



Faculty of Computer Science and Information Technology

**PRODUCT AND OFFER CATEGORIZATION SYSTEM  
(POCaS) FOR COMPARISON SHOPPING ENGINE WEBSITES**

Sharvin Rathanasamy (71539)

Bachelor of Computer Science with Honours

(Software Engineering)

2023

Supervised By

Dr. Mohamad Imran bin Hj.Bandan

**PRODUCT AND OFFER CATEGORIZATION SYSTEM  
(POCaS) FOR COMPARISON SHOPPING ENGINE WEBSITES**

SHARVIN RATHANASAMY

This project is submitted in partial fulfilment of  
the requirements for the degree of Bachelor of Computer Science with Honours  
(Software Engineering)

Faculty of Computer Science and Information Technology

UNIVERSITI MALAYSIA SARAWAK

2023

**SISTEM KATEGORI PRODUK DAN TAWARAN (POCaS) UNTUK LAMAN WEB  
ENJIN BELAH PERBANDINGAN**

**SHARVIN RATHANASAMY**

Projek ini merupakan salah satu keperluan untuk  
Ijazah Sarjana Muda Sains Komputer dengan Kepujian (Kejuruteraan Perisian)

Fakulti Sains Komputer dan Teknologi Maklumat

UNIVERSITI MALAYSIA SARAWAK

2023

UNIVERSITI MALAYSIA SARAWAK

THESIS STATUS ENDORSEMENT FORM

TITLE PRODUCT AND OFFER CATEGORIZATION SYSTEM (POCAS)  
FOR COMPARISON SHOPPING ENGINE WEBSITES

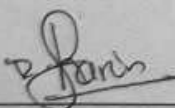
ACADEMIC SESSION: 2022/2023

SHARVIN A/L RATHANASAMY  
(CAPITAL LETTERS)

hereby agree that this Thesis\* shall be kept at the Centre for Academic Information Services, Universiti Malaysia Sarawak, subject to the following terms and conditions:

1. The Thesis is solely owned by Universiti Malaysia Sarawak
2. The Centre for Academic Information Services is given full rights to produce copies for educational purposes only
3. The Centre for Academic Information Services is given full rights to do digitization in order to develop local content database
4. The Centre for Academic Information Services is given full rights to produce copies of this Thesis as part of its exchange item program between Higher Learning Institutions [ or for the purpose of interlibrary loan between HLI ]
5. \*\* Please tick ( ✓ )

- CONFIDENTIAL (Contains classified information bounded by the OFFICIAL SECRETS ACT 1972)
- RESTRICTED (Contains restricted information as dictated by the body or organization where the research was conducted)
- UNRESTRICTED



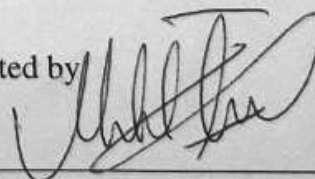
(AUTHOR'S SIGNATURE)

Permanent Address

31, LORONG BELIMBING 4, TAMAN  
BELIMBING, 14000 BUKIT MERTAJAW,  
PULAU PINANG

Date: 25/07/2023

Validated by



(SUPERVISOR'S SIGNATURE)

**Ts. Dr Mohamad Imran bin Hj Bandan**  
Senior Lecturer  
Faculty of Computer Science and Information  
Technology  
Universiti Malaysia Sarawak

Date: 25/07/2023

Note \* Thesis refers to PhD, Master, and Bachelor Degree

\*\* For Confidential or Restricted materials, please attach relevant documents from relevant organizations / authorities

## DECLARATION

I hereby declare that this research, inclusive with all related content is none other than that of my own work, with consideration of the exception of research-based information which were adapted and extracted from relevant source, which have evidently been stated respectively.

Signed,

A handwritten signature in black ink, appearing to read 'Sharvin', written over a horizontal line.

Sharvin Rathanasamy

71539

Faculty of Computer Science and Information Technology

University Malaysia Sarawak

1<sup>st</sup> July 2023

## **ACKNOWLEDGEMENT**

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Mohamad Imran bin Bandan, for their continuous support, advice, and guidance throughout my Final Year Project 2 and provided me with resources as well as essential information that was needed to complete my project task successfully. Next, I would like to express my appreciation to my Madam Izzatul Nabila bt Sarbini, who helped me provided insightful advice and opinions, which enabled and inspired me to complete my final year project and very grateful that she provided valuable suggestions for the betterment of the project, which we greatly appreciate for doing so.

