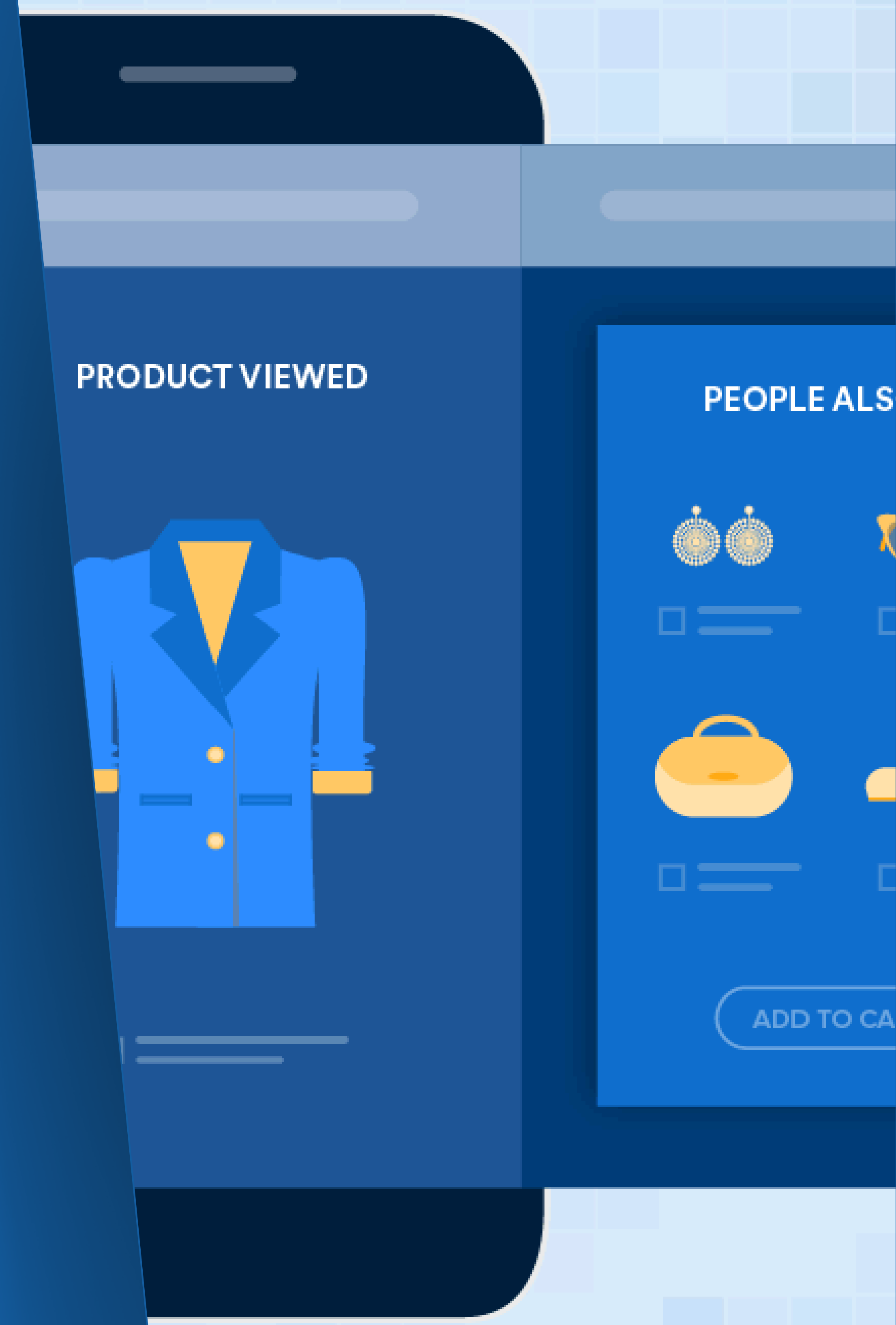
PERSONALIZED PRODUCT RECOMMENDATION SYSTEM

ABOUT ME

- M.M.Chathuni Devindi Manage
- Undergraduate at Informatics Institute of Technology in affiliation with Robert Gordon university
- IID ID : 20200997
- RGU ID : 2017915

INTRODUCTION

This project enhances e-commerce engagement by improving recommendation systems using Graph Neural Networks (GNNs) and K-means clustering. These techniques refine user-item interactions and customer segmentation, boosting the accuracy and personalization of product suggestions.



