

ANALYZING AND CUSTOMER DATA LABELING IN CRM SYSTEM

Pham Thi Lien*, Nguyen Quang Hiep, Nguyen Thu Phuong

TNU – University of Information and Communication Technology

ARTICLE INFO	ABSTRACT
Received: 22/12/2021	Customer Relationship Management (CRM) systems have been studied and applied by many businesses. Currently in the world there are many free and paid CRM software providers in the world such as: Zoho CRM, Hubspot CRM, SugarCRM, Microsoft Dynamics CRM, etc. But in Vietnam there are thousands of medium and large enterprises. For these businesses, the use of such large CRM software will be ineffective and too redundant compared to the size of the business. The purpose of this paper is to create an application CRM system for small and medium enterprises based on analysis of data sources from e-commerce websites. We analyze and label data automatically based on collected raw data. Testing at EXP Thai Nguyen Service and Technology Development Co., Ltd. The system works effectively, and the results achieved are statistics of customer data according to the purposes of business use.
Revised: 31/5/2022	
Published: 31/5/2022	
KEYWORDS	
Customer relationship management system	
CRM	
Data labeling	
Data collection	
Data analysis	

PHÂN TÍCH VÀ GÁN NHÃN DỮ LIỆU KHÁCH HÀNG TRONG HỆ THỐNG CRM

Phạm Thị Liên*, Nguyễn Quang Hiệp, Nguyễn Thu Phương

Trường Đại học Công nghệ thông tin và Truyền thông - ĐHTN Thái Nguyên

THÔNG TIN BÀI BÁO	TÓM TẮT
Ngày nhận bài: 22/12/2021	Hệ thống quản lý quan hệ khách hàng (Customer Relationship Management-CRM) đã và đang được nhiều doanh nghiệp tìm hiểu và áp dụng. Hiện nay trên thế giới có rất nhiều đơn vị cung cấp phần mềm CRM miễn phí lẫn trả phí có thể kể tới như: Zoho CRM, Hubspot CRM, SugarCRM, Microsoft Dynamics Crm,... Nhưng tại Việt Nam có hàng ngàn doanh nghiệp vừa và nhỏ, đối với những doanh nghiệp này việc sử dụng các phần mềm CRM lớn như trên sẽ không mang lại hiệu quả và quá dư thừa so với quy mô doanh nghiệp. Mục đích của bài báo này là tạo ra một hệ thống CRM ứng dụng cho các doanh nghiệp vừa và nhỏ dựa trên phân tích các nguồn dữ liệu từ các website thương mại điện tử. Chúng tôi phân tích và gán nhãn dữ liệu tự động dựa trên các dữ liệu thô đã thu thập được. Thử nghiệm tại Công ty TNHH dịch vụ và phát triển công nghệ EXP Thái Nguyên. Hệ thống hoạt động hiệu quả và Kết quả đạt được là thống kê dữ liệu khách hàng theo mục đích sử dụng của doanh nghiệp.
Ngày hoàn thiện: 31/5/2022	
Ngày đăng: 31/5/2022	
TỪ KHÓA	
Hệ thống quản lý quan hệ khách hàng	
CRM	
Gán nhãn dữ liệu	
Thu thập dữ liệu	
Phân tích dữ liệu	

DOI: <https://doi.org/10.34238/tnu-jst.5370>

* Corresponding author. Email: ptlien@ictu.edu.vn

1. Giới thiệu

CRM là hệ thống quản lý quan hệ khách hàng được rất nhiều doanh nghiệp ứng dụng vì tính năng ưu việt của nó. Dựa trên các tài liệu, chúng tôi khái niệm CRM như một tập hợp các hoạt động kinh doanh liên quan đến bốn lĩnh vực can thiệp khác nhau (bốn khía cạnh CRM: chiến lược, tổ chức, công nghệ và quản lý khách hàng) với mục tiêu cuối cùng là cải thiện mối quan hệ với khách hàng. Kumar và Reinartz [1] nhấn mạnh rằng mỗi khía cạnh CRM là cần thiết, mặc dù không đủ, để triển khai CRM. Để triển khai CRM thành công và có khả năng cạnh tranh, một công ty phải ngang bằng với đối thủ trên tất cả các khía cạnh CRM của mình và tạo điều kiện thuận lợi cho việc điều phối chúng một cách tích cực.

Các nghiên cứu [2] đã chỉ ra rằng CRM đang đóng một vai trò quan trọng trong việc tăng thị phần, nó nâng cao năng suất, tinh thần của nhân viên cao hơn đồng thời cải thiện kiến thức chuyên sâu về khách hàng và cũng như sự hài lòng của khách hàng cao hơn để cải thiện lòng trung thành của khách hàng, công ty cũng sẽ có thông tin rõ ràng rằng khách hàng của họ là gì, nhu cầu của họ là gì và điều gì sẽ khiến họ hài lòng hơn. Trong [3] đã chỉ ra sự yếu kém trong quản lý tiếp thị của doanh nghiệp đồng thời chứng minh sự thành công khi sử dụng hệ thống CRM.

