

A
Major Project
On
**CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR
EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE**
(Submitted in partial fulfillment of the requirements for the award of Degree)

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

BY

B. VINEETH

(167R1A0514)

K. DURGA PRADEEP

(167R1A0530)

N B VISHWA TEJA

(157R1A0540)

Under the Guidance of

Mr. KOTHA MAHESH

(Assistant Professor)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
CMR TECHNICAL CAMPUS
UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)
Recognized Under Section 2(f) & 12(B) of the UGC Act.1956,
Kandlakoya (V), Medchal Road, Hyderabad-501401.
2016-2021

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

This is to certify that the project entitled “**CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE**” being submitted by **B. Vineeth, K. Durga Pradeep, N B Vishwa Teja** bearing the **167R1A0514, 167R1A0530, 157R1A0540** roll number in partial fulfillment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering of the Jawaharlal Nehru Technological University Hyderabad, during the year 2021-22. It is certified that they have completed the project satisfactorily.

INTERNAL GUIDE
Kotha Mahesh
Assistant Professor

DIRECTOR
Dr. A. Raji Reddy

Head of Department
Dr. Srujan Raju

EXTERNAL EXAMINER

Submitted for viva voce Examination held on _____

ACKNOWLEDGEMENT

Apart from the efforts of us, the success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project. We take this opportunity to express my profound gratitude and deep regard to my guide

Mr. Kotha Mahesh , Assistant Professor for his exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by him shall carry us a long way in the journey of life on which we are about to embark.

We also take this opportunity to express a deep sense of gratitude to Project Review Committee (PRC) Coordinators: **Mr. B P Deepak Kumar, Mr. J. Narasimha Rao** and **Ms. V. Sandya** for their cordial support, valuable information and guidance, which helped us in completing this task through various stages.

We are also thankful to the Head of the Department **Dr. K. Srujan Raju** for providing excellent infrastructure and a nice atmosphere for completing this project successfully.

We are obliged to our Director **Dr. A. Raji Reddy** for being cooperative throughout the course of this project. We would like to express our sincere gratitude to our Chairman **Sri. Ch. Gopal Reddy** for his encouragement throughout the course of this project

The guidance and support received from all the members of **CMR TECHNICAL CAMPUS** who contributed and who are contributing to this project, was vital for the success of the project. We are grateful for their constant support and help.

Finally, we would like to take this opportunity to thank our family for their constant encouragement without which this assignment would not be possible. We sincerely acknowledge and thank all those who gave support directly and indirectly in completion of this project.

B Vineeth (167R1A0514)

K durga Pradeep (167R1A0530)

N B Vishwa Teja (157R1A0540)

TABLE OF CONTENTS

	Page No's
ABSTRACT	i
LIST OF FIGURES	ii
LIST OF SCREENSHOTS	iii
LIST OF TABLES	iv
1. INTRODUCTION	1
1.1 INPUT DESIGN	4
1.2 OBJECTIVES	4
1.3 OUTPUT DESIGN	5
2. SYSTEM ANALYSIS	6
2.1 PROBLEM DEFINITION	6
2.2 EXISTING SYSTEM	6
2.2.1 LIMITATIONS OF EXISTING SYSTEM	7
2.3 PROPOSED SYSTEM	7
2.3.1 ADVANTAGES	7
2.4 FEASIBILITY STUDY	8
2.4.1 ECONOMIC FEASIBILITY	8
2.4.2 TECHNICAL FEASIBILITY	8
2.4.3 BEHAVIORAL FEASIBILITY	8
2.5 HARDWARE & SOFTWARE REQUIREMENTS	9
2.5.1 HARDWARE REQUIREMENTS	9
2.5.2 SOFTWARE REQUIREMENTS	9
3. ARCHITECTURE	10
3.1 PROJECT ARCHITECTURE	10
3.2 DATA FLOW DIAGRAM	10
3.3 MODULES DESCRIPTION	12
3.3.1 UPLOAD PRODUCTS	15
3.3.2 PRODUCT REVIEW BASED ORDER	12
3.3.3 RATINGS AND REVIEWS	13
3.3.4 DATA ANALYSIS	13

3.4 UML DIAGRAMS	13
3.5 USE CASE DIAGRAMS	13
3.6 CLASS DIAGRAM	15
3.7 SEQUENCE DIAGRAM	16
3.8 ACTIVITY DIAGRAM	17
4. IMPLEMENTATION	19
4.1 SAMPLE CODE	19
5. SCREEN SHOTS	21
6. TESTING	26
6.1 INTRODUCTION TO TESTING	26
6.2 TYPES OF TESTING	27
6.2.1 UNIT TESTING	27
6.2.2 INTEGRATION TESTING	27
6.2.3 FUNCTIONAL TESTING	27
6.2.4 MODULE TESTING	28
6.2.5 SUB-SYSTEM TESTING	28
6.2.6 SYSTEM TESTING	28
6.2.7 ACCEPTANCE TESTING	28
6.3 TEST CASES	29
7. CONCLUSION	30
8. BIBLIOGRAPHY	31
8.1 REFERENCES	31