Furthermore, I would want to express my gratitude to University Malaysia Sarawak (UNIMAS) and Faculty of Computer Science and Information Technology (FCSIT) for the platform provided to complete my Final Year Project. This platform allowed me test out my current skillset and knowledge as well as exposed to new valuable knowledge and technology and without their support, neither the artistic nor intellectual growth that is evident now would have been possible. Finally, yet importantly, my utmost gratitude towards my family for their unconditional support during the completion of my Final Year Project. I am also grateful to my friend Krishna Raj who played an important role when it came down to giving motivation and supported me along the way and making himself available at every stages.

## ABSTRACT

Ecommerce plays an important role in modern-day life. During the COVID-19 pandemic, the use of e-commerce platforms increased. There are billions of products that have been sold on this platform. Comparing prices across the e-commerce platform has been a difficult task. The Comparison-Shopping Engine (CSE) websites are tools to compare pricing, shipping options, and service provided for each product or service from multiple retailers from various e-commerce platforms to identify the best offer or deal for the specific product or service. Categorization produces a well-organized product catalogue which will reduce users (customers) burden to navigate through the catalogue to find the desired product. Moreover, it also reduces the workload of the teams (organization which build and maintain the site) itself. Categorizing the product catalogue properly helps the team to monitor the traffic towards the product, generate reports, decision-making about the products, and help to understand the users need better which improve the productivity of the organization. Most importantly, categorization helps to identify and push the pages with high value in Google search result. The Product and Offer Categorization System (POCaS) is a categorization system that's developed to categorise the products and offers of comparison-shopping engines. POCaS is a manual categorization system where the categorization is done manually by the categorization team. This system will be used by the categorization team to set rules for each category by defining the conditions for reaching each category and subcategory. Several works of literature were selected and reviewed to get a better insight into the proposed system and applied it in the development of the system. A web-based system is developed and named as POCaS. Based on the UML Diagrams and methodology, all the functionalities have been created and applied in the system. Moreover, functional and usability testing are performed to ensure the system meets the criteria and functions satisfactorily. The plan and tests are used to enhance the system further.

## **ABSTRAK**

E-dagang memainkan peranan penting dalam kehidupan moden. Semasa pandemik COVID-19, penggunaan platform e-dagang meningkat. Terdapat berbilion produk yang telah dijual di platform ini. Membandingkan harga merentas platform e-dagang telah menjadi tugas yang sukar. Comparison-Shopping Engine (CSE) ialah alat untuk membandingkan harga, pilihan penghantaran dan perkhidmatan yang disediakan untuk setiap produk atau perkhidmatan daripada berbilang peruncit daripada pelbagai platform e-dagang untuk mengenal pasti tawaran atau tawaran terbaik untuk produk atau perkhidmatan tertentu. Pengkategorian menghasilkan katalog produk yang teratur yang akan mengurangkan beban pengguna (pelanggan) untuk menavigasi melalui katalog untuk mencari produk yang dikehendaki. Selain itu, ia juga mengurangkan beban kerja pasukan (organisasi yang membina dan menyelenggara tapak) itu sendiri. Mengkategorikan katalog produk dengan betul membantu pasukan memantau trafik ke arah produk, menjana laporan, membuat keputusan tentang produk dan membantu memahami keperluan pengguna dengan lebih baik yang meningkatkan produktiviti organisasi. Paling penting, pengkategorian membantu mengenal pasti dan menolak halaman yang mempunyai nilai tinggi dalam hasil carian Google. Sistem Pengkategorian Produk dan Tawaran (POCaS) ialah sistem pengkategorian yang dibangunkan untuk mengkategorikan produk dan tawaran enjin beli-belah perbandingan. POCaS ialah sistem pengkategorian manual di mana pengkategorian dilakukan secara manual oleh pasukan pengkategorian. Sistem ini akan digunakan oleh pasukan pengkategorian untuk menetapkan peraturan bagi setiap kategori dengan menentukan syarat untuk mencapai setiap kategori dan subkategori. Sistem berasaskan web dibangunkan dan dinamakan sebagai POCaS. Berdasarkan Rajah dan metodologi UML, semua fungsi telah dicipta dan digunakan dalam sistem. Selain itu, ujian kefungsiian dan kebolehgunaan dilakukan untuk memastikan sistem memenuhi kriteria dan fungsi dengan memuaskan. Pelan dan ujian digunakan untuk meningkatkan lagi sistem.



