

**The evaluation of the daily profits of the group of cosmetics Sephora branches by using distributed systems the K-means algorithm and WEKA visualization**

**\* Corresponding author: \*Llahm Omar Faraj Ben Dalla<sup>1</sup>, Entesar M.eljali <sup>2</sup>,  
Almhdie Aboubaker Ahmed Agila**

\*Software Engineering Department<sup>1</sup>, Computer Engineering Department<sup>2</sup>, Computer Engineering Department<sup>3</sup>

\*College of Technical Science- Sebha<sup>1</sup>, University of Gharyan<sup>2</sup>, College of Technical Science- Sebha<sup>3</sup>

**\*[llahmomarfaraj77@gmail.com](mailto:llahmomarfaraj77@gmail.com)<sup>1</sup>& [entisar.eljali@gmail.com](mailto:entisar.eljali@gmail.com)<sup>2</sup>,  
[almhdie172@gmail.com](mailto:almhdie172@gmail.com)<sup>3</sup>**

**\*[llahmomarfaraj77@ctss.edu.ly](mailto:llahmomarfaraj77@ctss.edu.ly)<sup>1</sup>&[tarik.ahmad@gu.edu.ly](mailto:tarik.ahmad@gu.edu.ly)<sup>2</sup>&  
[almhdie@ctss.edu.ly](mailto:almhdie@ctss.edu.ly)<sup>3</sup>**

\* Orcid.ID: 0009-0008-7624-7567<sup>1</sup>& Orcid.ID:0009-0003-3955-5695<sup>2</sup>

**Abstract**

This study delves into the evaluation of daily profits within the network of Sephora cosmetics branches, employing distributed systems, the K-means algorithm, and WEKA visualization techniques. The objective is to analyze the profitability patterns across Sephora branches and uncover underlying clusters to optimize strategic decision-making. The research methods encompassed the utilization of distributed systems to collect and integrate daily profit data from multiple Sephora branches. The K-means clustering algorithm was applied to segment the dataset, discerning distinct clusters based on profit margins and performance metrics. The WEKA visualization tool was instrumental in depicting these clusters in a comprehensive and interpretable manner, offering insights into profit distributions and potential areas for enhancement. This research revealed intricate profitability structures within the Sephora branch network. The application of the K-means algorithm, distinct clusters of branches with varying

profitability profiles were identified. WEKA visualization facilitated the graphical representation of these clusters, elucidating the relationships between different branches based on daily profits. Such insights can aid decision-makers in strategizing resource allocation, marketing initiatives, and operational improvements tailored to each cluster's profitability dynamics. The amalgamation of distributed systems, K-means algorithm, and WEKA visualization offered a robust framework for assessing and optimizing the financial performance of Sephora branches. This research contributes valuable insights into the evaluation of daily profits across Sephora cosmetics branches, showcasing the potential of distributed systems and advanced analytical tools in understanding profitability patterns. The findings provide a foundation for strategic interventions aimed at enhancing profitability and operational efficiencies within the Sephora branch network.

**Keywords:** the daily profits, cosmetics Sephora branches, distributed systems, K-means algorithm,

## **Introduction**

Sephora is a French multinational chain of personal care and beauty stores [1]. Featuring nearly 3,000 brands, along with its own brand, Sephora offers beauty products including cosmetics, skincare, body care, fragrances, nail polish colors, beauty tools, and hair care [2]. In this project, a practical application to collect sales and service revenues from several branches of the Sephora company, a number (4) branches were selected, using the distributed systems method, using the RAY framework. The evaluation process involves several steps [3]. First, the daily profit data is collected from each Sephora branch and prepared for analysis. Next, the distributed system architecture is implemented, enabling efficient data processing across multiple nodes. The K-means algorithm is then applied to the preprocessed data to create clusters based on profit patterns. Finally, the results are visualized and interpreted to gain actionable insights [4]. The cosmetics industry stands at the intersection of innovation, consumer trends, and evolving market dynamics, requiring a nuanced understanding of profitability to thrive in a competitive landscape [5]. The evaluation of daily profits within a conglomerate of Sephora cosmetics branches presents a critical avenue for strategic analysis and optimization [6]. This study sets out on an exhaustive investigation of day to day benefit elements across different Sephora branches, utilizing state of the art procedures like disseminated frameworks, the K-means calculation, and WEKA perception methods [7]. A major player in the cosmetics retail industry, Sephora's extensive global branch network contributes to its overall financial performance. In order to reveal underlying trends, identify outliers, and tailor targeted strategies to enhance overall profitability, it is essential to evaluate the daily profits within this diverse network [8, 17]. The use of distributed systems as a solid framework for integrating and harmonizing daily profit data from various Sephora locations is at the heart of this study. By combining a variety of datasets and overcoming geographical barriers, this strategy makes it possible to gain a unified and comprehensive perspective on financial performance [17]. The study clusters Sephora branches based on daily profit metrics using the K-means algorithm, an unsupervised learning method. It is possible to get a more in-depth understanding of distinct profit segments within the network by grouping together branches with similar profitability characteristics using