ABSTRACT

Online reviews have become an important source of information for users before making an informed purchase decision. Early reviews of a product tend to have a high impact on the subsequent product sales. In this paper, we take the initiative to study the behavior characteristics of early reviewers through their posted reviews on two real-world large e-commerce platforms, i.e., Amazon and Yelp. In specific, we divide product lifetime into three consecutive stages, namely early, majority and laggards. A user who has posted a review in the early stage is considered as an early reviewer. We quantitatively characterize early reviewers based on their rating behaviors, the helpfulness scores received from others and the correlation of their reviews with product popularity. We have found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our analysis of product reviews also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. By viewing the review posting process as a multiplayer competition game, we propose a novel margin-based embedding model for early reviewer prediction. Extensive experiments on two different e-commerce datasets have shown that our proposed approach outperforms a number of competitive baselines.

LIST OF FIGURES

FIGURE NO.	FIGURE NAME	PAGE NO
Fig. 3.1	Project Architecture	10
Fig. 3.2	Data Flow Diagram	11
Fig. 3.5.1	Use Case Diagram for User	14
Fig. 3.6	Class Diagram	15
Fig. 3.7	Sequence Diagram	16
Fig. 3.8.1	Activity Diagram for User	17
Fig. 3.8.2	Activity Diagram for Admin	17
Fig. 3.8.3	Activity Diagram for E-Commerce	18

LIST OF SCREENSHOTS

SCREENSHOT NO.	SCREENSHOT NAME	PAGE NO
5.1	USER LOGIN PAGE	21
5.2	VENDOR PRODUCT UPLOAD PAGE	21
5.3	USER/ADMIN PRODUCTS VIEW PAGE	22
5.4	USER/ADMIN PRODUCT'S REVIEW & RATING PAGE	22
5.5	USER PRODUCT RATING PAGE	23
5.6	SENTIMENTAL ANALYSIS PAGE	23
5.7	GRAPHICAL ANALYSIS PAGE	24
5.8	SENTIMENTAL ANALYSIS PAGE	24
5.9	GRAPHICAL ANALYSIS PAGE	25

LIST OF TABLES

TABLE NO.	TABLE NAME	PAGE NO
Table. 6.3.1	PRODUCT UPLOAD	29
Table. 6.3.2	User Rating Page	29
Table. 6.3.3	Analysis page	29

1. INTRODUCTION

The emergence of e-commerce websites has enabled users to publish or share purchase experiences by posting product reviews, which usually contain useful opinions, comments and feedback towards a product. As such, a majority of customers will read online reviews before making an informed purchase decision. It has been reported that 71% of global online shoppers read online reviews before purchasing a product. Product reviews, especially the early reviews (i.e., the reviews posted in the early stage of a product), have a high impact on subsequent product sales. We call the users who posted the early reviews early reviewers. Although early reviewers contribute only a small proportion of reviews, their opinions can determine the success or failure of new products and services. It is important for companies to identify early reviewers since their feedback can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products. For this reason, early reviewers become the emphasis to monitor and attract at the early promotion stage of a company. The pivotal role of early reviews has attracted extensive attention from marketing practitioners to induce consumer purchase intentions. For example, Amazon, one of the largest e-commerce companies in the world, has advocated the Early Reviewer Program¹, which helps to acquire early reviews on products that have few or no reviews. With this program, Amazon shoppers can learn more about products and make smarter buying decisions. As another related program, Amazon Vine² invites the most trusted reviewers on Amazon to post opinions about new and prerelease items to help their fellow customers make informed purchase decisions. Based on the above discussions, we can see that early reviewers are extremely important for product marketing. Thus, in this paper, we take the initiative to study the behavior characteristics of early reviewers through their posted reviews on representative e-commerce platforms, e.g., Amazon and Yelp. We aim to conduct effective analysis and make accurate predictions on early reviewers. This problem is strongly related to the adoption of innovations. In a generalized view, the review posting process can be considered as an adoption of innovations, which is a theory that seeks to explain how, why, and at what rate new ideas and technology spread. The analysis and detection of early adopters in the diffusion of innovations have attracted much attention from the research community. Three fundamental elements of a diffusion process have been studied: attributes of an innovation, communication channels, and social network structures. However, most of these studies are theoretical analysis at the macro level and there is a lack of quantitative investigations. With the rapid growth of online social platforms and the availability of a high volume of social networking data, studies of the diffusion of innovations have been widely conducted on social networks. However, in many application domains, social networking links or communication channels are unobserved. Hence, existing methods relying on social network structures or communication channels are not suitable in our current problem of

