN**

🙜🙢🙠🙞



**ĐỒ ÁN MÔN HỌC**

**TRÍ TUỆ NHÂN TẠO**

**ĐỀ TÀI: DỰ ĐOÁN KHẢ NĂNG MUA XE**

Giảng viên hướng dẫn: **Võ Thị Hồng Tuyết**

Sinh viên thực hiện:

1. 2001190215 – Trần Gia Phúc
2. 2001207294 – Trần Minh Quân

# **BẢNG PHÂN CHIA CÔNG VIỆC**

|  |  |
| --- | --- |
| 2001207294 -Trần Minh Quân | * Thiết kế ppt, word |
| 2001190215 - Trần Gia Phúc | * Xây dựng chương trình |

**MỤC LỤC**

[**DANH MỤC BẢNG BIÊU 5**](#_Toc18417)

[**DANH MỤC HÌNH ẢNH 6**](#_Toc13899)

[**LỜI CẢM ƠN 7**](#_Toc5433)

[**Chương 1 GIỚI THIỆU 8**](#_Toc5867)

[1.1 Giới thiệu đề tài 8](#_Toc8468)

[1.2 Nội dung nghiên cứu 8](#_Toc21706)

[1.3 Phương pháp nghiên cứu 8](#_Toc834)

[1.4 Giới hạn đề tài 8](#_Toc782)

[1.5 Bố cục báo cáo 9](#_Toc69)

[**Chương 2 CƠ SỞ LÝ THUYẾT 9**](#_Toc7965)

[3.1 Khái niệm cơ bản 9](#_Toc23858)

[2.1.1 Khái niệm dự báo 9](#_Toc20553)

[*2.1.1.1* Khái niệm dự báo 9](#_Toc14189)

[*2.1.1.2* Ý nghĩa 10](#_Toc10100)

[*2.1.1.3* Vai trò 10](#_Toc9151)

[3.2 Nghiên cứu liên quan 11](#_Toc29927)

[2.2.1 Các phương pháp dự báo 11](#_Toc31897)

[*2.2.1.1* Phương pháp dự báo định tính 11](#_Toc24225)

[*2.2.1.2* Lấy ý kiến của ban điều hành 11](#_Toc22293)

[*2.2.1.3* Lấy ý kiến của người bán hàng 11](#_Toc4300)

[*2.2.1.4* Phương pháp chuyên gia(delphi) 12](#_Toc19796)

[*2.2.1.5* Phương pháp điều tra người tiêu dùng 12](#_Toc28369)

[*2.2.1.6* Phương pháo dự báo định lượng 13](#_Toc25505)

[2.2.2 Thuật toán KNN 13](#_Toc28117)

[*2.2.2.1* Khái Niệm 13](#_Toc19582)

[*2.2.2.2* Quy trình thực hiện thuật toán 13](#_Toc6586)

[*2.2.2.4* Ưu điểm , nhược điểm của thuật toán 14](#_Toc4116)

[*2.2.2.5* Ứng dụng 15](#_Toc8005)

[**Chương 3 DỰ ĐOÁN KHẢ NĂNG MUA XE 16**](#_Toc346)

[*3.1* Form chương trình chính 16](#_Toc24876)

[*3.2* Chức năng mở file dữ liệu từ excel 16](#_Toc24458)

[*3.3* Chức năng tính khoảng cách giữa 2 điểm trong toạ độ 18](#_Toc11863)

[*3.4* Chức năng tính khoảng cách tất cả hàng dữ liệu 19](#_Toc30016)

[*3.5* Chức năng chọn dữ liệu theo k 21](#_Toc629)

[*3.6* Chức năng hiển thị dữ liệu đã chọn theo k và xuất kết quả dự đoán 22](#_Toc32079)

[**Chương 4 KẾT QUẢ - KẾT LUẬN 23**](#_Toc8618)

[*4.1* Kết quả - kết luận 24](#_Toc9608)

[*4.2* Ưu ,nhược điểm của đề tài 24](#_Toc5624)

[*4.3* Hướng phát triển của đề tài 25](#_Toc22310)

**DANH MỤC BẢNG BIÊU**

[***Bảng 3.1 Bảng dữ liệu khách hàng mua xe* 18**](#_Toc27865)

**DANH MỤC HÌNH ẢNH**

[***Hình 2.1 ví dụ minh hoạ thuật toán KNN* 15**](#_Toc29529)

[***Hình 3.1 From chương trình chính* 17**](#_Toc6591)

[***Bảng 3.1 Bảng dữ liệu khách hàng mua xe* 18**](#_Toc1331)

[***Hình3.3 chức năng tính khoảng cách toạ độ* 20**](#_Toc26820)

[***Hình 3.4 tính khoảng cách tiền lương* 21**](#_Toc26015)

[***Hình 3.5 chức năng chọn dữ liệu theo k* 22**](#_Toc8677)

[***Hình 4.1 kết quả “Mua”* 25**](#_Toc2785)

[***Hình 4.2 kết quả “ Không Mua”* 25**](#_Toc26038)

# **LỜI CẢM ƠN**

Chúng em xin chân thành cảm ơn thầy thuộc khoa Công Nghệ Thông Tin đã tận tình giảng dạy, trang bị những kiến thức vô cùng bổ ích trong thời gian qua, giúp nhóm có một kiến thức nền tảng để thực hiện được đề tài này.

Nhóm cũng xin chân thành cảm ơn quý thầy cô khoa Công Nghệ Thông Tin đã tạo điều kiện, thời gian tốt nhất cho chúng em có thể hoàn thành đồ án môn học này. Mặc dù vậy, nhưng với trình độ và kinh nghiệm chưa cao nên trong quá trình hoàn thành đồ án khó tránh những lỗi sai sót không đáng có. Chúng em kính mong nhận được sự thông cảm và đóng góp ý kiến của quý thầy cô và các bạn bè xung quanh để có thể nâng cao hơn trình độ của cá nhân nói riêng và của cả nhóm đề tài nói chung.

Một lần nữa, nhóm xin chân thành cảm ơn thầy.