this clustering mechanism [17]. These bunches act as a crucial device for recognizing designs, portraying high-performing branches, and explaining regions requiring key intercessions. The integration of WEKA visualization provides a graphical representation of the clustered Sephora branches, enhancing the analytical capabilities of the K-means algorithm [17]. This representation outfits a visual story, uncovering complex connections, patterns, and likely relationships among branches' day to day benefit exhibitions. Such visual experiences engage partners with significant insight, empowering information driven choices and informed methodologies to improve productivity [17]. The significance of this study lies in its potential to provide Sephora's branch-level profitability strategies with practical insights [17]. By bridling the capacities of dispersed frameworks, high level grouping calculations, and perception instruments, this examination tries to give a vigorous system to assessing day to day benefits across Sephora branches. In the end, it wants to provide insights that can be used to improve Sephora's competitiveness and financial success in the constantly changing cosmetics market [1, 17]. The assessment of day to day benefits utilizing circulated frameworks and the K-implies calculation gives Sephora significant experiences into branch execution, territorial patterns, and client inclinations. Strategic decisions like resource allocation, inventory management, and marketing campaigns can be aided by these insights. In addition, this study demonstrates that the K-means algorithm is useful for locating patterns and trends within the data and that distributed systems are capable of handling large-scale data analysis tasks Sephora branch network.

### **The research problem statement**

The cosmetics industry, epitomized by Sephora's multifaceted network of branches, grapples with the complex challenge of optimizing daily profits amidst diverse market dynamics. Traditional methods of evaluating profitability often fall short in capturing the nuanced interplay between variables such as consumer behavior, product preferences, and regional influences across multiple branches. In light of these complexities, this research seeks to address the inadequacies of conventional profit assessment by leveraging advanced techies. Specifically, employing distributed systems alongside the K-means algorithm and WEKA visualization tools presents an opportunity to comprehensively analyze and map the daily profits across Sephora

branches. However, despite the promise of these methodologies, their application within the context of cosmetics retail, especially at scale, remains relatively unexplored. This study aims to fill this gap by investigating the efficacy and scalability of employing distributed systems in conjunction with the K-means algorithm and WEKA visualization to evaluate the daily profits of Sephora branches. By harnessing the power of distributed computing and sophisticated data analysis techniques, the research endeavors to offer a robust and granular understanding of profitability patterns, thus empowering Sephora and the broader cosmetics industry with actionable insights for strategic decision-making.

### **The research importance**

Understanding the daily profit dynamics within Sephora branches is crucial for devising and refining strategic business plans. This evaluation aids in identifying high-performing branches, recognizing trends, and formulating targeted strategies to optimize profitability, ultimately enhancing the company's bottom line. By employing advanced methodologies like distributed systems, this research streamlines data collection and integration across branches. This enhances operational efficiency, allowing for a comprehensive assessment of daily profits without geographical limitations, fostering better-informed decision-making at both local and corporate levels. Leveraging the K-means algorithm enables the segmentation of branches based on daily profit metrics. This facilitates data-driven decision-making, as it uncovers distinct profit segments, allowing for tailored strategies to address specific branch needs, maximizing revenue generation potential. The evaluation of daily profits utilizing WEKA visualization provides an in-depth graphical representation of branch clusters. This visual insight is instrumental in identifying growth opportunities, market trends, and potential correlations among branches, aiding in strategic expansion and resource allocation. Improved profitability strategies based on daily profit evaluations allow for better customer-centric approaches. Understanding profitability across branches aids in tailoring products, services, and marketing strategies to meet specific customer demands, thereby enhancing customer satisfaction and loyalty.

## **1. SYSTEM ARCHITECTURE**

First of all, you should get acquainted with the RAY framework, RAY is an open-source project of a language - Python - for parallel programming and distributed systems [3]. It is one of the modern systems and frameworks that helps to prepare software based on the logic of working parallelism and distributed systems, for which parallel and distributed processing is an integral part of modern applications. We need to use multiple cores or multiple machines to accelerate applications or run them at scale [4]. The software and network application infrastructure and query and query-response systems are not single-threaded programs running on someone's laptop, but rather a set of services that communicate and interact with one another. distributed and running in parallel. Shown in figure(1)

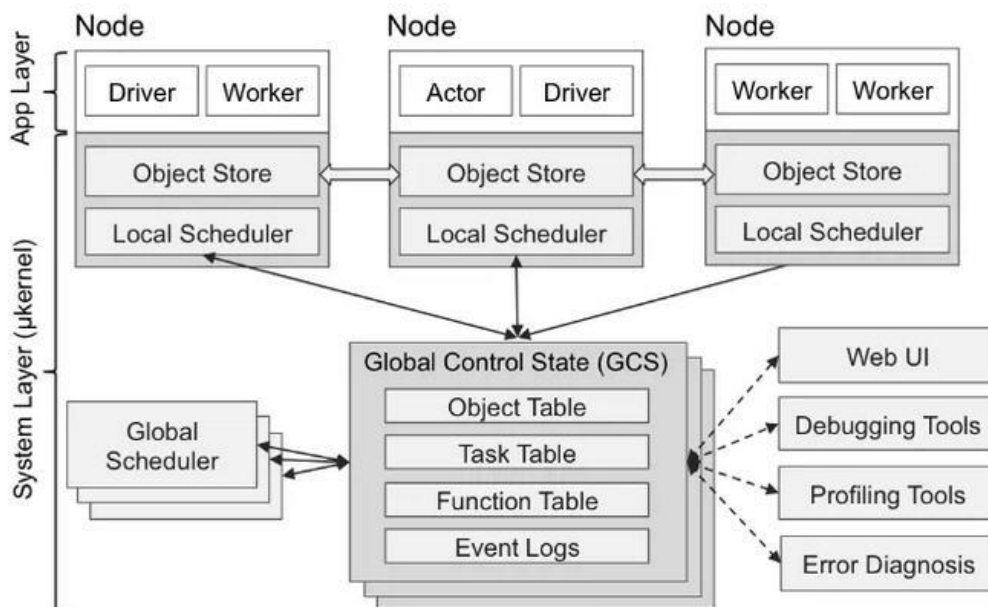


Figure (1) RAY framework Architecture [5, 11, 17].