PROBLEM IN BRIEF

In the dynamic world face the challenge of delivering precise and personalized product recommendations amidst vast product assortments. Current recommendation systems often fall short in accuracy and fail to effectively adapt to the rapidly changing preferences of users.

SOLUTION

This project addresses the need for more sophisticated algorithms that can handle complex user-item interactions more adeptly, thereby improving the personalization and relevance of product suggestions.

AIM

Primary Objectives

GNN & K-means clustering

Technical Improvements

Analyze and interpret complex user-item interactions

Business Impact

Increase sales conversions on e-commerce platforms

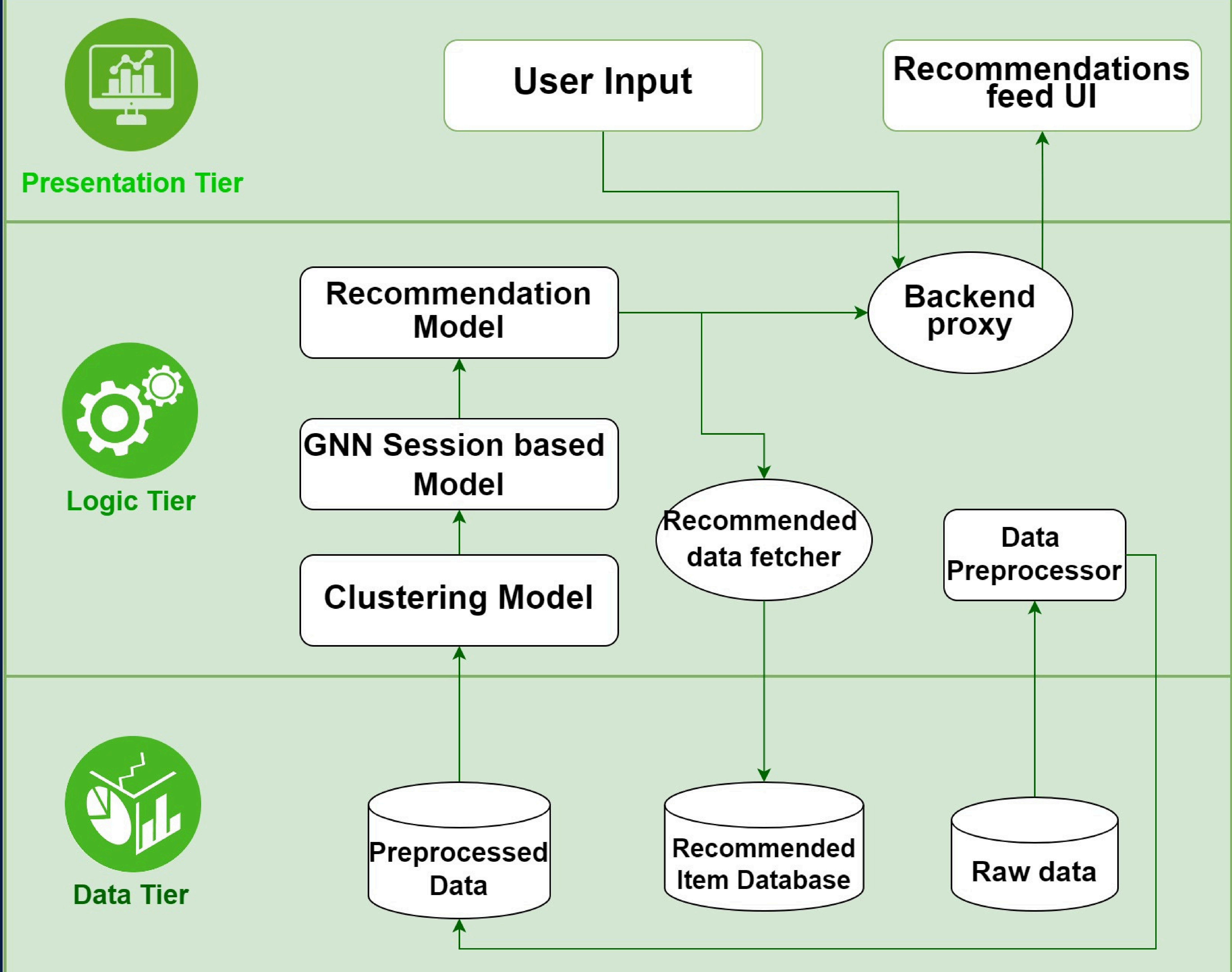
Contribution to Research

Fill the research gap in the use of GNNs within personalized recommendation systems