## Table of Contents

DECLARATION .....	4
ACKNOWLEDGEMENT .....	5
ABSTRACT.....	6
ABSTRAK.....	7
1.1 Introduction.....	16
1.2 Problem Statement.....	18
1.3 Scope.....	19
1.3.1 System .....	19
1.3.2 Users.....	19
1.4 Objective.....	20
1.5 Project Significance .....	21
1.6 Project Schedule.....	22
1.7 Expected Outcome .....	23
2.1 Introduction.....	24
2.1.1 Definition of categorization. ....	24
2.2 Existing System Review .....	25
2.2.1 Rules Exchange System (REX.....	25
2.2.1.1 Strength and weaknesses. ....	28
2.2.2 Pumice.AI.....	29
2.2.2.1 Strength and weaknesses. ....	32
2.2.3 Datafeedwatch.....	33
2.2.3.1 Strength and weaknesses. ....	35
2.3 Discussion.....	36
2.4 Summary.....	38
3.1 Introduction.....	39

3.2 Methodology .....	39
3.2.1 Pre-phase .....	40
3.2.2 Phase 1 Feasibility.....	41
3.2.2.1 Interview .....	42
3.2.2.2 Requirement analysis.....	44
3.2.3 Phase 2: Foundation .....	45
3.2.3.1 UML Use case. ....	45
3.2.3.2 Use case specifications. ....	47
3.2.3.3 Sequence Diagram.....	55
3.2.3.4 Activity diagram. ....	62
3.2.3.5 Entity Relationship Diagram .....	64
3.2.3.6 Class diagram .....	65
3.2.3.6 System’s non-functional prototype.....	66
3.2.4 Phase 3: Evolutionary Development .....	70
3.2.5 Phase 4: Deployment.....	71
3.2.6 Post-Project .....	71
3.3 Summary .....	72
4.1 Introduction.....	73
4.2 Softwares and Tools Used .....	73
4.2.1 XAMPP .....	74
4.2.2 Visual Studio Code.....	75
4.2.3 000webhost.....	76
4.3 POCaS Implementation .....	77
4.3.1 Login Function .....	77
4.3.2 Manage rules for category.....	78
4.3.3 Manage categories.....	85

4.3.4 View Product.....	93
4.3.5 Upload taxonomy and Printing Taxonomy .....	95
4.4 Functional Testing .....	97
4.4.1 Login & Logout.....	97
4.4.2 Mange rules for categories .....	99
4.4.3 Manage categories.....	105
4.4.4 Upload Taxonomy.....	111
4.4.5 Print Taxonomy.....	112
4.4.6 View Product.....	113
4.5 User testing .....	114
4.6 Summary .....	118
5.1 Introduction.....	119
5.2 Achievements in relation to Project Objectives.....	119
5.3 Future Works .....	120
5.4 Conclusion .....	121
Reference .....	122
Appendix.....	124
<b>Interview form 1 .....</b>	<b>124</b>
<b>Interview form 2.....</b>	<b>127</b>
<b>User Testing .....</b>	<b>130</b>
<b>POCaS Link: <a href="https://pocas-unimas.000webhostapp.com/index.html">https://pocas-unimas.000webhostapp.com/index.html</a>.....</b>	<b>133</b>

## **List of Figures.**

Figure 1.1: Table of Task in Gantt Chart.	22
Figure 1.2: Gantt Chart of POCaS.	22
Figure 2.1: Flow of REX	25
Figure 2.2: User interface of REX	26
Figure 2.3: Flow of Pumice.ai	29
Figure 2.4: Dashboard of Pumice.ai (Categorization Tab)	31
Figure 2.5: Dashboard of Pumice.ai (Uploading Taxonomy and product manual injection)	31
Figure 2.6: Flow of Datafeedwatch	33
Figure 2.7: Parent/high-level category Datafeedwatch	34
Figure 2.8: Rules creation for further categories Datafeedwatch	34
Figure 3.1: The DSDM process	40
Figure 3.2 Usecase diagram of POCaS	46
Figure 3.3 Login usecase of POCaS.	55
Figure 3.4 Logout usecase of POCaS.	56
Figure 3.5 Manage rules for categories usecase of POCaS.	57
Figure 3.6 Manage categories usecase of POCaS.	58
Figure 3.7 Upload Taxonomy usecase of POCaS.	59
Figure 3.8 Print Taxonomy usecase of POCaS.	60
Figure 3.9 View product usecase of POCaS.	61
Figure 3.10 Activity diagram for add category, upload taxonomy and upload feed	62
Figure 3.11 Activity diagram for managing category and managing rules	63
Figure 3.12 ERD for POCaS	64