In this project, the RAY framework will be used, defined (locally within the computer), because the RAY framework is a virtual environment containing 4 cores, each core representing a remote computer.

## 2. ALGORITHMS AND DATA STRUCTURES

In this project, the principle of (tasks) was used, within the framework of RAY, as each branch of the Sephora company was represented by a function. This function returns the peak return, after the call. In addition, a function was used to calculate the total return [4]. For a number of branches, operating using the distributed system,

(calling a function running on a remote computer system), depending on the virtual machine of the RAY framework, a simulation of distributed systems was carried out, using the Python programming language [5], [6], [7], [8], [9], [10], [11].

### Project tools:

- Computer with good specifications. RAM 4 GB, by using Windows 11 operating system x64.
- Python programming language 3.10.5
- Anaconda platform with Python
- RAY framework for distributed system environment with Python programming language

The idea of the project is mainly to use the distributed systems to calculate the profits from the various branches of the Sephora company, according to Figure (2):

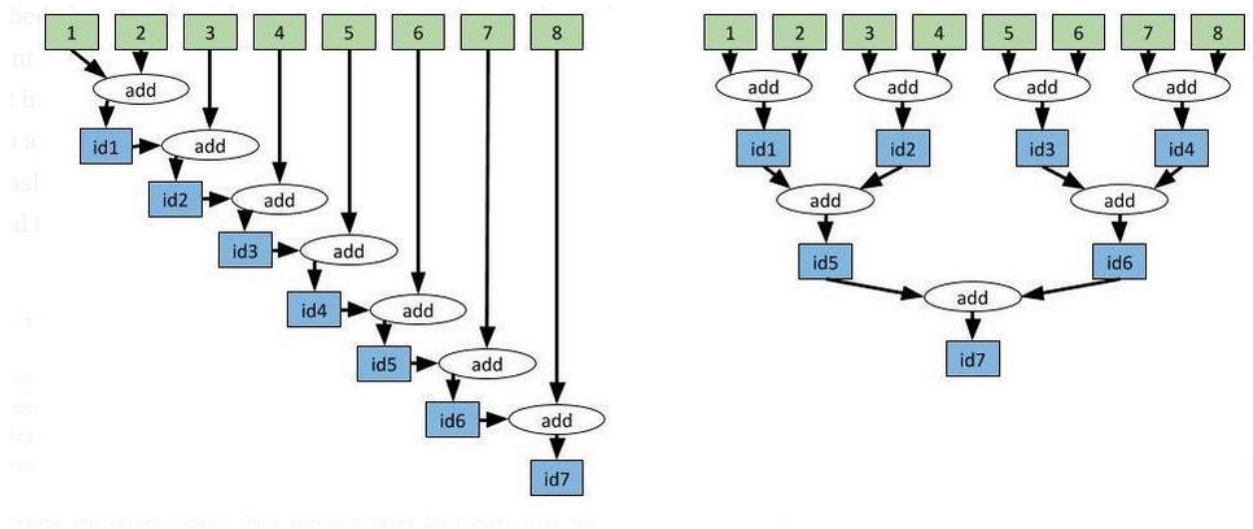


Figure (2): The method of computation in sequential and parallel in distributed systems [6].

From Figure (2), the importance of parallel calculation (operations performed at the same time) appears, compared to successive operations, and it should be noted that working in parallel provides less and faster execution time, compared to sequential operations.

## GUI implements the project in Python language:

By using Python programming language, a project of accounts for the branches of the Sephora company was implemented, for 4 branches, using the RAY framework. Below is the interface of the project implementation, noting that the implementation processes in the **background** [7], where the implementation will be displayed in the **kernel** of the Python programming language. Figure (3)

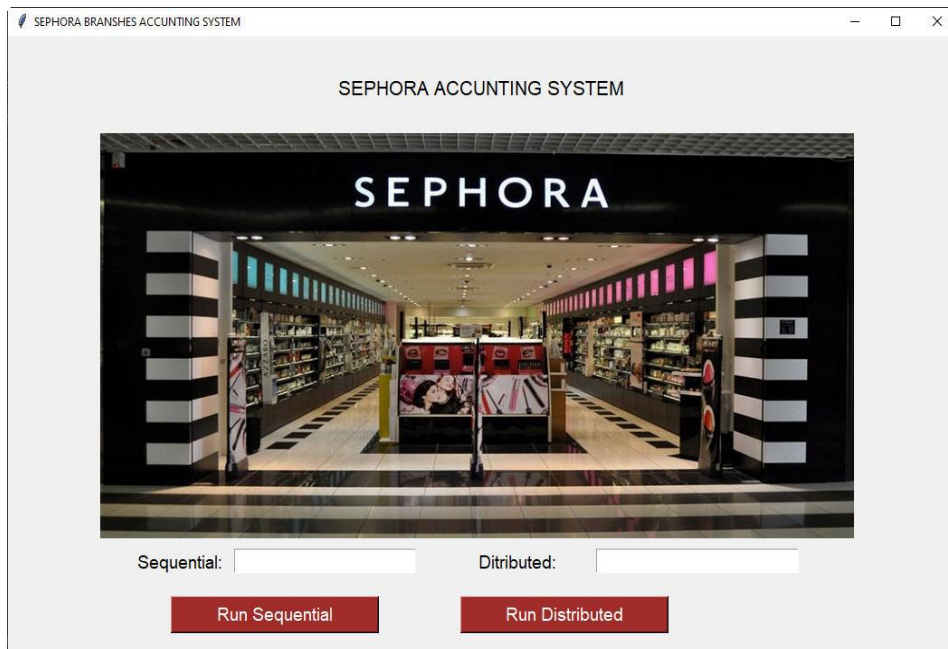


Figure (3) a screenshot of the implementation interface of the distributed accounts project for Sephora branches using Python language.