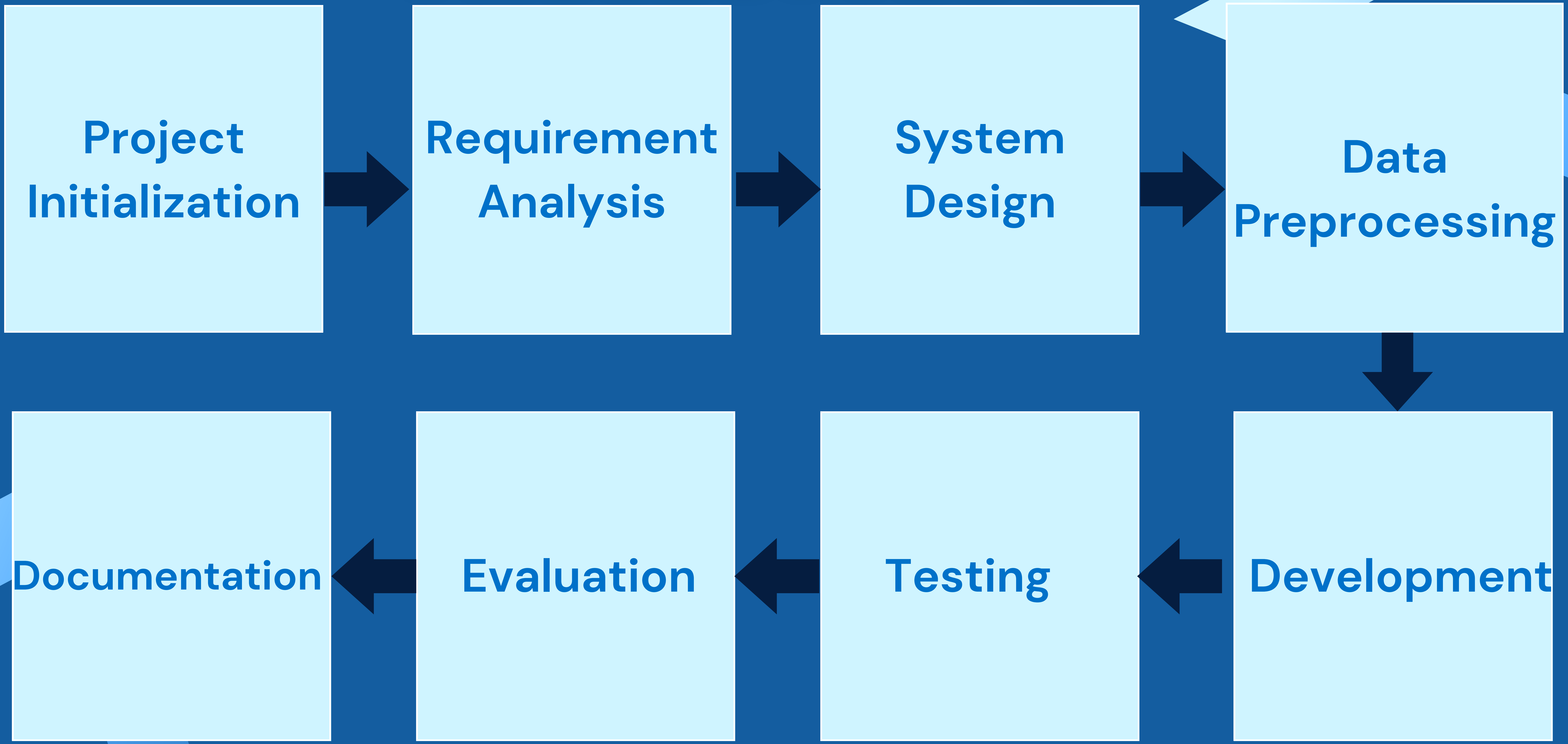
RESEARCH

Click here to view clearly.