predicting early reviewers from online reviews. To model the behaviors of early reviewers, we develop a principled way to characterize the adoption process in two real-world large review datasets, i.e., Amazon and Yelp. More specially, given a product, the reviewers are sorted according to their timestamps for publishing their reviews. Following [8], we divide the product lifetime into three consecutive stages, namely early, majority and laggards. A user who has posted a review in the early stage is considered as an early reviewer. In our work here, we mainly focus on two tasks, the first task is to analyze the overall characteristics of early reviewers compared with the majority and laggard reviewers. We characterize their rating behaviors and the helpfulness scores received from others and the correlation of their reviews with product popularity. The second task is to learn a prediction model which predicts early reviewers given a product. To analyze the characteristics of early reviewers, we take two important metrics associated with their reviews, i.e., their review ratings and helpfulness scores assigned by others. We have found that an early reviewer tends to assign a higher average rating score to products; and an early reviewer tends to post more helpful reviews. Our above findings can find relevance in the classic principles of personality variables theory from social science, which mainly studies how innovation is spread over time among the participants. Earlier adopters have a more favorable attitude toward changes than later adopters; and earlier adopters have a higher degree of opinion leadership than later adopters. We can relate our findings with the personality variables theory as follows: higher average rating scores can be considered as the favorable attitude towards the products, and higher helpfulness votes of early reviews given by others can be viewed as a proxy measure of the opinion leadership. Our analysis also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. We further explain this finding with the herd behavior widely studied in economics and sociology. Herd behavior refers to the fact that individuals are strongly influenced by the decisions of others. To predict early reviewers, we propose a novel approach by viewing the review posting process as a multiplayer competition game. Only the most competitive users can become the early reviewers w.r.t. to a product. The competition process can be further decomposed into multiple pairwise comparisons between two players. In a two-player competition, the winner will beat the loser with an earlier timestamp. Inspired by the recent progress in distributed representation learning we propose to use a margin-based embedding model by first mapping both users and products into the same embedding space, and then determining the order of a pair of users given a product based on their respective distance to the product representation. Previous studies have highly emphasized the phenomenon that individuals are strongly influenced by the decisions of others, which can be explained by herd behavior. The influence of early reviews on subsequent purchase can be understood as a special case of herding effect. Early reviews contain important product evaluations from previous adopters, which are valuable reference resources for subsequent purchase decisions. As shown in [19], when consumers use the product evaluations of others to estimate product quality on the Internet, herd behavior occurs in the online shopping process. Different from existing studies on herd behavior, we focus on quantitatively analyzing the overall characteristics of early reviewers using large-scale real-world datasets. In addition, we formalize the early reviewer prediction

task as a competition problem and propose a novel embedding based ranking approach to this task. To our knowledge, the task of early reviewer prediction itself has received very little attention in the literature. Our contributions are summarized as follows:

We present a first study to characterize early reviewers on an e-commerce website using two real-world large datasets.

- We quantitatively analyze the characteristics of early reviewers and their impact on product popularity. Our empirical analysis provides support to a series of theoretical conclusions from sociology and economics.
- We view the review posting process as a multiplayer competition game and develop an embedding-based ranking model for the prediction of early reviewers. Our model can deal with the cold-start problem by incorporating side information of products.
- Extensive experiments on two real-world large datasets, i.e. Amazon and Yelp have demonstrated the effectiveness of our approach for the prediction of early reviewers.

The emergence of e-commerce websites has enabled users to publish or share purchase experiences by posting product reviews, which usually contain useful opinions, comments and feedback towards a product. As such, a majority of customers will read online reviews before making an informed purchase decision . It has been reported that 71% of global online shoppers read online reviews before purchasing a product . Product reviews, especially the early reviews (i.e., the reviews posted in the early stage of a product), have a high impact on subsequent product sales . We call the users who posted the early reviews early reviewers. Although early reviewers contribute only a small proportion of reviews, their opinions can determine the success or failure of new products and services . It is important for companies to identify early reviewers since their feedback can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products. For this reason, early reviewers become the emphasis to monitor and attract at the early promotion stage of a company. The pivotal role of early reviews has attracted extensive attention from marketing practitioners to induce consumer purchase intentions . For example, Amazon, one of the largest e-commerce companies in the world, has advocated the Early Reviewer Program¹, which helps to acquire early reviews on products that have few or no reviews. With this program, Amazon shoppers can learn more about products and make smarter buying decisions. As another related program, Amazon Vine² invites the most trusted reviewers on Amazon to post opinions about new and prerelease items to help their fellow customers make informed purchase decisions.