1. **GIỚI THIỆU**
2. **Giới thiệu đề tài**

Ngày nay, việc ứng dụng trí tuệ nhân tạo (AI) được coi là một trong những yếu tố quyết định đến hoạt động của các tổ chức của công ty, nó có vai trò rất quan trọng, có thể tạo ra những bước đột phá mạnh mẽ. Hiện nay, trong cuộc sống của chúng ta có rất nhiều cửa hàng bán xe nổi tiếng khác nhau trong và ngoài nước vậy thì cách mà họ có thể quản lý được khả năng có thể mua xe của mình là như thế nào. Sử dụng phần mềm để dự đoán khả năng là sự lựa chọn hợp lý của họ. Vì vậy hiện nay nhóm chúng em đã phát triển và nghiên cứu một phần mềm để quản lý được khả năng mua xe một cách dễ dàng nhất, với phần mềm này bạn có thể dự đoán được khả năng có mua xe được hay không. Phần mềm có tên “DỰ BÁO KHẢ NĂNG MUA XE”.

1. **Nội dung nghiên cứu**

* Tổng quan phân tích, dự báo khả năng mua xe
* Khảo sát, phân tích hệ thống - thiết kế hệ thống dự đoán
* Xây dựng chương trình.

1. **Phương pháp nghiên cứu**

* Đặc tả lý thuyết.
* Nghiên cứu kỹ thuật phân tích dữ liệu và dự đoán về khách hàng
* Nghiên cứu và phân tích 1 số phương pháp hồi quy
* Thu thập số liệu thông tin khách hàng
* Xây dựng chương trình dự đoán khả năng mua xe của khách hàng để đẩy quảng cáo

1. **Giới hạn đề tài**

* Nghiên cứu tổng quan về phân tích dữ liệu và dự đoán về khả năng mua xe của khách hàng
* Nghiên cứu một số phương pháp liên quan và đề xuất phương pháp phù hợp để dự đoán khả năng mua xe của khách hàng

1. **Bố cục báo cáo**

* Chương 1 : Giới thiệu
* Chương 2 : Cơ sở lý thuyết.
* Chương 3 : Dự đoán khả năng mua xe
* Chương 4 : Kết quả - Kết luận.

1. **CƠ SỞ LÝ THUYẾT**
2. **Khái niệm cơ bản**
   1. **Khái niệm dự báo**
      1. **Khái niệm dự báo**

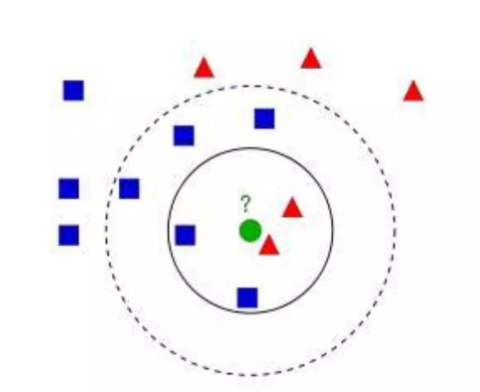
* Dự báo dã hình thành từ đầu nhữmg năm 60 của thế kỉ 20. Khoa học dự báo với tư cách một ngành khoa học độc lập có hệ thống lí luận, phương pháp luận và phương pháp hệ riêng nhằm nâng cao tính hiệu quà của dự báo. Người ta thường nhấn mạnh rằng một phuơng pháp tiếp cận hiệu quả đối với dự báo là phần quan trọng trong hoạch định. Khi các nhà quản trị lên kế hoạch, trong hiện tại họ xác định hướng tương lai cho các hoạt động mà họ sẽ thực hiện. Bước đầu tiên trong hoạch định là dự báo hay là ước lượng nhu cầu trong lai cho sản phẩm hoặc dịch vụ và các nguồn lực cần thiết để sản xuất sản phẩm hoặc dịch vụ đó.
* Như vậy, dự báo là một khoa học và nghệ thuật tiên đoán những sự việc sẽ xảy ra trong tong lai, trên cơ sở phân tích khoa học về các dữ liệu đã thu thập được.
* Khi tiến hành dự báo ta căn cứ vào việc thu nhập xử lý số liệu trong quá khứ và hiện tại để xác định xu hướng vận động của các hiện tượng trong tương lai nhờ vào một số mô hình toán học.
* Dự báo có thể là dự báo chủ quan hoặc trực quan về tương lai. Nhưng để lại cho dự báo được chính xác hơn, người ta cố loại trừ những tính chủ quan của người dự báo.
* Ngày nay dự báo là một nhu cầu không thể thiếu của mọi hoạt động kinh tế-xã hội, khoa học - kỹ thuật, được tất cả các ngành khoa học quan tâm nghiên cứu.
  + 1. **Ý nghĩa**
* Dùng để dự báo các mức độ tương lai của hiện tượng, qua đó giúp các nhà quản trị doanh nghiệp chủ động trong công việc đề ra các kế hoạch và các quyết định cần thiết phục vụ cho quá trình sản xuất, kinh doanh, đầu tư, quảng bá,… và chuẩn bị đầy đủ điều kiện cơ sở vật chất, kỹ thuật cho sự phát triển trong thời gian tới(kế hoạch cung cấp các yếu tố đầu vào như:lao động, nguyên vật liệu, tư liệu lao động….cũng như các yếu tố đầu ra dưới dạng sản phẩm vật chất và dịch vụ).
* Trong các doanh nghiệp nếu công tác các dự báo được thực hiện một cách nghiêm túc còn tạo điều kiện nâng cao khả năng cạnh tranh trên thị trường.
* Dự báo chính xác sẽ giảm bớt mức độ rủi ro cho doanh nghiệp nói riêng và toàn toàn bộ nền kinh tế nói chung.
* Dự báo chính xác là căn cứ để các nhà hoạch định các chính sách phát triển kinh tế văn hoá xã hội trong toàn bộ nền kinh tế quốc dân
* Nhờ có dự báo các chính sách kinh tế các kế hoạch và chương trình phát triển kinh tế được xây dựng có cơ sở khoa học và mang lại hiệu quả kinh tế cao.
* Nhờ có dự báo thường xuyên và kịp thời, các nhhaf quản trị doanh nghiệp có khả năng kịp thời đưa ra những biện pháp điều chỉnh các hoạt động kinh tế của đơn vị mình nhằm thu được hiểu quả sản xuâ kinh doanh cao nhất.
  + 1. **Vai trò**
* Dự báo tạo ra lợi thế cạnh tranh
* Công tác dự báo là một bộ phận không thể thiếu trong hoạt động của các doanh nghiệp, trong từng phòng ban như: phòng kinh doanh, phòng sản xuất,marketing….,

1. **Nghiên cứu liên quan**
   1. **Các phương pháp dự báo**
      1. **Phương pháp dự báo định tính**