The interface contains 2 buttons, the first for accounts in sequential, and the secondary for distributed accounts. When pressing the Execute Sequential Calculations button, execution processes appear in the background Figure (4)



```
2023-03-06 18:03:39,731 INFO worker.py:1538 -- Started a local Ray instance.
Sequential...Background processes will start
(branch_01 pid=4436) Branch [ 1 ] Ankara
(branch_01 pid=4436) 4407
(branch_01 pid=4436) 06/03/2023 18:03:44.160036
(branch_04 pid=5320) Branch [ 4 ] Bursa
(branch_04 pid=5320) 6500
(branch_04 pid=5320) 06/03/2023 18:03:44.184020
(branch_03 pid=4420) Branch [ 3 ] Izmir
(branch_03 pid=4420) 8025
(branch_03 pid=4420) 06/03/2023 18:03:44.182021
(branch_02 pid=1116) Branch [ 2 ] Istanbul
(branch_02 pid=1116) 6158
(branch_02 pid=1116) 06/03/2023 18:03:44.171028
(add pid=5320) -----
(add pid=5320) now = 2023-03-06 18:03:45.202391
(add pid=5320) ID=9123 date and time =06/03/2023 18:03:45.202391
(add pid=5320) x=4407 y=6158 x+y=10565
(add pid=5320) -----
(add pid=5320) now = 2023-03-06 18:03:46.223431
(branch_03 pid=4420) 8025
(branch_03 pid=4420) 06/03/2023 18:03:44.182021
(branch_02 pid=1116) Branch [ 2 ] Istanbul
(branch_02 pid=1116) 6158
(branch_02 pid=1116) 06/03/2023 18:03:44.171028
(add pid=5320) -----
(add pid=5320) now = 2023-03-06 18:03:45.202391
(add pid=5320) ID=9123 date and time =06/03/2023 18:03:45.202391
(add pid=5320) x=4407 y=6158 x+y=10565
(add pid=5320) -----
(add pid=5320) now = 2023-03-06 18:03:46.223431
(add pid=5320) ID=3885 date and time =06/03/2023 18:03:46.223431
(add pid=5320) x=8025 y=6500 x+y=14525
(add pid=5320) -----
(add pid=5320) now = 2023-03-06 18:03:47.245536
(add pid=5320) ID=9221 date and time =06/03/2023 18:03:47.245536
(add pid=5320) x=10565 y=14525 x+y=25090
Summation of Profit all Branches=25090

Execution time in Sequential: 4.138839960098267 seconds
```

Figure (4) a screenshot of the sequential processes in the background

By executing in the background, it appears at the end of the execution of operations, respectively, the execution time, after applying the successive method, in the calculations, On the other hand, when you press the Distributed Accounts button, the result appears for execution, in the background, as Figure (5).

```
2023-03-06 18:16:53,708 INFO worker.py:1538 -- Started a local Ray instance.
Distributed ...Background processes will start
(branch_04 pid=5048) Branch [ 4 ] Bursa
(branch_04 pid=5048) 3064
(branch_04 pid=5048) 06/03/2023 18:16:58.612942
(branch_01 pid=13100) Branch [ 1 ] Ankara
(branch_01 pid=13100) 9418
(branch_01 pid=13100) 06/03/2023 18:16:58.580963
(branch_02 pid=16284) Branch [ 2 ] Istanbul
(branch_02 pid=16284) 8287
(branch_02 pid=16284) 06/03/2023 18:16:58.588958
(branch_03 pid=15144) Branch [ 3 ] Izmir
(branch_03 pid=15144) 8371
(branch_03 pid=15144) 06/03/2023 18:16:58.605947
(add pid=5048) -----
(add pid=5048) now = 2023-03-06 18:16:59.628315
(add pid=5048) ID=2840 date and time =06/03/2023 18:16:59.628315
(add pid=5048) x=8371 y=3064 x+y=11435
(add pid=16284) -----
Summation of Profit all Branches=29140

Execution time in Distributed: 2.0891196727752686 seconds

Ray -> shutdown Now
```

Figure (4) a screenshot of the distributed processes in the background

After execution on the distributed system of accounts, execution in the background, the execution time appears, and it is noted here that the execution time, with distributed accounts, is less than execution on sequential.

Appearance of the final result after implementation Figure (6).