Figure 3.13 class diagram for POCaS	65
Figure 3.14 Login page of POCaS	66
Figure 3.15 Homepage of POCaS	67
Figure 3.16 Homepage of POCaS (rules)	67
Figure 3.17 Homepage of POCaS (Category details)	68
Figure 3.18 Taxonomy of POCaS	69
Figure 3.19 Product fees page of POCaS	58
Figure 4.1 XAMPP Local Host	74
Figure 4.2 POCaS programming in VS Code	75
Figure 4.3 Web hosting in 000webhost	76
Figure 4.4 Login Function code snippet	77
Figure 4.5 Login page	77
Figure 4.6 Category Page – Rules Table	78
Figure 4.7 Sub-category Page – Rules Table	78
Figure 4.8 Category rules display function.	79
Figure 4.9 Subcategory rules display function.	79
Figure 4.10 Add Rules-Category code snippet.	80
Figure 4.11 Add Rules-Subcategory code	80
Figure 4.12 Delete Rules-Category code	81
Figure 4.13 Delete Rules-Subcategory code	81
Figure 4.14 Implement Rules-Category code	82
Figure 4.15 Implement Rules-Subcategory	82
Figure 4.16 Reset Rules-Category code	83
Figure 4.17 Reset Rules-Subcategory code snippet.	83

Figure 4.18 Dashboard Page	85
Figure 4.19 Category Page – Category Details Table & Subcategory button	85
Figure 4.20 Subcategory Page – Subcategory Details Table	85
Figure 4.21 Get category and subcategory code snippet.	86
Figure 4.22 Get category details.	86
Figure 4.23 Get subcategory details.	87
Figure 4.24 Add category code snippet.	88
Figure 4.25 Add subcategory code snippet.	88
Figure 4.26 Add Category Model.	88
Figure 4.27 Add Subcategory	88
Figure 4.28 Edit Category code snippet.	90
Figure 4.29 Edit subcategory code snippet.	90
Figure 4.30 Edit Category in table.	91
Figure 4.31 Edit subcategory in table.	91
Figure 4.32 Delete Category code snippet.	92
Figure 4.33 Delete subcategory code snippet.	92
Figure 4.34 View product – category code snippet.	93
Figure 4.35 View product – subcategory code snippet.	93
Figure 4.36 View product panel	94
Figure 4.37 Upload taxonomy function code snippet.	95
Figure 4.38 Taxonomy Page	95
Figure 4.39 Print taxonomy function code snippet.	96

## **List of Tables**

Table 2.1 Functionalities of similar systems and proposed systems.	37
Table 3.1 Requirements of POCaS.	44
Table 3.2 Login usecase of POCaS.	47
Table 3.3 Logout usecase of POCaS.	48
Table 3.4 Manage rules for categories usecase of POCaS.	49
Table 3.5 Manage categories usecase of POCaS.	51
Table 3.6 Upload Taxonomy usecase of POCaS.	52
Table 3.7 Print Taxonomy usecase of POCaS.	53
Table 3.8 View Product usecase of POCaS.	54
Table 4.1 TC_F1.1	97
Table 4.2 TC_F2.1	98
Table 4.3 TC_F3.1	99
Table 4.4 TC_F3.2	100
Table 4.5 TC_F3.3	101
Table 4.6 TC_F3.4	102
Table 4.7 TC_F3.5	103
Table 4.8 TC_F3.6	104
Table 4.9 TC_F4.1	105
Table 4.10 TC_F4.2	106
Table 4.11 TC_F4.3	107
Table 4.12 TC_F4.4	108
Table 4.13 TC_F4.5	109
Table 4.14 TC_F4.6	110
Table 4.15 TC_F5.1	111
Table 4.16 TC_F6.1	112

Table 4.17 TC_F7.1	113
Table 4.18 Summary of user testing	116
Table 4.19 Summary of user testing	119