Với những lợi ích mà hệ thống CRM mang lại mà đã có rất nhiều các nghiên cứu với nhiều công nghệ khác nhau để xây dựng hệ thống CRM đặc trưng cho từng ngành nghề với dữ liệu khách hàng khác nhau. Trong [4] các nhà nghiên cứu đã sử dụng bảng câu hỏi để thu thập dữ liệu. Đối với phân tích dữ liệu, các mối tương quan hai biến đã được áp dụng. Kết quả chỉ ra rằng môi trường tổ chức và CRM có mối quan hệ tích cực và đáng kể với môi trường tổ chức. Trong [5] đã đưa bốn giả thuyết đã được xây dựng và dữ liệu cần thiết được thu thập thông qua một bảng câu hỏi có cấu trúc. Dữ liệu được phân tích định lượng. Theo các mối quan hệ rút ra từ việc phân tích dữ liệu, các nhà nghiên cứu kết luận rằng yếu tố thứ nhất (công nghệ) có tác động tích cực và quan trọng đến việc triển khai CRM, trong khi ba yếu tố còn lại (nguồn nhân lực, kiến thức về CRM và quản lý tri thức) có một yếu tố không chặt chẽ nhưng tích cực trong việc triển khai CRM. Trong [6] đã phân tích tổng hợp dựa trên lý thuyết thống nhất về việc chấp nhận và sử dụng công nghệ (meta-UTAUT) để dự đoán ý định hành vi của người dùng tổ chức và hành vi sử dụng của họ đối với hệ thống quản lý quan hệ khách hàng (CRM) tích hợp trí tuệ nhân tạo (AI).

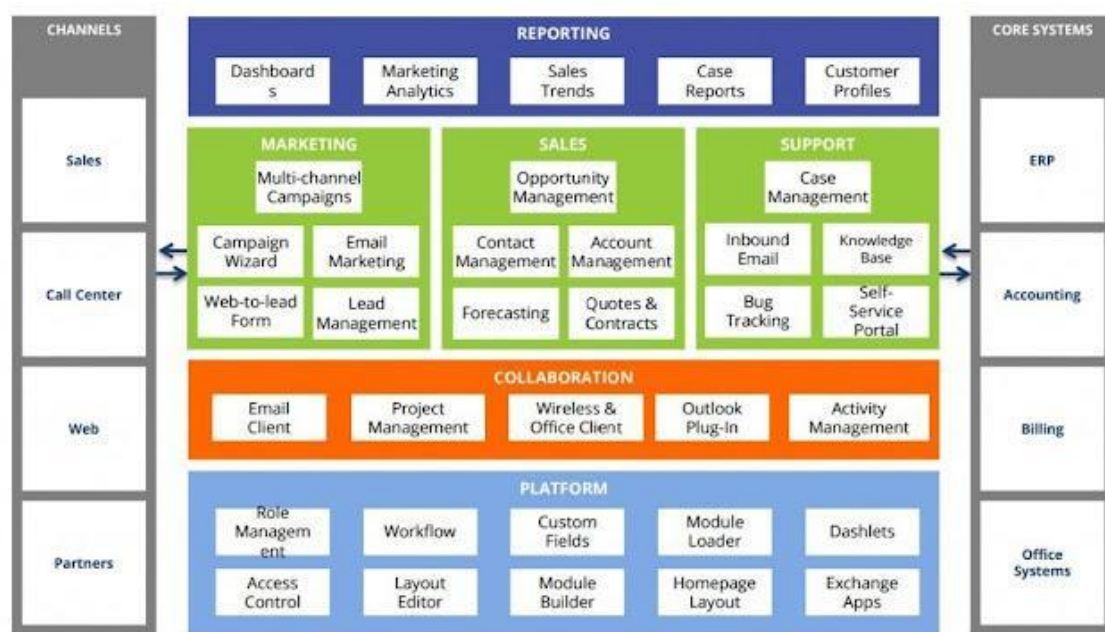
Đặc biệt trong [7], [8] đã chỉ ra tầm quan trọng của hệ thống CRM đối với các hệ thống kinh doanh nhỏ lẻ, chú trọng đến việc thiết kế mô hình CRM có kích thước và hiệu suất phù hợp với đặc điểm của doanh nghiệp.