* Các phương pháp này dựa trên cơ sở nhận xét của những nhân tố nhân quả dựa theo doanh số của từng sản phẩm hay dịch vụ riêng biệt và ựự trên những ý kiến về các khả năng có liên quan đến mức độ phức tạp khác nhau, từ những khảo sát ý kiến được tiến hành một cách khoa học để nhận biết sự kiện tương lai
* Ưu điểm: Dễ dàng thực hiện, không đòi hỏi kiến thức về các mô hình toán hoặc kinh tế lượng, thường được chấp nhận
* Nhược điểm: mang tính chủ quan rất cao, không chuẩn, mất nhiều năm để trở thành người có khả năng phán đoán đúng. Không có phương pháp hệ thống để đánh giá và cải thiện mức độ chính xác.
  + 1. **Lấy ý kiến của ban điều hành**
* Phương pháp này được sử dụng rộng rãi ở các doanh nghiệp. Khi tiến hành dự báo, họ lấy ý kiến của các nhà quản trị cấp cao , những người phụ trách các công việc, các bộ phận doanh nghiệp và sử dụng các số liệu thống kê về những chỉ tiêu tổng hợp:doanh số, chi phí, lợi nhuận…. Ngoài ra cấn lấy thêm ý kiến của các chuyên gia marketing, tài chính, sản xuất, kỹ thuật.
* Ưu điểm của phương pháp này là: thu nhập được nhiều kinh nghiêm từ nhiều chuyên gia khác nhau.
* Nhược điểm lớn nhất của phương pháp này là có tính chủ quan của các thành viên và ý kiến của người có chực vụ cao nhất thường chi phối ý kiến của những người khác
  + 1. **Lấy ý kiến của người bán hàng**
* Ưu điểm cảu phương pháp này là: những người bán hàng tiếp xúc thường xuyên với khách hàng, do đó họ hiểu rõ nhu cầu, thị hiếu của người dùng. Họ có thể dự đoán được lượng hàng tiêu thụ tại khu vực mình phụ trách.
* Tập hợp ý kiến của nhiều người bán hàng tại nhiều khu vực khác nhau, ta có được lượng dự báo tổng hợp về nhu cầu đối với loại sản phẩm đang xét.
* Nhược điểm của phương pháp này là phụ thuộc vào đánh giá chủ quan của người bán hàng. Một số có khuynh hướng lạc quan đánh giá cao lượng hàng bán ra của mình. Ngược lại, một số khác lại muốn giảm xuống để dễ đạt định mức
  + 1. **Phương pháp chuyên gia(delphi)**
* Phương pháp này thu nhập ý kiến của các chuyên gia trong hoặc ngoài doanh nghiệp theo từng khuôn mẫu câu hỏi được in sẵn và được thực hiện như sau:
* Mỗi chuyên gia được phát như một thư yêu cầu trả lời một số câu hỏi phục vụ cho việc dự báo
* Nhân viên dự báo tập hợp các câu trả lời, sắp xếp chọn lọc và tóm tắt các ý kiến của các chuyên gia.
* Dựa vào bảng tóm tắt nhân viên dự báo lại tiếp tục nêu ra các câu hỏi để chuyên gia trả lời trực tiếp,
* Tập hợp các ý kiến của các chuyên gia. Nếu chưa thoả mãn thì tiếp tục quá trình nêu trên đến khi đạt yêu cấu đự báo.
* Ưu điểm của phương pháp này là tránh được các liên hệ cá nhân với nhau, không xảy ra va chạm giữa các chuyên gia và họ không ảnh hưởng bởi ý kiến của một người nào đó có ưu thế trong số người được hỏi ý kiến.
  + 1. **Phương pháp điều tra người tiêu dùng**
* Phương pháo này sẽ thu nhập nguồn thông tin từ đối tượng người tiêu dùng về nhu cầu hiện tại cũng như tương lai. Cuộc điều tra nhu cầu được thực hiện bởi những nhân ciên bán hàng hoặc nhân viên nghiên cứu thị trường. Họ thu nhập ý kiến khách hàng thông qua phiếu điều tra, phỏng vấn trực tiếp hay điện thoại… Cách tiếp cận này không những giúp cho doanh nghiệp về dự báo nhu cầu mà cả trong việc cải tiến thiết kế sản phẩm. Phương pháp này mất rất nhiều thời gian, việc chuẩn bị phức tạp, khó khăn và tốn kém, có thể không chính xác trong các câu trả lời của người tiêu dùng.
* Ưu điểm: cách tốt nhất để dự báo nhu cầu, sở thích của họ qua dự định mua sắm của họ, điều tra được thị hiếu của khách hàng để cải tiến sản phẩm.
* Nhược điểm: phù hợp cho các sản phẩm công nghiệp, tính chính xác của dữ liệu.
  + 1. **Phương pháo dự báo định lượng**
* Mô hình dự báo định lượng dựa trên số liệu quá khứ, những số liệu này giả sử có liên quan đến tương lai và có thể tìm thấy được. Tất cả các mô hình dự báo theo định lượng có thể sử dụng thông qua chuỗi thời gian và các giá trị được quan sát do lường các giai đoạn theo từng chuỗi.

Ưu điểm:

* Kết quả dự báo có tính khách quan
* Có phương pháp đo lường độ chính xác dự báo
* Tốn it thời gian để tìm ra kết quả dự báo
* Nhược điểm: chỉ dự báo tốt trong thời gian ngắn và trung hạn
  1. **Thuật toán KNN**
     1. **Khái Niệm**
* Thuật toán K láng giềng gần nhất trong tiếng Anh là K-Nearest Neighbor, viết tắt là KNN.
* Thuật toán K láng giềng gần nhất là một kĩ thuật học có giám sát (supervised learning) dùng để phân loại quan sát mới bằng cách tìm điểm tương đồng giữa quan sát mới này với dữ liệu sẵn có.
  + 1. **Quy trình thực hiện thuật toán**
* B1: Xác định số K = số láng giềng gần nhất.
* B2: Tính khoảng cách từ đối tượng cần phân lớp với tất cả đối tượng trong training data.
* B3: Sắp xếp khoảng cách theo thứ tự tăng dần và xác định K láng giềng gần nhất đối với đối tượng cần phân lớp
* B4: Lấy tất cả các lớp của K láng giềng gần nhất.
* B5: Dựa vào phần lớn của lớp K để xác định lớp cho đối tượng cần phân lớp.
  + 1. **Ví dụ**

**

*Hình 2.1 ví dụ minh hoạ thuật toán KNN*

* + 1. **Ưu điểm , nhược điểm của thuật toán**

Ưu điểm

* Độ phức tạp tính toán nhỏ.
* Việc dự đoán kết quả của dữ liệu mới rất đơn giản.
* Thuật toán đơn giản dễ dàng triển khai, dễ sử dung, dễ cài đặt.