HIGH LEVEL ARCHITECTURE



PROJECT FLOW



DATASET

- Online Retail dataset from UC Irvine Machine Learning Repository

InvoiceNo	StockCode	Descriptio	Quantity	InvoiceDate	UnitPrice	CustomerI	C
536365	85123A	WHITE HAI	6	01-12-10 8:26	2.55	17850	U
536365	71053	WHITE MET	6	01-12-10 8:26	3.39	17850	U
536365	84406B	CREAM CU	8	01-12-10 8:26	2.75	17850	U
536365	84029G	KNITTED U	6	01-12-10 8:26	3.39	17850	U
536365	84029E	RED WOOL	6	01-12-10 8:26	3.39	17850	U
536365	22752	SET 7 BABU	2	01-12-10 8:26	7.65	17850	U
536365	21730	GLASS STA	6	01-12-10 8:26	4.25	17850	U
536366	22633	HAND WAI	6	01-12-10 8:28	1.85	17850	U
536366	22632	HAND WAI	6	01-12-10 8:28	1.85	17850	U
536367	84879	ASSORTED	32	01-12-10 8:34	1.69	13047	U
536367	22745	POPPY'S P	6	01-12-10 8:34	2.1	13047	U
536367	22748	POPPY'S P	6	01-12-10 8:34	2.1	13047	U
536367	22749	FELTCRAF	8	01-12-10 8:34	3.75	13047	U
536367	22310	IVORY KNI	6	01-12-10 8:34	1.65	13047	U
536367	84969	BOX OF 6 A	6	01-12-10 8:34	4.25	13047	U

Dataset Characteristics	Feature Type	Instances	Features
Multivariate, Sequential, Time-Series	Integer, Real	541909	7

DATA CLEANING & TRANSFORMATION

- Handling missing values
- Handling duplicates
- Treating cancelled transactions
- Correcting stockcode anomalies
- Cleaning description column
- Treating zero unit prices
- Outlier treatment

FEATURE ENGINEERING

Recency, Frequency, Monetary features

Geographic features

Product diversity

Cancellation insights

Behavioral features

Seasons & Trends

DATA PREPROCESSING

Feature scalling

Dimensionality reduction

CLUSTERING MODEL (K-MEANS)

- Segment customers based on their transaction behaviors to tailor marketing strategies and improve customer engagement.

- Parameter Selection:

  n_clusters: Determined through Elbow Method and Silhouette Analysis to identify the optimal number of clusters.

  Initialization: k-means++ to ensure a smarter initialization and better convergence.

CLUSTERING MODEL (K-MEANS)

- Evaluation

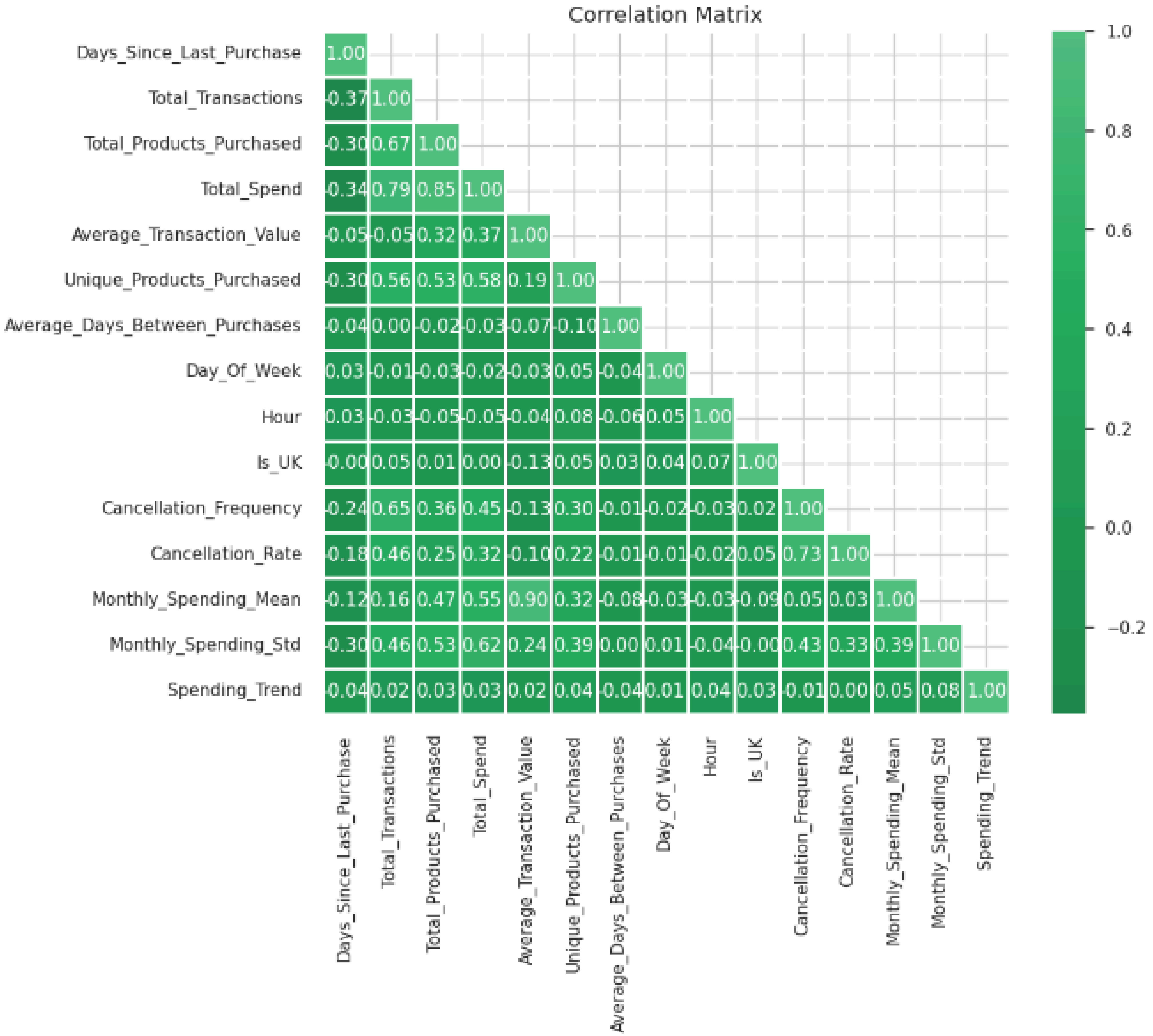
Silhouette Score: Measures how similar an object is to its own cluster compared to other clusters.