Trong bài báo này chúng tôi cũng nghiên cứu và thực nghiệm mô hình CRM cho doanh nghiệp vừa và nhỏ của Việt Nam

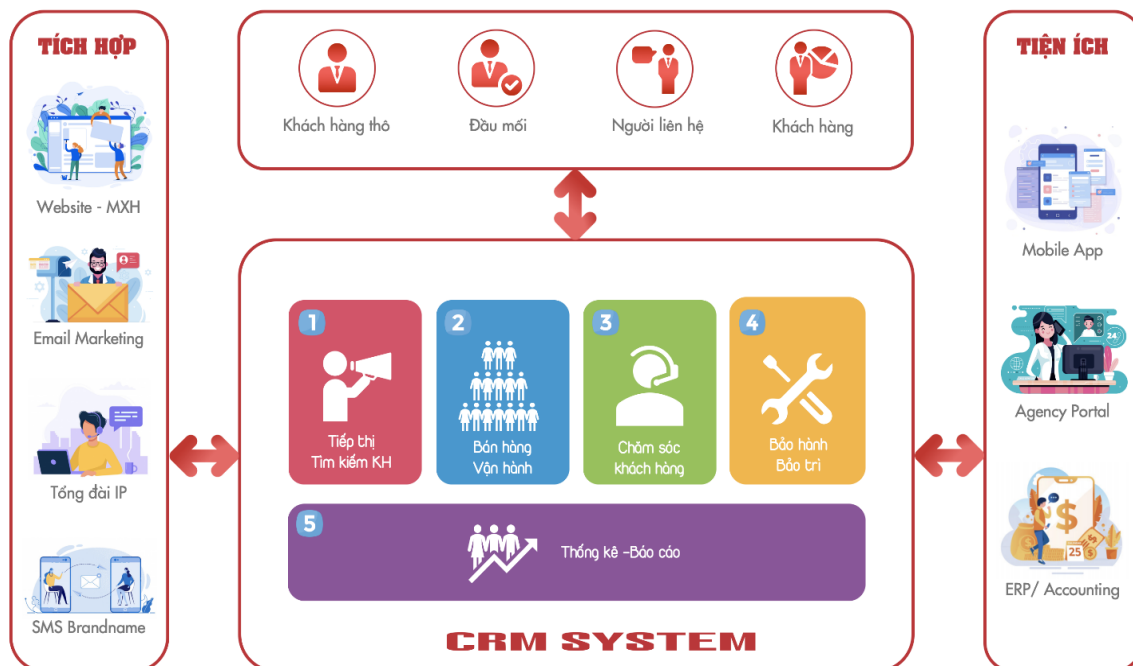
2. Mô hình CRM cho doanh nghiệp vừa và nhỏ của Việt Nam

Chúng tôi đã tiến hành khảo sát một số phần mềm CRM. Hình 1 là Kiến trúc tổng quan của SugarCRM [9]. SugarCRM là phần mềm CRM mã nguồn mở. Đây là một trong những phần mềm CRM lớn nhất thế giới. Sugar CRM chuyên về quản lý thực hiện các chương trình quảng bá; tiếp thị sản phẩm; quản lý chăm sóc khách hàng. Bên cạnh đó là thiết lập các chương trình hệ thống kinh doanh đặc thù. SugarCRM có thể tích hợp e-mail và thông tin lịch với Microsoft Outlook. Hình 2 đưa ra mô hình hệ thống CRM được thiết kế cho các doanh nghiệp Việt Nam với đầy đủ các chức năng như tiếp thị tìm kiếm khách hàng, Bán hàng vận hành, chăm sóc khách hàng, Bảo hành bảo trì, thống kê báo cáo. Đồng thời tích hợp nhiều hình thức tiếp cận với khách hàng.

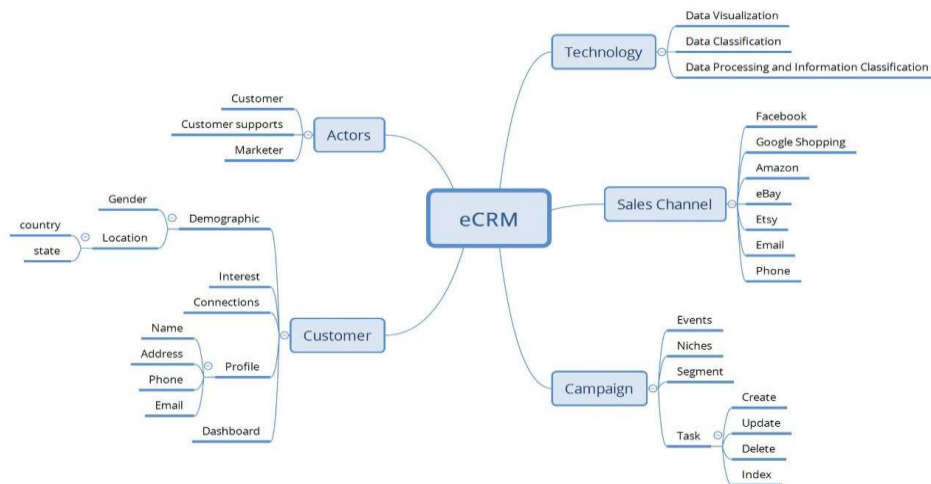
Qua khảo sát phân tích nhu cầu của khách hàng cũng như đặc điểm của các doanh nghiệp vừa và nhỏ ở Việt Nam. Với sự phát triển của công nghệ và mạng xã hội hiện nay, Việt Nam có hàng ngàn doanh nghiệp hoạt động trong lĩnh vực thương mại điện tử. Chúng tôi đã khảo sát và đưa ra mô hình cũng như lĩnh vực, nền tảng kinh doanh của doanh nghiệp thương mại điện tử (Hình 3).