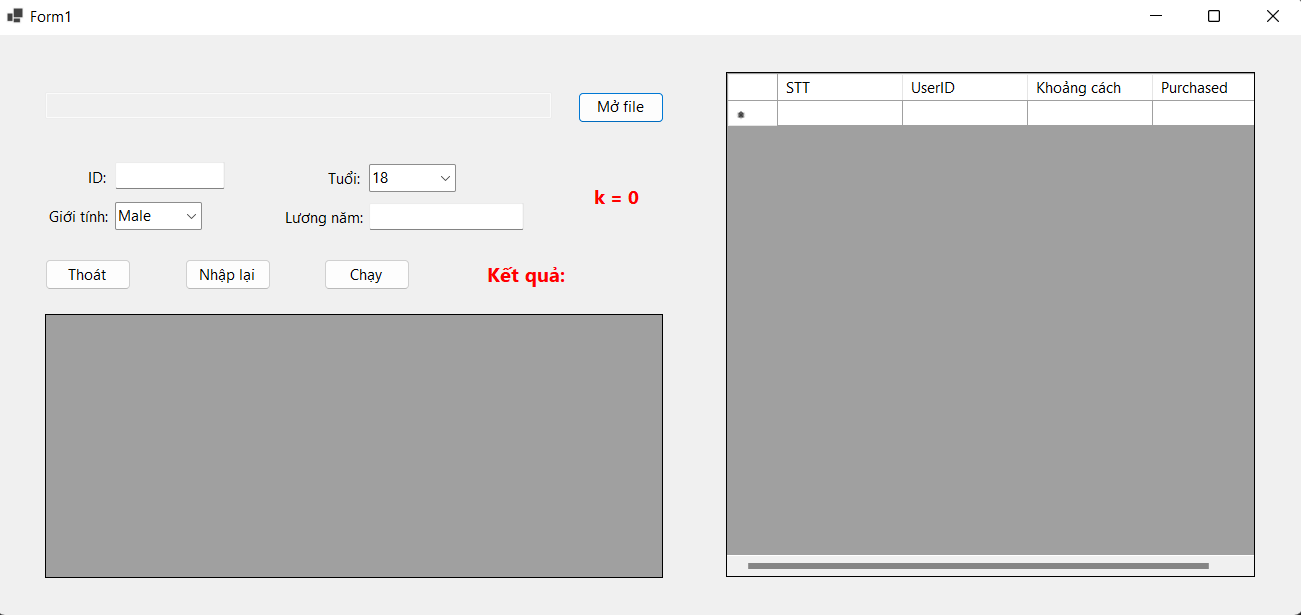
Nhược điểm

* Không thích hợp cho dữ liệu có kích thước lớn
* Tốn bộ nhớ
* Nhạy cảm với các dữ liệu bất thường (nhiễu)
* Sẽ trở nên rất chậm khi số lượng điểm dữ liệu tăng lên vì mô hình cần lưu trữ tất cả các điểm dữ liệu.
  + 1. **Ứng dụng**
* KNN là một mô hình đơn giản và trực quan nhưng vẫn có hiệu quả cao vì nó không tham số; mô hình không đưa ra giả định nào về việc phân phối dữ liệu. Hơn nữa, nó có thể được sử dụng trực tiếp để phân loại đa lớp.

Dưới đây là một số lĩnh vực mà thuật toán k-hàng xóm gần nhất có thể được sử dụng:

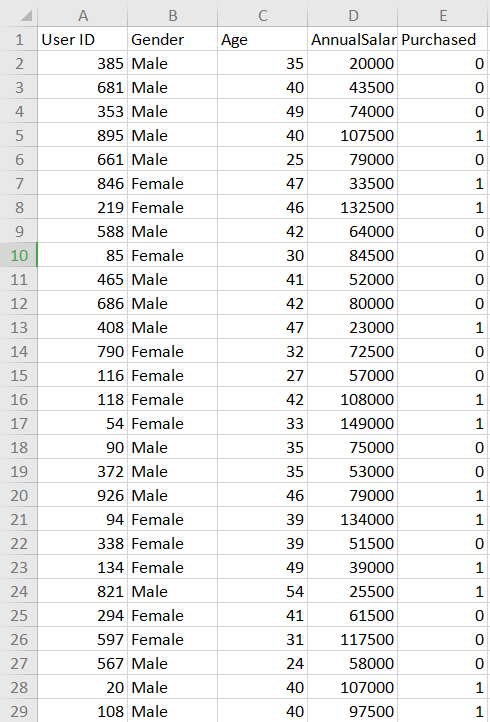
* **Xếp hạng tín dụng:** Thuật toán KNN giúp xác định xếp hạng tín dụng của một cá nhân bằng cách so sánh họ với những người có đặc điểm tương tự.
* **Phê duyệt khoản vay:** Tương tự như xếp hạng tín dụng, thuật toán k-gần nhất có lợi trong việc xác định các cá nhân có nhiều khả năng bị vỡ nợ hơn đối với các khoản vay bằng cách so sánh các đặc điểm của họ với những cá nhân tương tự.  
  **Tiền xử lý dữ liệu:** Tập dữ liệu có thể có nhiều giá trị bị thiếu. Thuật toán KNN được sử dụng cho một quá trình được gọi là truyền dữ liệu bị thiếu để ước tính các giá trị bị thiếu.
* **Nhận dạng mẫu:** Khả năng của thuật toán KNN để xác định các mẫu tạo ra một loạt các ứng dụng. Ví dụ, nó giúp phát hiện các mô hình sử dụng thẻ tín dụng và phát hiện các mô hình bất thường. Phát hiện mẫu cũng hữu ích trong việc xác định các mẫu trong hành vi mua hàng của khách hàng.
* **Dự đoán giá cổ phiếu:** Vì thuật toán KNN có khả năng dự đoán giá trị của các thực thể không xác định, nên nó hữu ích trong việc dự đoán giá trị tương lai của cổ phiếu dựa trên dữ liệu lịch sử.
* **Hệ thống khuyến nghị:** Vì KNN có thể giúp tìm kiếm những người dùng có đặc điểm tương tự, nó có thể được sử dụng trong hệ thống khuyến nghị. Ví dụ: nó có thể được sử dụng trong nền tảng phát video trực tuyến để đề xuất nội dung mà người dùng có nhiều khả năng xem hơn bằng cách phân tích những gì người dùng tương tự xem.
* **Thị giác máy tính:** Thuật toán KNN được sử dụng để phân loại hình ảnh. Vì nó có khả năng nhóm các điểm dữ liệu tương tự nhau, chẳng hạn như nhóm các con mèo lại với nhau và các con chó trong một lớp khác nhau, nó rất hữu ích trong một số ứng dụng thị giác máy tính .