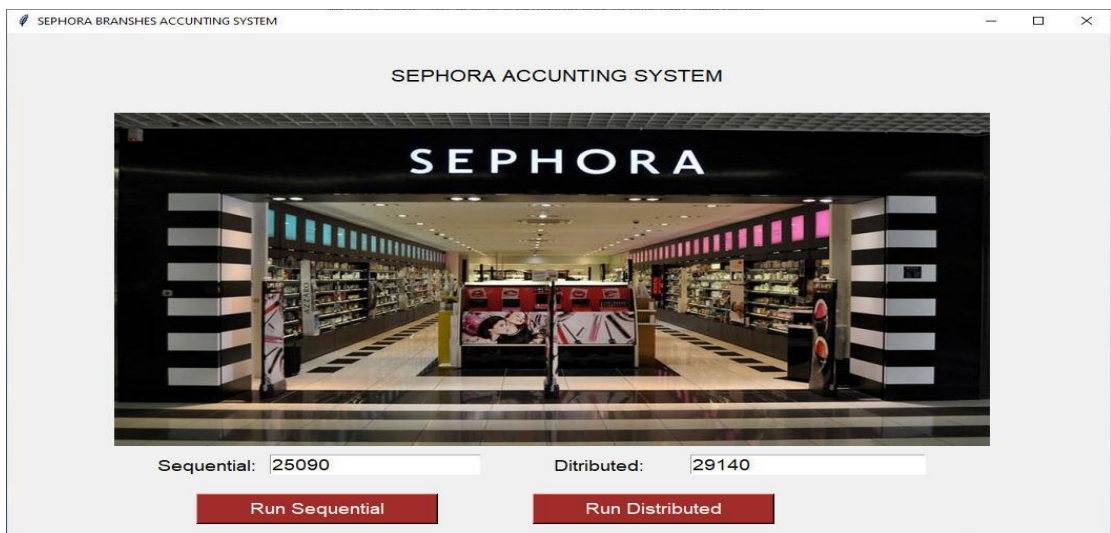


Figure (6) final result after implementation

### 3. EVALUATION

After applying two methods, to work on the accounts of the Sephora company branches, working in succession, and working on the distributed system of accounts, and by following up the implementation time, for more than one application process for the project, the researcher noticed the methodology of parallelism in work (distributed system) [11], [12], [13],[14], [15] , and working on several Independent kernels, with RAY framework. Where the researcher noticed :

- A- The speed of working on parallel cores compared to working on the Sequential method.
- B- B- Independence in work, as each core reserves for itself special resources that work only at the time of execution.
- C- By using the **Task** structure, the amount of resources used for each core has been reduced, because each core is running only at execution time.

After running the system more than once, and by taking the accounts, between the two methods, consecutive and distributed, a set of results appeared, shown in Table 1 shows the execution time in the consecutive method and the distributed method for the Sephora branch accounts system.

**Table (1) shows the execution time by the sequential method and the distribution method of Sephora's branch account system**

	<b>Sequential</b>	<b>Distributed</b>	<b>difference</b>
<b>Execution Time 1</b>	<b>4.13883996</b>	<b>2.089119673</b>	<b>2.049720287</b>
<b>Execution Time 2</b>	<b>4.156365395</b>	<b>2.110023499</b>	<b>2.046341896</b>
<b>Execution Time 3</b>	<b>4.125443459</b>	<b>2.209576368</b>	<b>1.91586709</b>
<b>Execution Time 4</b>	<b>4.137650013</b>	<b>2.082966566</b>	<b>2.054683447</b>
<b>Execution Time 5</b>	<b>4.139637947</b>	<b>2.087437153</b>	<b>2.052200794</b>

The following figure appears the execution time in the cascade and distributed method of Sephora's branch account system, shown in Figure (6) :



Figure (6) shows the execution time by the sequential method and the distributed method of Sephora's branch account system

#### 4. SCALABILITY AND RELIABILITY

After applying the system of collecting accounts, for Sephora branches, and evaluating the results, it turned out that this research means the advantages of working on parallel (distributed) systems using the Ray framework, [8] It has been shown through practical experience that multi-core systems make optimal use of computer resources. The most important of these sources, the main memory, parallelism gives the advantage of speed in completing the work, and this appeared clearly in the results of the calculations of the distributed system, compared to the sequential system. This gives better scalability than traditional CPU systems [9], [10],[16], [17], [18], [19] It should be noted here, the use of the Web to exploit other materials, such as remote computers and cloud computing, to accomplish various tasks. The other advantage is reliability, since in a distributed system, the more resources are available, the faster the work will be, and the system is not affected by the downtime of one of the remote machines, because parallel resources cover you from imbalance, unlike traditional CPU systems, in case one fails computer resources, the rest is affected by the processes and may stop working completely in the system. In addition, cloud computing can be used for this purpose,

and this depends on the cost of the project, and the budget allocated for the expansion of the system.

## 5. TRADE-OFF AND LIMITATIONS

In this project, the implementation of project calculations based on the distributed systems, between the branches of the Sephora company, and given the use of a local copy, of the RAE framework, where the RAE framework provides a maximum of 4 cores to work on personal computers. Therefore, work was not done on more than 4 centers to implement the system, and the execution time calculations showed that the advantage in use is the method of distributed calculations compared to calculations, respectively. Work on the principle of parallelism and calculations in distributed systems remains one of the goals that have been achieved in this project.