Hình 1. Kiến trúc tổng quan của SugarCRM



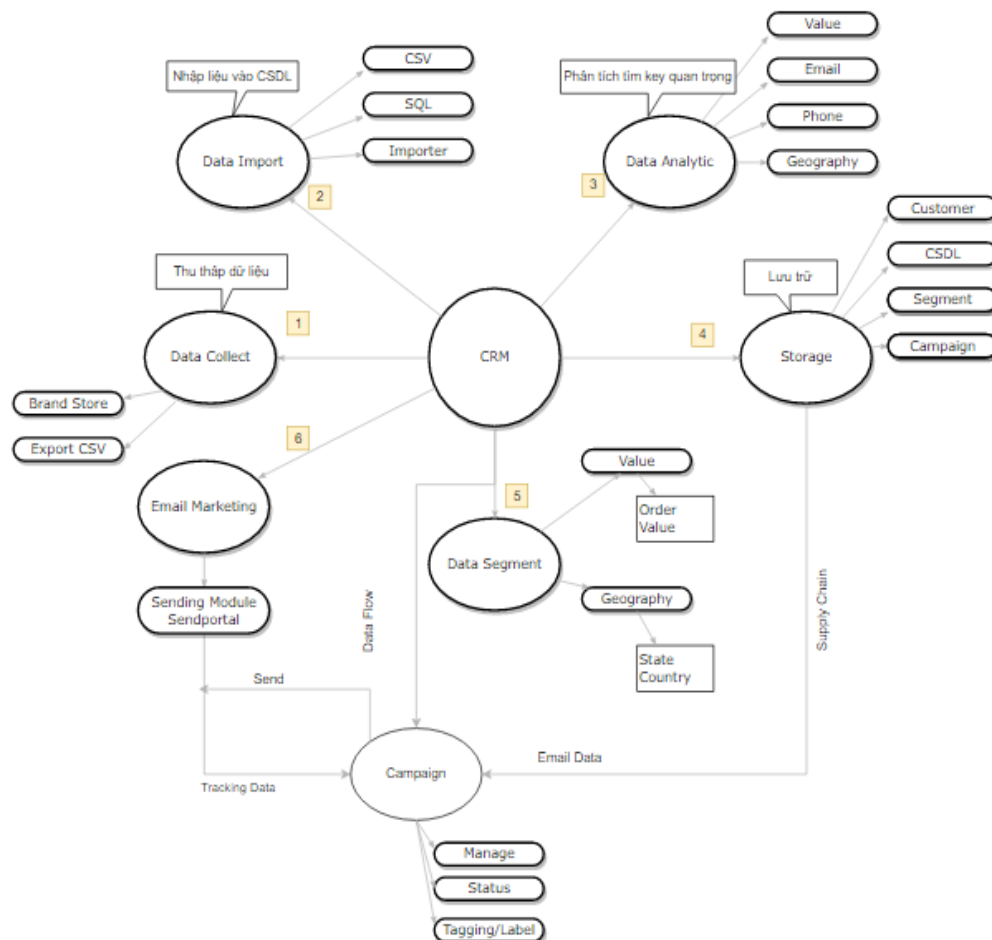
Hình 2. Hệ thống CRM của onlineCRM



Hình 3. Mô hình kinh doanh thương mại điện tử của doanh nghiệp vừa và nhỏ tại Việt Nam (kết quả của nghiên cứu năm 2021)

3. Quy trình xây dựng hệ thống CRM

Mô hình hệ thống CRM được mô tả trong hình 4.



Hình 4. Mô hình hệ thống CRM (kết quả của nghiên cứu năm 2021)

Thu thập dữ liệu (hình 5).

Name	Date modified	Type	Size
1. annobros.csv	8/19/2020 6:09 AM	Microsoft Excel C...	5,492 KB
2. angles.csv	8/19/2020 6:06 AM	Microsoft Excel C...	16,147 KB
3. bnapth.csv	8/19/2020 6:15 AM	Microsoft Excel C...	3,303 KB
4. jenged.csv	8/19/2020 6:09 AM	Microsoft Excel C...	20,339 KB
5. k.csv	8/19/2020 6:04 AM	Microsoft Excel C...	3,396 KB
6. memec.csv	8/19/2020 6:41 AM	Microsoft Excel C...	14 KB
7. PETERLNU2.csv	8/20/2020 4:01 AM	Microsoft Excel C...	242,041 KB
8. rdnatmap.csv	8/19/2020 5:43 AM	Microsoft Excel C...	328,148 KB
9. rdnatmap.csv	8/19/2020 5:50 AM	Microsoft Excel C...	32,870 KB
10. southbeach.csv	8/19/2020 5:06 AM	Microsoft Excel C...	13,183 KB
11. sportmansden.csv	8/19/2020 4:44 AM	Microsoft Excel C...	1,268 KB
12. sportmansden.csv	8/19/2020 4:31 AM	Microsoft Excel C...	5,787 KB
13. substatyevision.csv	8/19/2020 4:13 AM	Microsoft Excel C...	15,464 KB