1. **DỰ ĐOÁN KHẢ NĂNG MUA XE**
2. **Form chương trình chính**

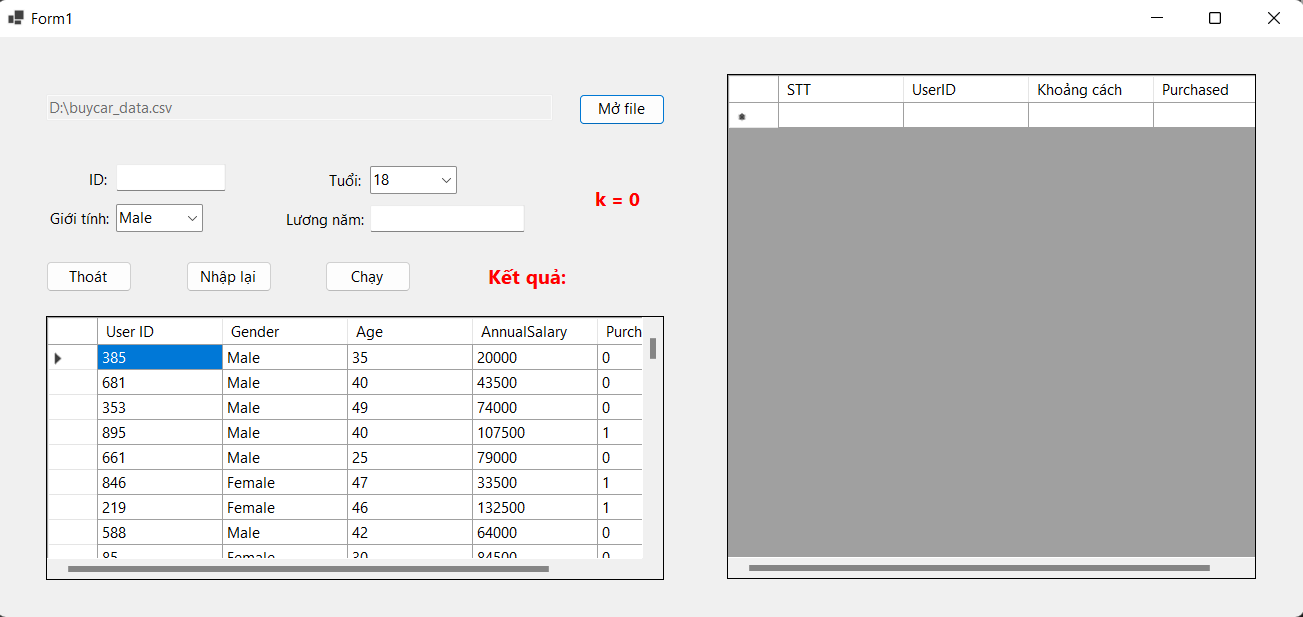


*Hình 3.1 Form chương trình chính*

1. **Chức năng mở file dữ liệu từ excel**



*Hình 3.1 Bảng dữ liệu khách hàng mua xe*



*Hình 3.2 Mở file*

*Input:*

Khi click vào nút Mở file sẽ hiện hộp thoại để chọn file dữ liệu .csv

*Output:*

Hiển thị địa chỉ của file đó lên textbox địa chỉ và đồng thời hiển thị dữ liệu lên datagridview ở dưới.

*Mã tự nhiên:*

* Đọc tuần tự ở dòng tiêu đề trong file
* Đọc từng cột tiêu đề trong 1 hàng, ngăn cách bởi dấu “,”.
* Thêm cột vào bảng DataTable.
* Đọc tuần tự các hàng dữ liệu trong file
* Ở mỗi hàng dữ liệu, đọc từng cột dữ liệu trong 1 hàng, ngăn cách bởi dấu “,”.
* Thêm hàng vào bảng DataTable.
* Hiển thị lên datagridview bên dưới từ bảng DataTable.

*Mã giả:*

Tạo bảng DataTable.

Foreach từng dòng tiêu đề trong file (lấy trong dòng 1)

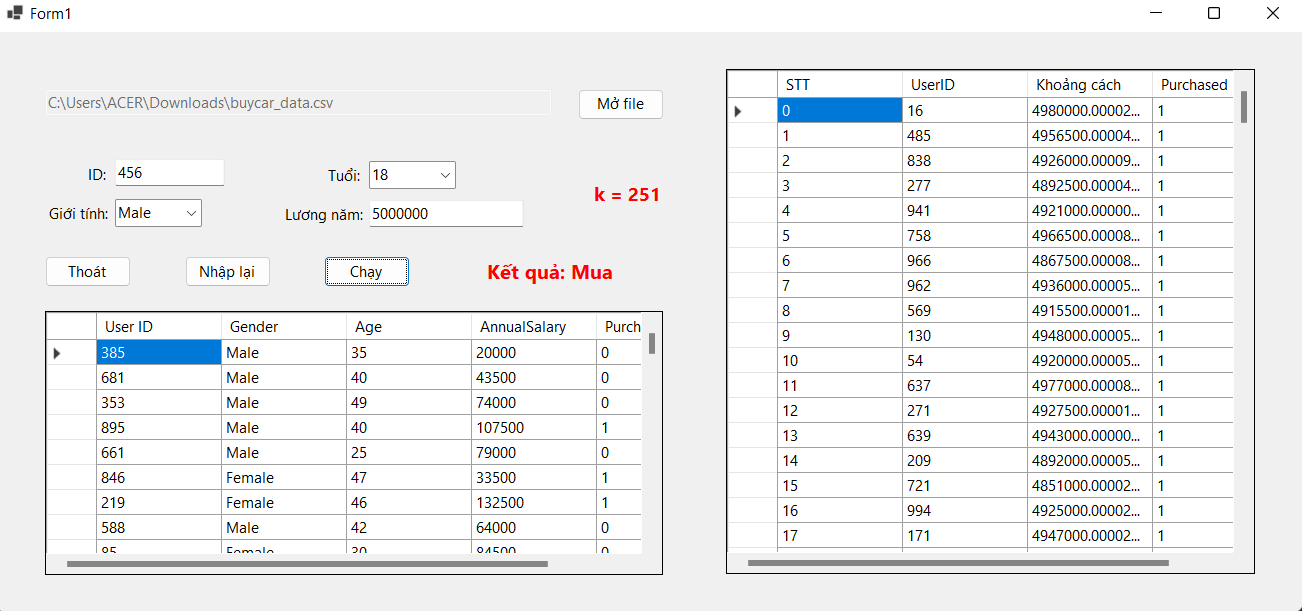
Foreach từng cột tiêu đề trong dòng tiêu đề, ngăn cách bởi dấu “,”, có loại bỏ tiêu đề trống