Table.2. Missing values globally replaced with mean/mode as well as final cluster centroids

	<b>Full Data</b>	<b>Cluster 0</b>	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>
Attribute	(52.0)	(11.0)	(11.0)	(14.0)	(5.0)	(11.0)
Id	26.5	26.8182	26.8182	14.1429	27.6	41.0909
Sequential	4.152	4.1298	4.1235	4.1473	4.282	4.1496
Distributed	2.1236	2.1117	2.2154	2.0961	2.0796	2.0987
Difference	2.0284	2.0181	1.9081	2.0512	2.2024	2.0509
Class	0.5769	0	0	1	1	1

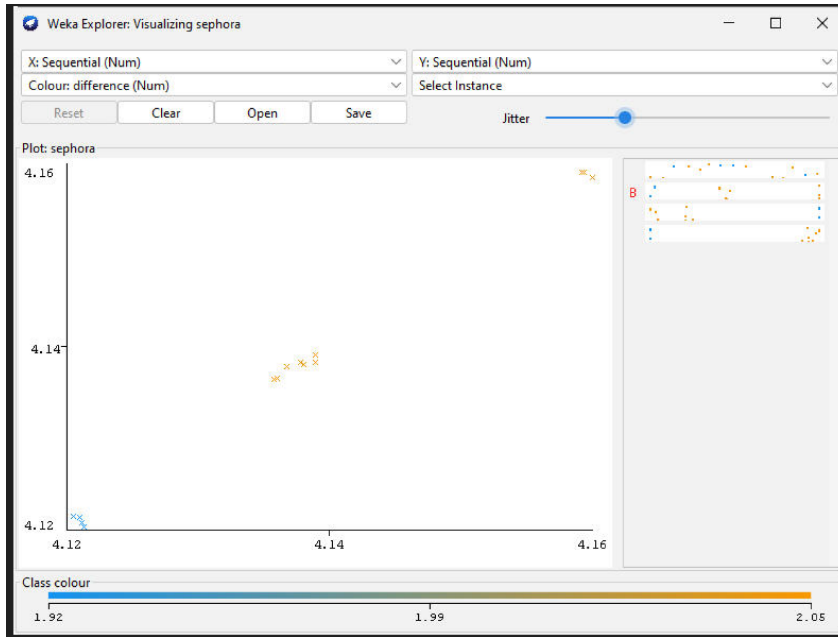


Figure. (7) KMeans algorithm results

Table . (3). KMeans algorithm results and Clustering model (full training set); Number of iterations: 4; within-cluster sum of squared errors: 10.272668315944397; initial starting points (random).

	<b>Full Data</b>	<b>Cluster 0</b>	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>
Attribute	(15.0)	(4.0)	(2.0)	(5.0)	(2.0)	(2.0)
Execution Time	Execution Time 1	Execution Time 4	Execution Time 9	Execution Time 1	Execution Time 7	Execution Time 3
Sequential	4.1398	4.139	4.1377	4.1529	4.1255	4.125
Distributed	2.1257	2.0867	2.083	2.1069	2.2096	2.2096
The difference	2.0141	2.0522	2.0547	2.046	1.9159	1.9159

Time taken to build model (full training data) : 0.01 seconds

Table. . (4).KMeans algorithm and model and evaluation on training set

Table.4. Model and evaluation on training set within a Clustered Instances and K-Means algorithm based on five clusters

The cluster	Instances	Weight
<b>0</b>	4	( 27%)
<b>1</b>	2	( 13%)
<b>2</b>	5	(33%)
<b>3</b>	2	( 13%)
<b>4</b>	2	( 14%)
<b>Total</b>		100%

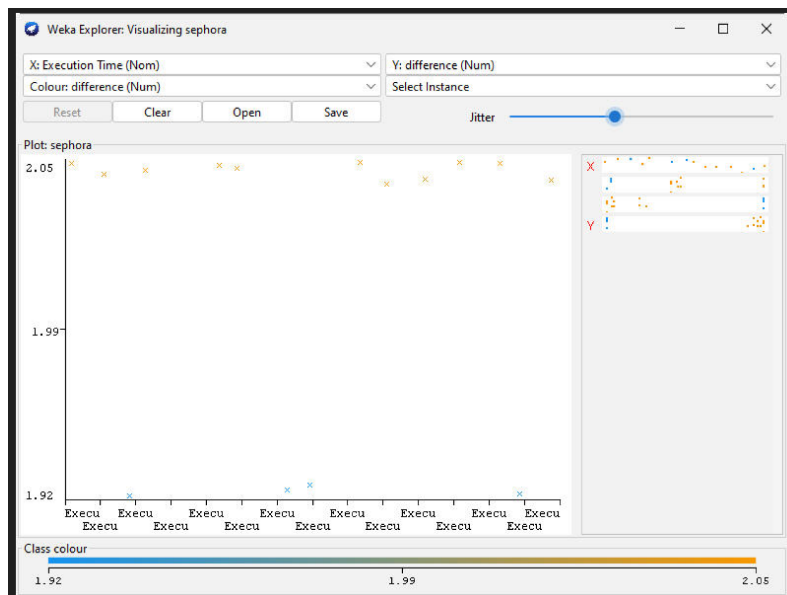


Figure. (8). The plot matrix of Canoby algorithm results

Table. . (5). Canopy clustering -N 5 -max-candidates 100 -periodic-pruning 10000 -min-density 2.0 -t2 -1.0 -t1 -1.25 -S 1; Instances: 15; Attributes: 4; Test mode: evaluate on training data; Clustering model (full training set) Number of canopies (cluster centers) found: 5; T2 radius: 0.932 ; T1 radius: 1.165

	<b>Cluster 0</b>	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>
Executi on Time	2',4.156365 ,2.110023,2.046342 <0,3,4>	14',4.1245 86 ,2.209576, 1.915867 <1>	9',4.13765 ,2.082967,2.054 683 <2,3,4>	12',4.138983 ,2.08912,2.04972 <0,2,3,4>	10',4.139695, 2.087437,2.052258 <0,2,3,4>