Autosave	Search (Ctrl-C)	nguyen hiep	File	Home	Insert	Page Layout	Formulas	Data	Review	View	Help
----------	-----------------	-------------	------	------	--------	-------------	----------	------	--------	------	------

id	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		
1	CUS	ADDRESS	first_name	last_name	birth	mor	birth	day	birth	year	id	state	id	exp	id	type	ADDRESS1	ADDRESS2	CITY	STATE	ZIP	PHONE1	PHONE2	COUNTRY	username	password	email1	email2
131	261	1761	TODD	SUSAN																								
132			2541	TIMOTHY	SMITH																							
133	31	1501	CHRISTIA	SLOAD																								
134	256	1726	RICHARD	HOLLINGER																								
135	43	1533	BEN	LANE																								
136		2549	TONNY	KILGORE																								
137	54	1534	MICHELLE	BOWMAN																								
138	340	1839	BILL	HOLDSHEL																								
139	345	1819	CARLA	HOWELL																								
140		2550	AMY	HILL																								
141	48	1518	EMILY	VASQUEZ																								
142		2729	VICKI	MOTTER																								
143	263	1713	SUSAN	HOLMES																								
144		2551	KATHERIN	COMINSKY																								

Hình 5. Danh sách khách hàng, dữ liệu thô đầu vào (kết quả của nghiên cứu năm 2021)

Tại bước này dữ liệu được tổng hợp từ nhiều nguồn khác nhau chủ yếu từ các dữ liệu bán hàng của công ty trong thời gian trước đây. Các dữ liệu nằm ở nhiều mẫu khác nhau do được thu thập từ nhiều kênh:

- Kênh web store bán hàng trực tuyến nền tảng Shopify và Woocommerce.
- Kênh bán hàng trên eBay.
- Kênh Facebook Marketplace.
- Kênh Amazon.
- Kênh Etsy.

Các biểu mẫu được thu thập nằm ở các dạng data chính: MySQL database, csv, excel, txt. Các dữ liệu này sẽ được xử lý thô thông qua module import theo một chuẩn thông nhất và được build lại dưới dạng máy tính có thể đọc được.

Phân tích Key data (hình 6).

Email Change										Total Paid - Customer Value										Customer Labels										S T U V W X Y Z																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
Name	email1	email2	Financial	First	F	F	F	F	F	Current	Shipping	Subtotal	Shipping	id	Account	Discount	Account	Shipping	Created	Label	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	LineItem	Line

Hình 6. Phân tích Key data (kết quả của nghiên cứu năm 2021)

Sau khi dữ liệu được import và chuẩn hóa, các key metadata chính sẽ được xác lập để đánh giá giá trị của dữ liệu[10].

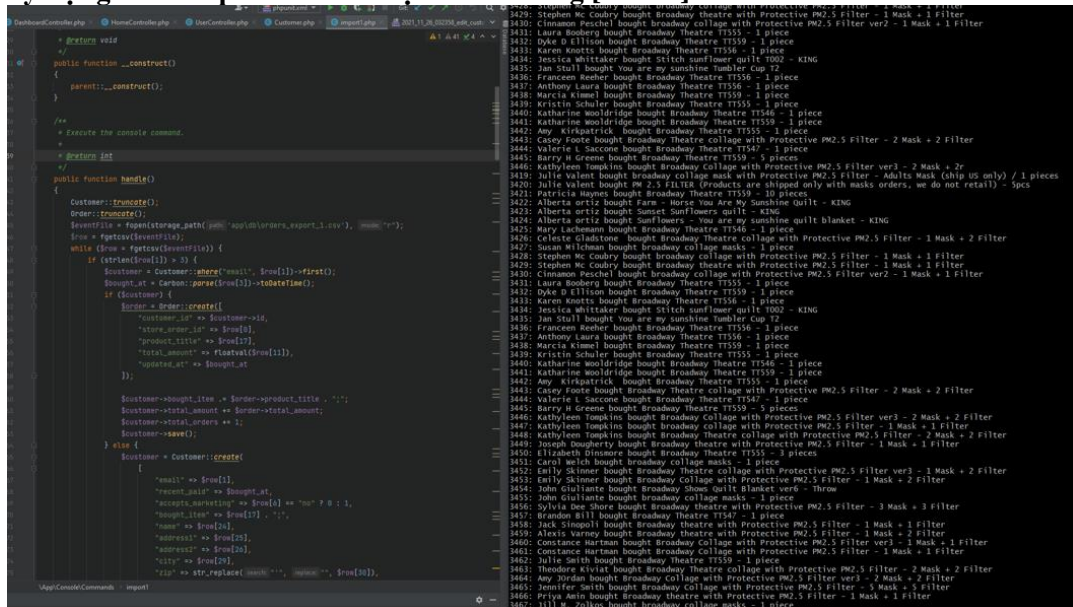
- Key dữ liệu liên quan tới số lần khách hàng mua sản phẩm. Dữ liệu này nếu có sẽ được tổ chức theo thời gian để nắm được thời gian cũng như tần suất của khách hàng sử dụng dịch vụ hoặc mua sản phẩm của công ty.