Thêm cột tiêu đề vào bảng DataTable.

Đọc dữ liệu trong file (skip dòng 1), ngăn cách bởi dấu “,”, foreach từng hàng thêm hàng vào bảng DataTable.

Hiển thị lên datagridview bên dưới từ bảng DataTable.

1. **Chức năng tính khoảng cách giữa 2 điểm trong toạ độ**



*Hình 3.3 Chức năng tính khoảng cách toạ độ*

*Input:*

Nhập vào toạ độ đầu tiên biến x, y, z và toạ độ thứ 2 biến x1, y1, z1

*Output:*

Tính ra khoảng cách giữa 2 điểm trong toạ độ bằng công thức Euclidean.

*Mã tự nhiên:*

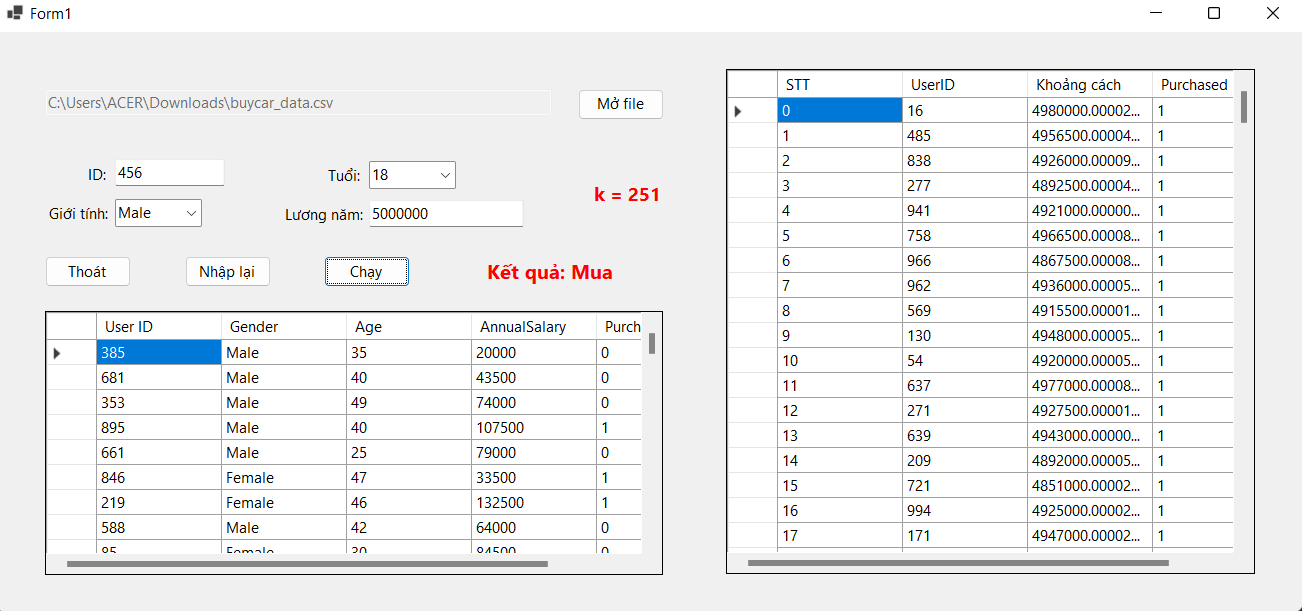
- Nhập vào toạ độ đầu tiên biến x, y, z và toạ độ thứ 2 biến x1, y1, z1

- Tính ra khoảng cách giữa 2 điểm trong toạ độ bằng công thức Euclidean, cụ thể là

*Mã giả:*

Return Math.sqrt của

1. **Chức năng tính khoảng cách tất cả hàng dữ liệu**



*Hình 3.4 Tính khoảng cách tiền lương*

*Input:*

Nhập vào các thông tin giới tính, tuổi, tiền lương.

*Output:*

Tính tất cả khoảng cách của tất cả hàng dữ liệu và đưa vào mảng khoảng cách.

*Mã tự nhiên:*

- Trong từng hàng dữ liệu trong datagridview bên dưới, lấy các thông tin dữ liệu giới tính, tuổi, tiền lương.

- Tính khoảng cách của các dữ liệu này cùng với dữ liệu người dùng mới nhập vào bằng chức năng tính khoảng cách 3.3 ở trên và gán vào mảng khoảng cách.

- Lặp lại đến khi duyệt hết datagridview bên dưới.

*Mã giả:*

Tạo mảng khoảng cách có kích thước phần tử n bằng với datagridview bên dưới.