Calinski-Harabasz Score: Ratio of sum of between-clusters dispersion and of within-cluster dispersion for all clusters where higher scores denote better defined clusters.

Davies-Bouldin Score: The average 'similarity' between clusters, where lower values indicate better separation.

PERFORMANCE EVALUATION OF CLUSTERING MODEL (K-MEANS)

Metric	Value
Number of Observations	4067
Silhouette Score	0.25619706506491186
Calinski Harabasz Score	1327.024060259708
Davies Bouldin Score	1.299281578901572



Correlation Matrix
Days_Since_Last_Purchase 1.00
Total_Transactions -0.37 1.00
Total_Products_Purchased -0.30 0.67 1.00
Total_Spend -0.34 0.79 0.85 1.00
Average_Transaction_Value -0.05 -0.05 0.32 0.37 1.00
Unique_Products_Purchased -0.30 0.56 0.53 0.58 0.19 1.00
Average_Days_Between_Purchases -0.04 0.00 -0.02 -0.03 -0.07 -0.10 1.00
Day_Of_Week 0.03 -0.01 -0.03 -0.02 -0.03 0.05 -0.04 1.00
Hour 0.03 -0.03 -0.05 -0.05 -0.04 0.08 -0.06 0.05 1.00
Is_UK -0.00 0.05 0.01 0.00 -0.13 0.05 0.03 0.04 0.07 1.00
Cancellation_Frequency -0.24 0.65 0.36 0.45 -0.13 0.30 -0.01 -0.02 -0.03 0.02 1.00
Cancellation_Rate -0.18 0.46 0.25 0.32 -0.10 0.22 -0.01 -0.01 -0.02 0.05 0.73 1.00
Monthly_Spending_Mean -0.12 0.16 0.47 0.55 0.90 0.32 -0.08 -0.03 -0.03 -0.09 0.05 0.03 1.00
Monthly_Spending_Std -0.30 0.46 0.53 0.62 0.24 0.39 0.00 0.01 -0.04 -0.00 0.43 0.33 0.39 1.00
Spending_Trend -0.04 0.02 0.03 0.03 0.02 0.04 -0.04 0.01 0.04 0.03 -0.01 0.00 0.05 0.08 1.00
Days_Since_Last_Purchase
Total_Transactions
Total_Products_Purchased
Total_Spend
Average_Transaction_Value
Unique_Products_Purchased
Average_Days_Between_Purchases
Day_Of_Week
Hour
Is_UK
Cancellation_Frequency
Cancellation_Rate
Monthly_Spending_Mean
Monthly_Spending_Std
Spending_Trend
1.0
0.8
0.6
0.4
0.2
0.0
−0.2

GNN MODEL

- build_graph()

Each sequence contributes to the edge weights between nodes (products), which represent transitions from one product to another.