- Key dữ liệu liên quan tới tổng giá trị các đơn hàng mà khách hàng đã đặt. Dữ liệu này đánh giá rõ nhất về giá trị của một khách hàng. Giúp ta phân loại được từng lớp khách hàng và có phương án chăm sóc cần thiết.

- Key dữ liệu liên quan tới các sản phẩm khách hàng đã mua. Dữ liệu này giúp cho chúng ta có thể hiểu được nhiều hơn về sở thích cũng như những giá trị mà công ty mang lại cho khách hàng, được khách hàng tiếp nhận. Key dữ liệu này sẽ sử dụng làm key chính để gán nhãn cho khách hàng theo sở thích.

- Key dữ liệu liên quan tới nhân khẩu học. Là các dữ liệu có thể thu thập được về địa chỉ, giới tính, độ tuổi và các thuộc tính khác có thể thu thập được. Các key này sẽ được sử dụng trong các query đặc biệt của các chiến dịch chăm sóc khách hàng của hệ thống.

Xây dựng module phân tích dữ liệu khách hàng [hình 7]



Hình 7. Module phân tích dữ liệu khách hàng (kết quả của nghiên cứu năm 2021)

Module này được xây dựng nhằm đọc dữ liệu thô khách hàng (đã được quy chuẩn hóa) và thực hiện các phân tích cần thiết nhằm tìm ra các Key metadata đã được phân tích ở trên[11]. Sau đó đưa toàn bộ vào một bảng cơ sở dữ liệu khách hàng để xử lý trong các bước sau.

Import nhãn đúng để gán nhãn dữ liệu (hình 8).

Để có thể gán nhãn, phân loại khách hàng, việc hiểu và thống nhất một hệ thống label là cần thiết. Khách hàng từ đó sẽ được phân loại theo một danh sách các tiêu chí nhất định (Label). Hệ thống Label này được xây dựng dựa theo các từ khóa keyword đã được tổng hợp lại theo các niche hẹp của công ty khi triển khai các chiến dịch marketing. Nó luôn được đánh giá theo volume và thời gian (Google Analytic - Google Keyword Planner) [12], [13].

```

Importing Label : (5000) blizzard beach tickets
Importing Label : (500) disney mulan 2020
Importing Label : (5000) winnie the pooh and piglet
Importing Label : (5000) disney dollars
Importing Label : (5000) disney questions
Importing Label : (500) snow white and the seven dwarfs story
Importing Label : (50000) lightning mcqueen cars
Importing Label : (5000) cheap disney world tickets
Importing Label : (5000) winnie the pooh and friends
Importing Label : (5000) walt disney world packages
Importing Label : (5000) periwinkle tinkerbelle
Importing Label : (5000) homecoming disney springs
Importing Label : (5000) frozen book
Importing Label : (5000) hamilton on disney plus
Importing Label : (5000) lady in the tramp
Importing Label : (5000) beaches and cream
Importing Label : (50000) disney's soul
Importing Label : (5000) pirates and princesses
Importing Label : (5000) princess mulan
Importing Label : (5000) remy's ratatouille adventure
Importing Label : (50000) minnie and mickey mouse
Importing Label : (5000) lightning mcqueen characters
Importing Label : (500) film raya
Importing Label : (50000) disney's blizzard beach
Importing Label : (5000) disney maleficent
Importing Label : (5000) matterhorn disneyland
import success disney with 652 Label
Admin@Hiep-OfficePC MINGW64 /e/xampp/htdocs/ecrm (master)

```

Hình 8. Import nhãn đúng để gán nhãn dữ liệu (kết quả của nghiên cứu năm 2021)