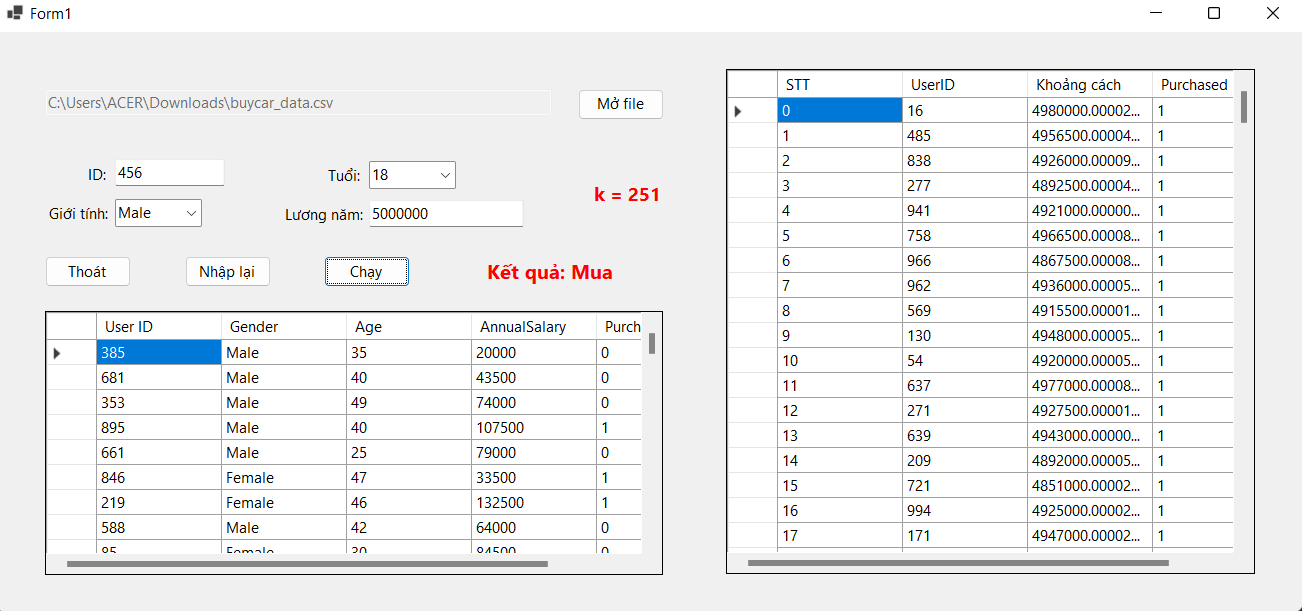
Nhận các thông tin giới tính (Male là 1, Female là 0), tuổi, tiền lương mà người dùng vừa nhập vào.

For từng hàng dữ liệu trong n

Nhận các thông tin giới tính (Male là 1, Female là 0), tuổi, tiền lương của hàng dữ liệu này.

Tính khoảng cách của các dữ liệu này cùng với dữ liệu người dùng mới nhập vào bằng chức năng tính khoảng cách 3.3 ở trên và gán vào mảng khoảng cách.

1. **Chức năng chọn dữ liệu theo k**



*Hình 3.5 chức năng chọn dữ liệu theo k*

*Input:*

Nhận k bằng 25% giá trị n (n là tổng hàng trong datagridview bên dưới, tự động).

*Output:*

Gán các vị trí của hàng dữ liệu có khoảng cách nhỏ nhất theo top k vào mảng vị trí.

*Mã tự nhiên:*

- Chọn k bằng 25% giá trị n, nếu k chẵn thì k cộng thêm 1, hiển thị k lên label “k = “.

- Tạo mảng trống (có k phần tử) để có thể nhận vị trí dữ liệu theo k.

- Lặp lại các câu lệnh dưới đây (từ 1 cho đến khi bằng k):

+ Duyệt vị trí từng hàng dữ liệu datagridview ở dưới, nếu mảng vị trí không chứa vị trí hàng dữ liệu này:

Nếu chưa có min, hoặc nếu đã có min và khoảng cách < min thì gán min bằng với khoảng cách này, và nhận vị trí này làm vị trí min.

+ Nếu tìm ra vị trí min, gán vị trí min này vào mảng vị trí.

*Mã giả:*

Chọn k bằng 25% giá trị n, nếu k chẵn thì k cộng thêm 1, hiển thị k lên label “k = “.

Tạo mảng chứa vị trí trống (có k phần tử) để có thể nhận vị trí dữ liệu theo k.

Tạo id = 0 (để xử lý mảng chứa vị trí này).

For step = 1 và step <= k

Gán min = -1, viTriI = -1

For i = 0 và i < n

If mảng chứa vị trí trống

If min = -1, hoặc min khác -1 và khoảng cách của i < min

Gán viTriI = vị trí i

Gán min = khoảng cách của i

If viTriI khác -1

Gán viTriI vào mảng chứa vị trí

Tăng id của mảng chứa vị trí lên 1

1. **Chức năng hiển thị dữ liệu đã chọn theo k và xuất kết quả dự đoán**

*Input:*

Lấy từ mảng chứa vị trí (đã được xử lý ở chức năng 3.5)

*Output:*

Hiển thị tất cả dữ liệu thông qua mảng chứa vị trí theo k lên datagridview bên phải, và hiển thị kết quả bằng label “Kết quả = “.

*Mã tự nhiên:*

- Tạo mảng nhận giá trị đã mua xe (có k phần tử), tạo biến đếm mua xe.

- Duyệt từng vị trí trong mảng chứa vị trí:

+ Gán giá trị đã mua xe vào mảng mua xe từ dữ liệu datagridview bên dưới.

+ Thêm hàng dữ liệu của vị trí này vào datagridview bên phải.

+ Nếu vị trí này không mua xe thì biến đếm trừ 1, ngược lại cộng 1.

- Xét nếu biến đếm lớn hơn 0 (tức có nhiều người mua xe hơn trong k) thì xuất ra label kết quả là “Mua”, ngược lại là “Không mua”.

*Mã giả:*

Tạo mảng nhận giá trị đã mua xe (có k phần tử)

Tạo biến đếm mua xe = 0.

For i = 0 và i < độ dài mảng chứa vị trí

Gán giá trị đã mua xe vào mảng mua xe từ dữ liệu datagridview bên dưới.

Thêm hàng dữ liệu của vị trí này vào datagridview bên phải.