## **CHAPTER 1: INTRODUCTION**

### **1.1 Introduction**

The last two decades have seen the growth of e-commerce into new industries, markets, and product categories. E-commerce, often known as electronic commerce, is the act of purchasing and reselling products or services online. It includes a wide range of information, systems, and tools for online buyers and sellers, such as encrypted online payment systems and mobile shopping. (“Ecommerce 101: The History and Future of Online Shopping”, n.d.) Additionally, e-commerce platforms are the system, tool, or software where the purchasing and selling of product or services occurs over internet. It is the platform where the seller and consumer play their roles. Some examples of E-commerce platforms are Shopee, Lazada, eBay and Yubeli.

On the other hand, Comparison shopping engines (CSEs) are tools to compare pricing, shipping options and service provided for each product or service from multiple retailers from various e-commerce platforms to identify the best offer or deal for the specific product or services. CSE helps to find the best offer for iPhone 14 across the e-commerce platform such as Shopee and Lazada to find the best deal by a retailer. iPrice is one of the successful CSE in Malaysia which helps its users to find the best offer across the deals offered by its partners.

For both CSE websites and e-commerce platform, product categorization plays an essential part. Categorization produces a well-organized product catalogue which will reduce users (customers) burden to navigate through the catalogue to find the desired product. Moreover, it also reduces the workload of the teams (organization which build and maintain the site) itself. Categorizing the product catalog properly helps the team to

monitor the traffic towards the product, generate reports, decision-making about the products, and help to understand the users need better which improve the productivity of the organization. Most importantly, categorization helps to identify and push the pages with high value in Google search result. This can be done by prioritizing categories or pages which brings value to the organization and allow the Google bot to crawl on those sites and push the Google Indexing. (Gibbons, 2021)

Product catalogue or offer catalogue (for CSE websites) categorization can be done through a backend system. This backend system, used to set rules for categories and subcategories for the product or offer catalogue. Usually, the categorization system will be used by the categorization teams where their responsibility is to manage the product catalogue. Moreover, this system is also used to manage the products and offers in the product catalogue.

In a nutshell, a proper categorization system assists the categorization team to organize and manage the products and product catalogue. A good practice of categorization decreases the users' cognitive load to interact with the e-commerce platform or CSE websites, thus increase its productivity.

## 1.2 Problem Statement

A well-organized product catalogue helps the customers navigate through the ecommerce or CSE (comparison shopping engine) websites without getting lost in the product pool. Categorization also makes the Google bot push the website to a higher place in the Google search console (GSC). On CSE websites, the navigation through the product catalogue is more complicated compared to e-commerce websites. This is caused by the large number of offers and products in the uncategorized catalogue. The relationship between a product and offers is one-to-many and one-to-one, where one product can have many offers, but one offer can only refer to one product. So, if the offers and products on a CSE website aren't well put together, it's easy for the user to get lost and confused.

On the other hand, hard coding the categorization rules is not an optimal solution. The rules for each category, and subcategory are different, and the size of the product and offer pool is large. The loading time of the website will be affected if the products and offers are categorised through hard coding every time the request is made.

Additionally, a dedicated team for categorization is needed to manage the overall product catalogue and details of products and offers. This team consists of members from various educational backgrounds. Therefore, hardcoding the rules is not an optimal solution. Also, the team should be able to see the offers that follow the rules that were made to avoid strange results.

For the categorization team to be able to manage the product and offer catalogue, there needs to be a back-end categorization system with an easy-to-use interface.

## **1.3 Scope**

### **1.3.1 System**

Product and Offer Categorization System (POCaS) used to set rules for categorization for product and offer catalogue. It helps the categorization team to manage the product and offer catalogue where they can create, edit and delete the rules for categorization. Additionally, the team can also manage the details of products and offers. The user can add, edit, delete, and view the details of the categories and subcategories.

This system also provides a visualization of the products and offers that adhere to the rules created for verification.

### **1.3.2 Users**

POCaS is dedicated to the categorization team where they are responsible for product catalogue organization and management. This team consists of members from various educational backgrounds. Therefore, categorization is difficult to do through hardcoding. There should be a dedicated system with Graphical User Interface for the categorization team to manage, organize and maintain the product catalogue.