Phân tích và gán nhãn dữ liệu tự động cho khách hàng (hình 9).

```

Family added Customer ID: 12,20,30,39,553,1654,2004,4228,5169,5283,5446,6157,6907,8004,8386,8434,35425,35579,35761,Process label sis
family added Customer ID: 19777,26154,28647,30967,31656,32393,32950,33099,34843,36270,36879,37157,37249,37684,37685,37916,43193,Process
label soulmate
Process label roommate
Process label partner
Process label nephew
Process label roomie
Process label pal
Family added Customer ID: 27115,27268,37223,Process label sister
Family added Customer ID: 19777,30967,Process label stepmom
Process label bff
Process label father
Family added Customer ID: 12104,18224,18243,18264,18317,18384,18549,22212,22272,22539,22541,22579,22634,22744,22958,22994,23013,23086,23
250,23797,23841,23905,23914,24108,24382,30311,38583,42206,42819,45850,46059,46274,46277,46286,46408,46415,46468,46490,46492,46498,46547,
46558,46904,46922,46956,47118,47349,47878,49504,50614,Process label wife
family added Customer ID: 55,61,63,69,71,253,35039,35463,37081,44692,45120,45171,45382,47586,Process label mummy
Process label companion
Process label sweetie
Process label bride
Family added Customer ID: 354,7064,20930,21140,21176,22281,22334,22833,24448,24455,24559,24785,26088,26114,26276,26316,26347,26389,26419,
26464,26484,26556,26563,26614,26634,26680,26684,26715,26776,26836,27150,27176,27178,27258,27291,27318,27343,27363,27364,27367,27376,274
28,27459,27479,27498,27520,27530,27535,27543,27551,27558,27566,27571,27581,27590,27593,27595,27597,27607,27610,27615,27637,27644,27648,2
7650,27665,27669,27673,27709,27719,27721,27724,27727,27728,27729,27740,27759,27760,27761,27786,27789,27790,27794,27808,27867,27886,27899,
27922,27931,27958,27971,27988,28022,28037,28042,28071,28097,28104,28126,28150,28151,28155,28160,28184,28352,28360,28399,28569,28636,287
73,28782,28808,28813,28823,28869,28910,29041,29087,29093,29104,29128,29158,29211,29244,29245,29255,29256,29305,29343,29375,29402,29463,2
9495,29522,29611,29637,29662,29696,29705,29726,29788,29815,29818,29830,30015,30068,30368,30411,30448,30523,30557,30571,30874,31049,31372,
31398,31623,32170,33055,33992,34400,38144,39152,39503,39733,39961,40176,40462,40930,41121,42767,46299,46625,47101,47846,48822,49468,498
53,Process label uncle
Family added Customer ID: 32381,Process label husband

```

Hình 9. Phân tích và gán nhãn dữ liệu tự động cho khách hàng (kết quả của nghiên cứu năm 2021)

Các bước gán nhãn được xây dựng như sau:

Bước 1. Xây dựng module Label Processor: Module này sẽ lấy một Label được chỉ định cùng với một dữ liệu khách hàng. Sau đó nó thực hiện so khớp các tiêu chí nhằm đưa ra được câu trả lời. Khách hàng có thuộc nhóm label đó hay không.

Bước 2. Quá trình thực hiện chạy. Lấy từng label trong danh sách các label đã được xây dựng và cho Label Processor chạy quét trên toàn bộ khách hàng.

Bước 3. Từng khách hàng ID sau khi được gán nhãn sẽ được lưu trữ lại trong cơ sở dữ liệu.

4. Kết quả thực nghiệm

Tất cả các thí nghiệm được thực hiện bằng hệ thống PC Win10 với CPU Intel Core i3 @ 3,00 GHz và RAM 2,00 GB. Hệ thống thông tin sử dụng ngôn ngữ lập trình PHP và một số ngôn ngữ khác như Nodejs, MySQL, kiến trúc MVC và Framework Laravel để lập trình các lớp mã.

Danh sách khách hàng có giá trị lớn nhất (hình 10).

Customer manage

Id	Name	Address	Email	TotalOrders	TotalAmount	Actions
8596	Patsy Hampton	UT US	hamptonfarming@gmail.com	4	779.70	 
11617	Monte Wilcox	CO US	montewilcox@hotmail.com	6	751.50	 
54	Laurie Ackerman	CA US	lzackerman@yahoo.com	26	694.11	 
11958	Darlene Wise	LA US	dolwise1@cox.com	2	636.55	 
9535	Jorge Munoz	TX US	tioperfumes07@gmail.com	11	597.50	 
12828	antonio diaz	NL MX	gil.rost3r1980@hotmail.com	2	510.38	 
180	Victoria Margulin	TX US	vickyjon@sbcglobal.net	19	493.35	 
8661	Sheri Flister	CA US	Flickaby7@aol.com	15	481.65	 

Hình 10. Danh sách khách hàng có giá trị lớn nhất (sản phẩm của nghiên cứu năm 2021)
Danh sách khách hàng mua nhiều hàng nhất (hình 11).

Id	Name	Address	Email	TotalOrders	TotalAmount	Actions
28216	Kimberly Moore	US	customer@gmail.com	112		 
24288	Dr. Sharen Barboza	US		45		 
229	Paul Connelly	MI US	dancerman2005@gmail.com	30	477.51	 
493	Veronica Alatorre	CA US	alator5@aol.com	28	422.85	 
54	Laurie Ackerman	CA US	lzackerman@yahoo.com	26	694.11	 
180	Victoria Margulin	TX US	vickyjon@sbcglobal.net	19	493.35	 
3861	Maria Martinez	CA US	mariavmartinez79@sbcglobal.net	15	190.30	 
32	Jennifer Keating	CT US	jenniferrkeating@yahoo.com	15	333.16	 
2603	Pricilla Perez	CA US	ppricilla55@yahoo.com	15	230.19	 