- data_masks()

It prepares the data for training by padding sequences to the same length, using **item_tail** as padding. It also creates masks to indicate the actual length of each sequence before padding.

GNN MODEL

- Data class

Encapsulates the data handling for the neural network, including batching and slicing of data. It also converts input sequences into graph adjacency matrices for the GNN.

- GNN class

Defines a GNN cell, similar to an LSTM cell but modified to work with graph data. It uses edge-weighted inputs to update node states.

- SessionGraph class

Main model class that integrates the GNN with embeddings and linear transformations to compute scores for recommendation.

PERFORMACE EVALUATION OF GNN MODEL

Testing Method	Recall	Precision	MRR	Epoch
GNN-based RecSys	62.12%	72.57%	31.97%	29

```
start training:  2024-04-26 13:08:14.054108
[0/12] Loss: 5.4256
[3/12] Loss: 5.3548
[6/12] Loss: 5.2842
[9/12] Loss: 5.2396
[12/12] Loss: 5.2127
start predicting:  2024-04-26 19:08:14.054108
Best Result:
        MMR@20: 0.3197  Precision@20:   0.7257  Average Loss:   2.0398  Epoch:  29
```

Let's see the demonstration...

- If there is any issue with the live demonstration, click <u>here</u>

TECHNOLOGY STACK

Presentation Tier

React JS HTML CSS matplotlib plotly

Logic Tier

kaggle Flask CO scikit learn seaborn NumPy SciPy PyTorch

Logic Tier

pandas Yellowbrick

BENCHMARK

Benchmarked the SR-GNN model against other neural-network-based recommendation systems such as NARM and STAMP to evaluate its effectiveness in session-based recommendations.

Method	Precision	MRR
NARM	68.32	28.63
STAMP	68.74	29.67
SG-GNN	72.57	31.97

LIMITATIONS

- Lack of Real-Time Data Processing
- Scalability Concerns
- Complexity in Integration
- Dependency on Quality Data

EVALUATION

- Research paper

Read full paper clicking here

Graph Neural Network and K-means Clustering based Recommendation System

Devindi Manage
Informatic Institute of Technology (of Aff.)
Robert Gordon University
Aberdeen, UK
chathuni.20200997@iit.ac.lk

Prashan Rathnayake
School of Computing Informatics
Institute of Technology
Colombo, Sri Lanka
prashan.r@iit.ac.lk

***Abstract*—Product recommendation systems are pivotal in e-commerce platforms, significantly influencing user experience and business revenue by suggesting relevant products to users. These systems harness user data and machine learning to predict and serve user-specific interests. This project focuses on enhancing the effectiveness of product recommendations in e-commerce by employing Graph Neural Networks (GNNs) and K-means customer segmentation techniques. GNNs are adept at processing the graph-structured data that naturally arises from user-item interactions, capturing complex relationships and user preferences. Concurrently, K-means segmentation further refines the recommendation process by grouping customers into clusters based on their behavior, which allows for more targeted and meaningful suggestions. The system developed as part of this thesis demonstrates a significant leap in performance and accuracy in product recommendations. It successfully identifies patterns in user behavior and predicts products that align with individual preferences, leveraging a robust dataset and state-of-the-art algorithms. The system's architecture is designed to be scalable and can be integrated seamlessly into existing e-commerce platforms to enhance the user experience. The evaluation of the system utilized a combination of quantitative metrics such as precision, Mean Reciprocal Rank (MRR), and several cluster analysis scores like Silhouette, Calinski-Harabasz, and Davies-Bouldin indices. These metrics collectively helped assess the accuracy, efficiency, and scalability of the recommendation model. However, the current implementation of the system does not include real-time data processing, which is identified as a key area for future enhancement. Real-time data integration would allow the system to adapt recommendations on the fly, further improving its relevance and accuracy.**

***Index Terms*—Graph Neural Networks, Personalized Recommendation Systems, E-commerce, Session-based Recommendations, Artificial Intelligence**

I. Introduction

In the rapidly evolving world of e-commerce, the ability to deliver personalized experiences stands as a critical determinant of success. Among the technologies driving this personalization, product recommendation systems play a pivotal role. These systems not only enhance user experience by providing tailored suggestions but also significantly contribute to increased business revenue. Leveraging user data and machine learning technologies, modern recommendation systems predict and serve user-specific preferences with remarkable accuracy.