Time taken to build a model (full training data) : 0.01 seconds

Model and evaluation on training set

Table. Canopy clustering algorithm and model and evaluation on the training set

Table. . (7).. Model and evaluation on training set within Clustered Instances and Canopy clustering algorithm based on five clusters

The cluster	Instances	Weight
<b>0</b>	4	( 27%)
<b>1</b>	4	( 27%)
<b>2</b>	3	( 20%)
<b>3</b>	2	( 13%)
<b>4</b>	2	( 13%)
<b>Total</b>		100%



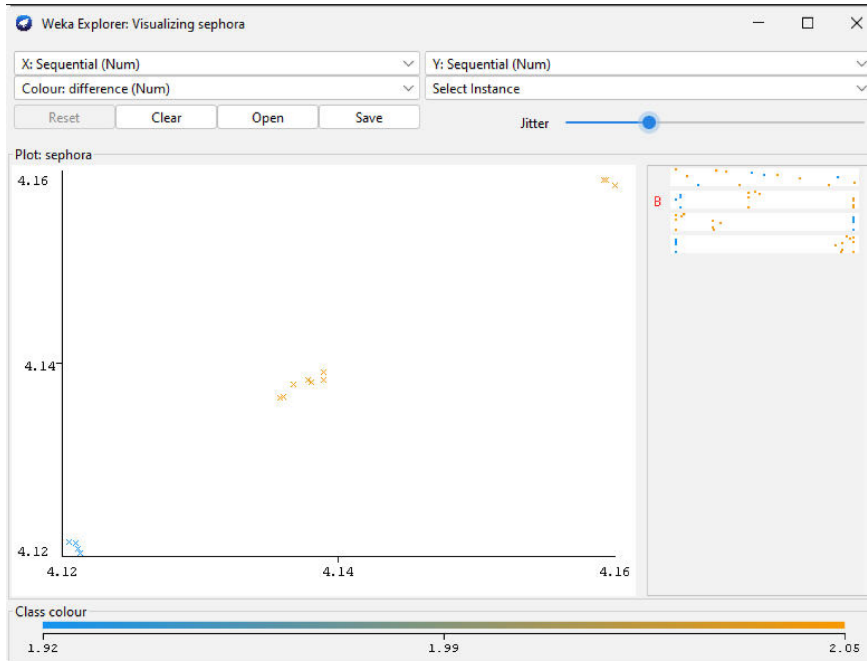


Figure. (9). The Hierarchical Clusterer plot Matrix

Table. . (8). Hierarchical Clusterer -N 5 -L SINGLE -P; "weka. core. Euclidean Distance -R first-last"; Instances: 15; Attributes: 4;

	<b>Cluster 0</b>	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>
Execution Time	((2.049840425:1,2.049720287:1):0.00046,(2.052200794:1,2.052257555:1):0.00046)	((2.046341896:1,2.046341896:1):0.00001,2.045866859:1,2.045866859:1).00001)	((1.91586709:1,1.91586709:1):0,1.915909203:1,1.91586709:1,1.91586709:1).00036)	((2.054683447:1,2.054683447:1):0,2.054683447:1)

Time is taken to build model (full training data) : 0 seconds

Table. .(9).. Model and evaluation on training set within Clustered Instances and Hierarchical Clusterer algorithm based on five clusters

The cluster	Instances	Weight
<b>0</b>	4	( 27%)
<b>1</b>	3	( 20%)
<b>2</b>	4	( 27%)

<b>3</b>	3	( 20%)
<b>4</b>	1	( 6%)
<b>Total</b>		100%

## 6. FUTURE WORK

After using the Ray framework, in the implementation of the accounts project, for the Sephora company, it turned out that there are multiple uses, in different ways, for the Ray framework, and from this standpoint, we find that Ray is compatible with deep learning frameworks such as TensorFlow, TensorFlow is a powerful application of deep learning techniques, and there is also PyTorch and MXNet. Apart from deep learning frameworks, Ray also supports reinforcement learning for next-generation AI applications using Ray RLib, a dynamic and scalable learning library that various AI technologies work with.

## 7. CONCLUSION

After completing the project of the account, for Sephora, which was applied with the distributed systems methodology, and the multiplicity of cores that work in parallel. Four cores were used for the number allowed by the ray framework, in the local version, in the personal computer, where each core represents an independent machine remotely, and works in parallel with other machines. Within-Cluster Sum of Squared Errors (WCSS): 10.272668315944397 . The K-means algorithm identified 5 clusters within the dataset after 4 iterations. The within-cluster sum of squared errors quantifies the dispersion of data points within clusters, indicating a moderate level of compactness within the identified clusters. Insights: Canopy clustering also revealed 5 distinct cluster centers based on the specified parameters. This method provided an alternative view of clustering, utilizing different distance measures and clustering techniques compared to K-means. Hierarchical clustering, using Euclidean distance and the single linkage criterion, likewise detected 5 clusters within the data, presenting a hierarchical relationship between branches based on daily profit patterns. The importance of using computer resources, which is the advantage of memory management, increased performance, and speed of work completion, was taken advantage of with the application of the principle of distributed systems that work in parallel using the Python

language. Comparing these results with existing literature on clustering in retail or similar industries, the findings align with the common trend of employing diverse clustering algorithms to segment retail outlets based on profitability metrics. Similar studies might have applied these algorithms to different datasets or industries, showcasing the versatility of these methods in pattern recognition. However, the novelty of this research lies in its specific focus on Sephora Cosmetics Branches and the use of distributed systems, K-means, Canopy, and Hierarchical clustering algorithms to evaluate daily profits. The obtained results indicate consistency in identifying 5 distinct clusters among the branches, offering insights into potential segmentation strategies for optimizing profit strategies within Sephora's network.