Hình 11. Danh sách khách hàng mua nhiều hàng nhất (sản phẩm của nghiên cứu năm 2021)

5. Kết luận

Bài báo đã đề cập đến một số mô hình CRM đang được sử dụng phổ biến ngày nay như sugerCRM hình 1, CRM SYSTEM hình 2. Cũng đã phân tích và đưa ra mô hình CRM hình 4 phù hợp với các doanh nghiệp thương mại điện tử vừa và nhỏ ở Việt Nam. Bài báo cũng đã phân tích và xây dựng thuật toán gán nhãn dữ liệu tự động cho dữ liệu dạng văn bản. Kết quả của quá trình phân loại dữ liệu ta sẽ thu được dữ liệu khách hàng theo mục đích sử dụng của từng doanh nghiệp. Quá trình thực nghiệm ở mục 4 đã chứng minh phạm vi ứng dụng của hệ thống CRM là phù hợp với các doanh nghiệp vừa và nhỏ.

Lời cảm ơn

Bài báo là sản phẩm của đề tài có mã số T2021-07-05, được tài trợ bởi kinh phí của trường Đại học Công nghệ Thông tin và Truyền thông – ĐHTN. Nhóm tác giả xin chân thành cảm ơn sự tài trợ của quý Nhà trường.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] V. Kumar and Werner Reinartz, *Customer Relationship Management: A Databased Approach*, John Wiley&Sons, 2006.
- [2] R. S. Hassan, A. Nawaz, M. N. Lashari, and F. Zafar, "Effect of Customer Relationship Management on Customer Satisfaction," *ScienceDirect*, vol. 23, pp. 563-567, 2015.
- [3] M. Kubina and V. Lendel, "Successful Application of Social CRM in The Company," *ScienceDirect*, vol. 23, pp. 1190-1194, 2015.
- [4] A. A. Aladwan and K. Mahmoud, "Measuring the Relationship between Organizational Climate and Customer Relationship Management," *Cihan University-Erbil Journal of Humanities and Social Sciences*, vol. 8, no. 2, pp. 17-28, December 2019.
- [5] Q. H. Al-Salami, I. A. Saadi, Z. T. Sawadi, and R. Kh. Saleh, "The Factors Affecting CRM Implementation at DHL Company in Baghdad, Iraq," *Cihan University-Erbil Journal of Humanities and Social Sciences*, vol. 3, pp. 43-51, 2019.
- [6] S. Chatterjee, N. P. Rana, S. Khorana, P. Mikalef, and A. Sharma, "Assessing Organizational Users' Intentions and Behavior to AI Integrated CRM Systems: a Meta-UTAUT Approach," *Springer Nature*, vol. 8, no. 2, pp. 17-28, August 2021.
- [7] J. Mathew, "Impact of Customer Relationship Management Initiatives in Indian Retail Scenario: A Review," *Pacific Business Review International*, vol. 10, no. 1, July 2017. [Online serial]. Available: <https://www.researchgate.net/publication/342698754/>. [Accessed January, 2021].
- [8] A. A. Mohammed and B. bin Rashid, "Customer Relationship Management (CRM) in Hotel Industry: A framework Proposal on the Relationship among CRM Dimensions, Marketing Capabilities and Hotel Performance," *International Review of Management and Marketing*, vol. 2, no. 4, pp. 220-230 2012.
- [9] PC World, "Architecture overview of CRM system". Available: <https://Onlinecrm.Vn/>. [Accessed January, 2021].
- [10] K. V. Nguyen, D. V. Nguyen, A. G. T. Nguyen, and N. L. T. Nguyen, "A Vietnamese Dataset For Evaluating Machine Reading Comprehension," *Barcelona, Spain*, pp. 2595-2605, December, 2020. [Online Serial]. Available: <https://Aclanthology.Org/2020.Coling-Main.233.Pdf>. [Accessed January, 2021].
- [11] T. M. H. Dinh, "Application of Data Classification Techniques for Ethnic Minority Language Data Mining for Natural Language Processing, Proceedings of the CITA National Scientific Conference 2017," *CNTT And Applications In The Fields*, pp. 237-244, 2017. [Online Serial]. Available: <http://Thuvien.Vku.Udn.Vn/Bitstream/>. [Accessed January, 2021].
- [12] S. T. Luu, K. V. Nguyen, A. G. T. Nguyen, and N. L. T. Nguyen, "An Experimental Study Of Deep Neural Network Models For Vietnamese Multiple-Choice Reading Comprehension," vol. 1, August 2020. [Online Serial]. Available: <https://Researchgate.Net/Publication/343785987/>. [Accessed January, 2021].
- [13] C. D. Tran and N. K. Pham, "Text Classification With Machine Learning Support Vectors And Decision Trees," *Science magazine*, vol. 21a, pp. 52-63, 2012. [Online Serial]. Available: <https://Sj.Ctu.Edu.Vn/Qt/Docgia/Tacgia-1133/Baibao-5654.html/>. [Accessed January, 2021].