The significance of effective recommendation systems is underscored by the shift in consumer behavior towards online platforms, where the sheer volume of options can overwhelm users. Traditional recommendation methods, while effective, often struggle to manage the dynamic and complex nature of user interactions and preferences in real-time. This project proposes an advanced recommendation system that integrates Graph Neural Networks (GNNs) and customer segmentation techniques to address these challenges. GNNs are particularly adept at processing the graph-structured data that naturally arises from user-item interactions, capturing intricate relationship patterns that are not easily deciphered by conventional predictive models.

Furthermore, the segmentation of customers into distinct groups based on behavioral data allows for more targeted and meaningful recommendations. This approach not only promises to enhance the accuracy of product suggestions but also ensures that recommendations are contextually relevant, thereby improving user engagement and satisfaction.

This paper details the development of a sophisticated e-commerce recommendation system that employs state-of-the-art machine learning techniques to improve product alignment with individual user preferences. By focusing on the integration of GNNs within a robust architecture designed for scalability and real-time data processing, this research contributes to the ongoing advancement of recommendation technologies in digital marketplaces. The system's performance, evaluated through precision metrics and user feedback, demonstrates significant improvements over traditional recommendation models, offering insights into the potential for future enhancements that could further refine the personalization process.

The following sections will outline the research methodology, system design, implementation details, and the comprehensive evaluation of the system's effectiveness, laying the groundwork for future work in this vital area of e-commerce.

II. Literature Review

1. Evolution of Recommendation Systems in E-commerce

Recommendation systems have become a cornerstone of e-commerce platforms, significantly enhancing user experiences by suggesting products tailored to their preferences and browsing habits. Initially, these systems primarily utilized

EVALUATION

- Review paper (Proof for the research)

Read full paper clicking here

Review of Graph Neural Network based Recommendation Systems

Devindi Manage
Informatic Institute of Technology (of Aff.)
Robert Gordon University
Aberdeen, UK
chathuni.20200997@iit.ac.lk@gma

Prashan Rathnayake
School of Computing Informatics
Institute of Technology
Colombo, Sri Lanka
prashan.r@iit.ac.lk

Nipuna Senanayake
School of Computing Informatics
Institute of Technology
Colombo, Sri Lanka
nipuna.s@iit.ac.lk

***Abstract*—This review paper presents a comprehensive overview of recent advancements in recommendation systems powered by Graph Neural Networks (GNNs). It explores how GNNs leverage complex user-item interaction graphs to significantly enhance recommendation accuracy and personalization. By integrating structural information and user-item relationships, GNN-based models outperform traditional recommendation systems.The paper categorizes various GNN architectures, discusses their strengths in different recommendation scenarios, and highlights future directions for research, underscoring the transformative potential of GNNs in recommendation systems.**

***Index Terms*—Graph Neural Networks (GNNs), Recommendation Systems, Personalization and Scalability**

I. Introduction

Recommendation systems have become essential tools in many different domains, such as e-commerce, social networking, and entertainment, during the enormous span of the digital era.These systems are pivotal in enhancing user experiences by providing personalized content and recommendations, thus increasing user engagement and driving business revenues. As the internet becomes increasingly saturated with information, the ability of recommendation systems to filter through the noise and present users with content that aligns with their preferences has become crucial for maintaining user interest and loyalty.

The evolution of recommendation techniques has been marked by significant milestones, transitioning from traditional methods such as collaborative filtering and content-based filtering to more sophisticated algorithms enabled by the advent of machine learning and deep learning. Traditional recommendation methods, while foundational, often fall short in handling the complexities of large-scale, high-dimensional data and fail to capture the nuanced preferences of individual users. This has spurred the need for more advanced recommendation algorithms that can effectively parse through vast datasets to deliver highly personalized recommendations.

Enter Graph Neural Networks (GNNs), a revolutionary framework that has transformed the landscape of recommendation systems. GNNs are adept at learning representations of data structured as graphs, encompassing components such as graph convolutional layers, message passing mechanisms, and graph-level aggregation. This framework excels in modeling data with inherent relational structures, making it uniquely suited for recommendation systems where user-item interactions naturally form a graph.