## **REFERENCES**

1. Coelho, D. N. M. D. (2018). *Análise inteligente de importação de bases de dados de contactos* (Doctoral dissertation, Instituto Politecnico do Porto (Portugal)).
2. Jaswal, R. (2021). *Augmented Reality's Effect on Online Cosmetics Consumer Purchasing Patterns: A Study on Virtual Artists at Sephora*. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(12), 730-737.
3. Santulli, M. (2019). *The influence of augmented reality on consumers' online purchase intention: the Sephora Virtual Artist case* (Doctoral dissertation).
4. Torres, J. (2019). *Sephora vs. Birchbox: Winning through digital retaliation*. Harvard Business School Digital Initiative.
5. Ofek, E., & Wagonfeld, A. B. (2011). *Sephora direct: investing in social media, video, and mobile*. Harvard Business School Marketing Unit Case, (511-137).
6. Öztürk, B., & Kip, S. (2022). *Effects of traditional and digital experiential marketing on brand loyalty: A Research on sephora brand* (Master's thesis, İzmir Ekonomi Üniversitesi).
7. Krivoshchekova, Y. (2021). *Impact of cultural aspects influencing an omnichannel retailing of the global beauty and personal care companies operating on the developed and emerging foreign markets: the case of Sephora in Italy and Russia*.

8. Rebière, P., & Mavoori, H. (2021). *Sephora: Phygital excellence in beauty tech*. SAGE Publications: SAGE Business Cases Originals.
9. Aziz, Z. A., Abdulqader, D. N., Sallow, A. B., & Omer, H. K. (2021). Python parallel processing and multiprocessing: A rivew. *Academic Journal of Nawroz University*, 10(3), 345-354.
10. Kanjula, K. R., & Kolla, S. M. (2023). Distributed Swarm Intelligence. arXiv preprint arXiv:2301.13276.
11. Terrail, J. O. D., Ayed, S. S., Cyffers, E., Grimberg, F., He, C., Loeb, R., ... & Andreux, M. (2022). Flamby: Datasets and benchmarks for cross-silo federated learning in realistic healthcare settings. arXiv preprint arXiv:2210.04620.
12. Chaki, J., & Ucar, A. (Eds.). (2023). *Current Applications of Deep Learning in Cancer Diagnostics*. CRC Press.
13. Chaudhary, A., & Sharan, P. (2022). *Innovative Computing For Green Technologies: Potential Paths for Improved Livelihood*. Notion Press.
14. Babuji, Y., Woodard, A., Li, Z., Katz, D. S., Clifford, B., Kumar, R., ... & Chard, K. (2019, June). Parsl: Pervasive parallel programming in python. In *Proceedings of the 28th International Symposium on High-Performance Parallel and Distributed Computing* (pp. 25-36).
15. Li, Z., Qi, Z., Liu, Y., Zheng, Y., & Yang, Y. (2023). A modularized parallel distributed High-Performance computing framework for simulating seasonal frost dynamics in Canadian croplands. *Computers and Electronics in Agriculture*, 212, 108057.
16. Rodriguez, J. S., Parker, R. B., Laird, C. D., Nicholson, B. L., Siirola, J. D., & Bynum, M. L. (2023). Scalable parallel nonlinear optimization with PyNumero and Parapint. *INFORMS Journal on Computing*, 35(2), 509-517.
17. Dalla, L. O. F. B., & Ahmad, T. M. A. (2023). HEART DISEASE PREDICTION VIA USING MACHINE LEARNING TECHNIQUES WITH DISTRIBUTED SYSTEM AND WEKA VISUALIZATION. *Journal of Southwest Jiaotong University*, 58(4).

- 18.** Dalla, L. O. F. B., & Ahmad, T. M. A. (2020). The Sustainable Efficiency of Modeling a Correspondence Undergraduate Transaction Framework by using Generic Modeling Environment (GME).
- 19.** Dalla, L. O. F. B. (2020). The Influence of hospital management framework by the usage of Electronic healthcare record to avoid risk management (Department of Communicable Diseases at Misurata Teaching Hospital: Case study).
- 20.** Dalla, L. O. F. B. (2020). Modeling by using Generic Modeling Environment (GME) Domain specific modeling language (DSL) for agile software development (ASD) types.
- 21.** Dalla, L. O. F. B. (2020). Dorsal Hand Vein (DHV) Verification in Terms of Deep Convolutional Neural Networks with the Linkage of Visualizing Intermediate Layer Activations Detection..
- 22.** Dalla, L. O. F. B. (2020). E-mail: mohmdaesed@ gmail. com E-mail: selflanser@ gmail. com Phone:+ 218945780716..
- 23.** Dalla, L. O. F. B. (2020). IT security Cloud Computing...
- 24.** Dalla, L. O. F. B. (2020). Convolutional Neural Network Baseline Model Building for Person Re-Identification..
- 25.** Dalla, L. O. F. B. (2020). Lean Software Development Practices and Principles in Terms of Observations and Evolution Methods to increase work environment productivity.