If vị trí này không mua xe

Biến đếm trừ 1

Else biến đếm cộng 1

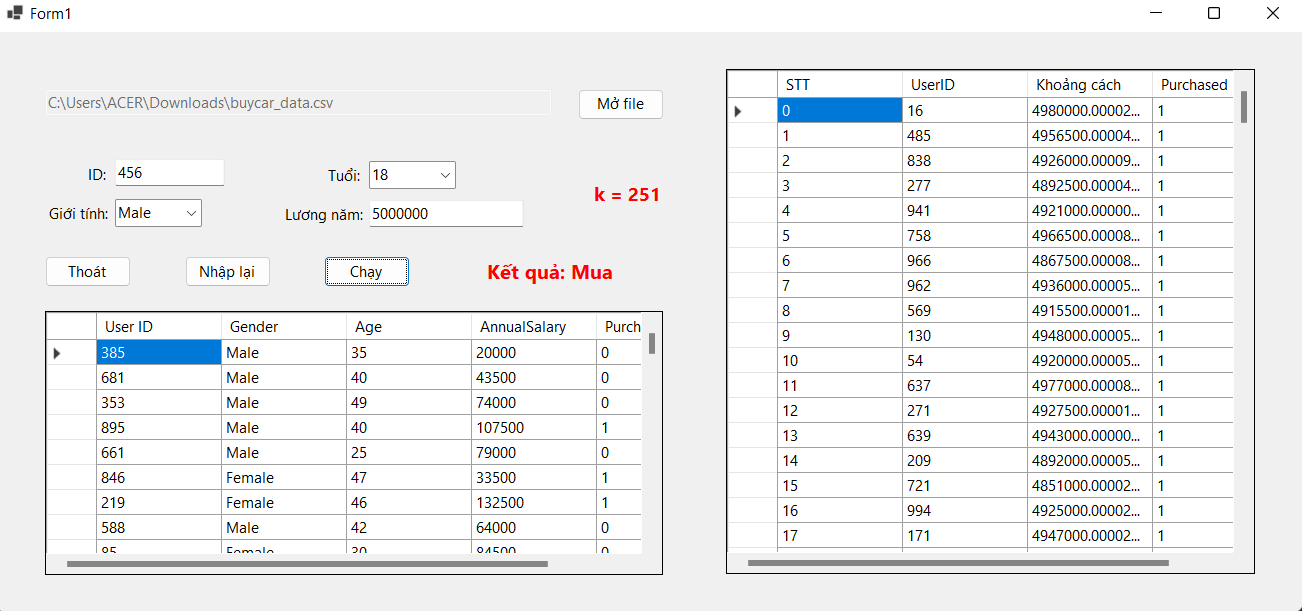
If đếm > 0

Hiển thị kết quả là “Mua”

Else

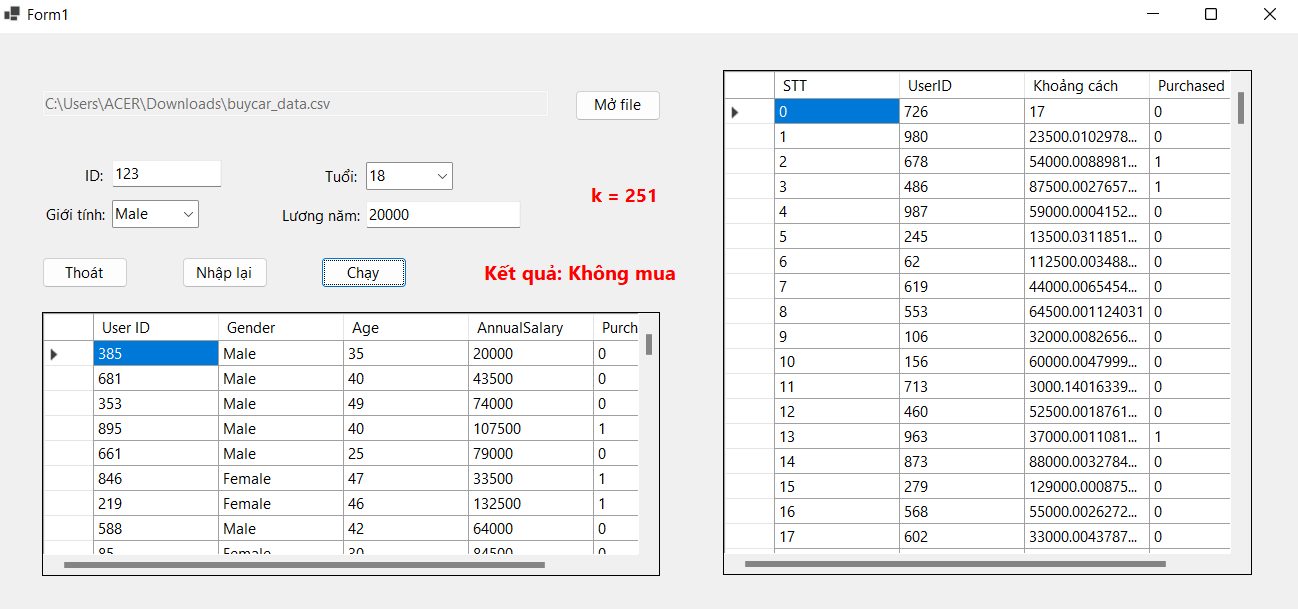
Hiển thị kết quả là “Không mua”

1. **KẾT QUẢ - KẾT LUẬN**
2. **Kết quả - kết luận**



*Hình 4.1 kết quả “Mua”*

* Với k= 101
* Khi nhập vào ID của khách hàng là “456” và mức lương “5000000” thì chương trình sẽ trả về kết quả là khách hàng đó sẽ “mua xe”



*Hình 4.2 kết quả “ Không Mua”*

* với khách hàng có ID là “123” và mức lương là “20000” thì chương trình trả về kết quả là khách hàng đó sẽ “không mua”

1. **Ưu ,nhược điểm của đề tài**

Ưu điểm:

* Giúp các doanh nghiệp dễ tiếp cận với khách hàng
* Đẩy quảng cáo sản phẩm cho khách hàng
* Chương trình dễ thực thi

Nhược điểm

* Giá trị K được set mặc định không set được K lên hoặc xuống

1. **Hướng phát triển của đề tài**

* Trong bài báo này, nhóm đã đề xuất một ứng dụng của thuật toán k-NN và mô hình “Dự đoán khả năng mua xe” dựa trên dữ liệu có sẵn nhằm phân lớp dữ liệu để chọn ra các phân loại khách hàng mua hay không mua.
* Hướng phát triển tiếp theo là sẽ trích xuất đặc trưng phù hợp với từng khách hàng đồng thời truy vấn ngữ nghĩa của các phân lớp hình ảnh khách hàng trên Ontology để tạo ra các ngữ nghĩa liên quan với các đối tượng trên ảnh.

**TÀI LIỆU THAM KHẢO**

<https://www.slideshare.net/trongthuy3/luan-van-chuong-trinh-du-bao-doanh-thu-ban-hang-cho-sieu-thi>

<https://phantichkinhdoanh.net/knn-thuat-toan-lang-gieng-gan-nhat-k-nearest-neighbor-la-gi/#Ung_dung_thuat_toan_KNN>

https://jstf.hufi.edu.vn/uploads/files/so-tap-chi/nam-2020/Tap-20-So-4/20\_4\_9\_89-101.pdf