The motivation for leveraging GNN-based recommendation systems arises from the limitations of traditional recommendation techniques to capture complex user-item interactions and utilize rich contextual information. Traditional methods often rely on simplistic assumptions and fail to consider the intricate web of relationships and dependencies within the data. GNNs, with their ability to directly operate on the graph structure of recommendation data, offer a promising solution. By exploiting the graph-based nature of user-item interactions, GNNs can unearth deeper insights into user preferences and item characteristics, facilitating more accurate and personalized recommendations.

This review paper aims to delve into the rapidly developing field of GNN-based recommendation systems. It seeks to provide a comprehensive overview of recent advancements, dissecting key methodologies and analyzing their performance in contrast to conventional approaches. The scope of this paper is focused on exploring how GNNs are applied across different stages of recommendation, adapted to various recommendation scenarios, and utilized to achieve diverse recommendation objectives.

The organization of the paper is structured to guide the reader through the landscape of GNN-based recommendation systems methodically. It will commence with a detailed review of recent research papers that have made significant contributions to the field, discussing their methodologies, innovations, and the impact of their findings. Subsequent sections will highlight key findings, emerging trends, challenges, and potential avenues for future research in GNN-based recommendation systems. Through this exploration, the paper aims to not only illuminate the current state of the art but also to inspire further innovation and research in this dynamic and rapidly evolving field.

This introductory exploration sets the stage for a deeper dive into the transformative impact of GNNs on recommendation systems, promising a journey through the latest in graph-based recommendation methodologies, their applications, and the future they portend for personalized recommendation systems.

- Got acceptance from South Asian Research Center for present for the conference.

ACCEPETANCE AND REGISTRATION FORM (03rd May-04th May 2024 at Chongqing,China.)

External Inbox ×

SARC Conferences <sarc.net.in@gmail.com> Tue, Apr 16, 12:18 PM

to me

Dear Researcher,

Many Congratulations to you!!
We are happy to inform you that your paper has been selected for the upcoming conference organized by

South Asian Research Center- SARC.

We will be happy to guide during your registration and answer if you have any query. Please call us Mob: **+91-8280047487 (Call/Whatsapp)** for any assistance or email us--

Kindly find the attached file:

--

Thanks & Regards
Conference Co-ordinator,
South Asian Research Center(SARC)
Call/Whatsapp: +91-8280047487
Email: sarc.net.in@gmail.com (Mail your queries related to Conference, Registration, Venue etc.)
Website: http://www.sarc.net.in/

LEGAL, ETHICAL, SOCIAL & PROFFESIONAL ISSUES

LEGAL

Compliance: Adhered to GDPR and other relevant data protection laws to ensure user data confidentiality and integrity.

Intellectual Property: Utilized open-source tools under GPL and LGPL licenses and followed proper citation practices.

ETHICAL

Data Anonymity: Ensured all data used was anonymized to protect participant privacy.

Informed Consent: Obtained informed consent for all data used in the project.

Harm Prevention:

PROFFESIONAL

Transparency: Maintained honesty and transparency in presenting research findings.

Responsibility: Took responsibility for the system's impact on end-users, adhering to the BCS Code of Conduct.

SOCIAL

Environmental Impact: Minimized environmental footprint by optimizing algorithms for energy efficiency and using sustainable cloud services.

Contribution to

FUTURE WORK

- Implement real-time data processing capabilities to dynamically update recommendations based on user interactions.
- Develop and integrate more efficient data handling and processing methods to enhance the scalability of the system.
- Simplify the integration process of GNNs into existing systems by developing modular, plug-and-play components.

RESEARCH CONTRIBUTION

Technical Contribution

- How Graph Neural Networks (GNNs) can be utilized in e-commerce to enhance the accuracy and personalization of product recommendations.
- Integration of GNNs with K-means clustering for improved customer segmentation.

RESEARCH CONTRIBUTION

Research Contribution

- Methodological Advancement: Introduces a significant shift by using latent vectors of items within sessions to generate recommendations, moving away from traditional user-based representations.

Domain Contribution

- Targets e-commerce applications, specifically the challenges of providing personalized recommendations without full user profiles during short sessions.

DIFFICULTIES

- Data Quality and Availability
- Computational Demands
- Real-Time Data Processing
- User Privacy and Data Security
- Time Management Difficulties

REFERENCES

Paper Categories	Paper Count (Read)
• Graph Neural Network in Different recommendation stages.	24 papers
• Graph Neural Network in Different Recommendation Scenarios	65 papers
• Graph Neural Network in Different Recommendation Objectives	21 papers

Thanks For Listing