## **1.4 Objective**

- To develop a backend tool for the categorization team to categorize, manage, and organise products and offers catalogues in comparison shopping engine websites.
- To develop a tool to create, edit, and delete rules for product and offer categories, and subcategories in the product catalogue of Comparison-Shopping Engines websites.
- To test functional and user experience requirements of the categorization system.

## **1.5 Project Significance**

Categorisation of offers and products are very crucial for CSE websites. With proper backend system for categorization, the product and offer catalogue can be managed well. Moreover, it also helps to capture the odd product pages which affects the ranking in Google Search Index. POCaS provides an efficient and convenient way to manage and organise the product and offer catalogue in CSE websites. Furthermore, POCaS provides a User-friendly Graphical User Interface for the Categorization team where hardcoding not required. POCaS help the user to visualize the product categorization which adhere to it rules that been created. In conclusion, POCaS improves overall categorization for CSE websites, thus reduces the cognitive loads of the customer visit the site.

## 1.6 Project Schedule



















	Task Mode	Task Name	Duration	Start	Finish
		Product and Offer Categorization System	156 days	Mon 9/5/22	Mon 4/10/23
		Identifying Project and Lecturer	25 days	Mon 9/5/22	Fri 10/7/22
		Finding Lecturer	2 days	Mon 9/5/22	Tue 9/6/22
		Finding Topic	3 days	Sun 9/25/22	Tue 9/27/22
		Preparation of proposal	8 days	Wed 9/28/22	Fri 10/7/22
		Chapter 1: Introduction	6 days	Mon 10/10/22	Mon 10/17/22
		Chapter 2: literature Review	15 days	Tue 10/18/22	Mon 11/7/22
		Chapter 3: System Requirement & Design	11 days	Tue 11/8/22	Tue 11/22/22
		Final FYP 1 report	5 days	Wed 11/23/22	Tue 11/29/22
		Chapter 4: Implementation & Testing	67 days	Fri 12/16/22	Mon 3/20/23
		Chapter 5: Conclusion	10 days	Tue 3/21/23	Mon 4/3/23
		Final project Presentation	5 days	Tue 4/4/23	Mon 4/10/23

Figure 1.1: Table of Task in Gantt Chart.

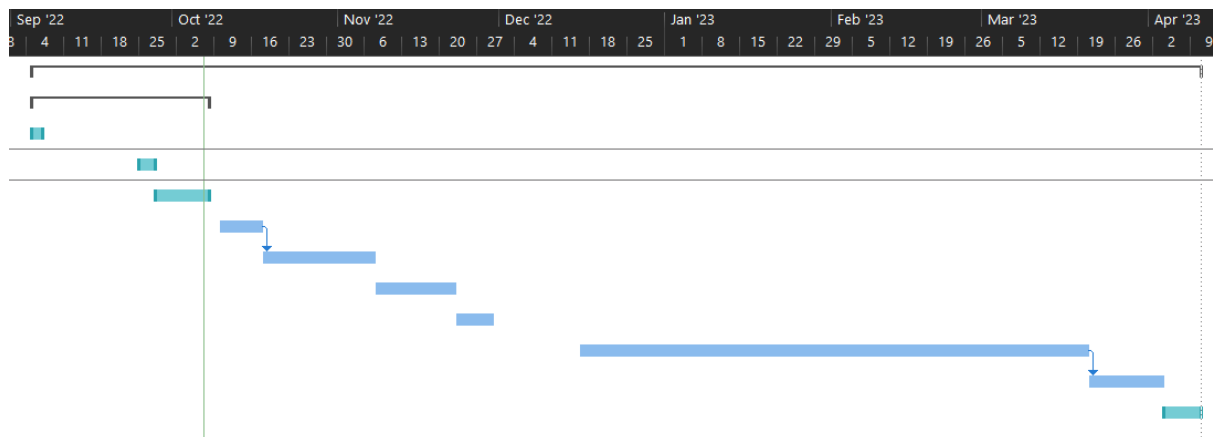


Figure 1.2: Gantt Chart of POCaS.

## **1.7 Expected Outcome**

A web-based backend system will be outcome of the project for the Categorization team of a CSE website organisation. This team will be able to manage and organize their product catalogue by setting the rules for categories and subcategories for the products and offers. The user also able to get a preview of the products and offers that adhere to the rule that has been created or modify. Thus, this system helps to create a well organised and maintained product and offer catalogue in a CSE website.