1.1 INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps necessary to put transaction data into a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining privacy. Input Design considered the following things:

- What data should be given as input?
- How should the data be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

1.2 OBJECTIVES

- Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system
- It is achieved by creating user-friendly screens for the data entry to handle large volumes of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data can be performed. It also provides record viewing facilities.
- When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize instant. Thus the objective of input design is to create an input layout that is easy to follow

1.3 OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other systems through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source of information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

- Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use it easily and effectively. When analyzing computer output, they should Identify the specific output that is needed to meet the requirements.
- Select methods for presenting information.
- Create documents, reports, or other formats that contain information produced by the system.

2. SYSTEM ANALYSIS

System Analysis is the important phase in the system development process. The System is studied to the minute details and analyzed. The system analyst plays an important role of an interrogator and dwells deep into the working of the present system. In analysis, a detailed study of these operations performed by the system and their relationships within and outside the system is done. A key question considered here is, “what must be done to solve the problem?” The system is viewed as a whole and the inputs to the system are identified. Once analysis is completed the analyst has a firm understanding of what is to be done.

2.1 PROBLEM DEFINITION

A detailed study of the process must be made by various techniques like interviews, questionnaires etc. The data collected by these sources must be scrutinized to arrive at a conclusion. The conclusion is an understanding of how the system functions. This system is called the existing system. Now the existing system is subjected to close study and problem areas are identified. The designer now functions as a problem solver and tries to sort out the difficulties that the enterprise faces. The solutions are given as proposals. The proposal is then weighed with the existing system analytically and the best one is selected. The proposal is presented to the user for an endorsement by the user. The proposal is reviewed on user request and suitable changes are made. This loop ends as soon as the user is satisfied with the proposal.

2.2 EXISTING SYSTEM

Previous studies have highly emphasized the phenomenon that individuals are strongly influenced by the decisions of others, which can be explained by herd behavior. The influence of early reviews on subsequent purchase can be understood as a special case of herding effect. Early reviews contain important product evaluations from previous adopters, which are valuable reference resources for subsequent purchase decisions. As shown in, when consumers use the product evaluations of others to estimate product quality on the Internet, herd behavior occurs in the online shopping process. Different from existing studies on herd behavior, we focus on quantitatively analyzing the overall characteristics of early reviewers using large-scale real-world datasets. In addition, we formalize the early reviewer prediction task as a competition problem and propose a novel embedding based ranking approach to this task. To our knowledge, the task of early reviewer prediction itself

has received very little attention in the literature. Our contributions are summarized as follows: We present a first study to characterize early reviewers on an e-commerce website using two real-world large datasets. We quantitatively analyze the characteristics of early reviewers and their impact on product popularity. Our empirical analysis provides support to a series of theoretical conclusions from sociology and economics. We view the review posting process as a multiplayer competition game and develop an embedding-based ranking model for the prediction of early reviewers. Our model can deal with the cold-start problem by incorporating side information of products. Extensive experiments on two real-world large datasets, i.e. Amazon and Yelp have demonstrated the effectiveness of our approach for the prediction of early reviewers.

2.2.1 LIMITATIONS OF EXISTING SYSTEM

- By using herding effect not all aspects for predicting are covered.
- Thus the efficiency of the result is low

2.3 PROPOSED SYSTEM

To predict early reviewers, we propose a novel approach by viewing the review posting process as a multiplayer competition game. Only the most competitive users can become the early reviewer's w.r.t. to a product. The competition process can be further decomposed into multiple pairwise comparisons between two players. In a two-player competition, the winner will beat the loser with an earlier timestamp. Inspired by the recent progress in distributed representation learning, we propose to use a margin-based embedding model by first mapping both users and products into the same embedding space, and then determining the order of a pair of users given a product based on their respective distance to the product representation.

2.3.1 ADVANTAGES

- Predicting the Product quality by viewing the review posting process as a multiplayer competition game.
- Thus efficiency is high
- This Process provides the best method compared to existing system

2.4 FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company.

Three key considerations involved in the feasibility analysis are

- Economic Feasibility
- Technical Feasibility
- Social Feasibility

2.4.1 ECONOMIC FEASIBILITY

The developing system must be justified by cost and benefit. Criteria to ensure that effort is concentrated on a project, which will give best, return at the earliest. One of the factors, which affect the development of a new system, is the cost it would require. The following are some of the important financial questions asked during preliminary investigation:

- The costs conduct a full system investigation.
- The cost of the hardware and software.
- The benefits in the form of reduced costs or fewer costly errors.

Since the system is developed as part of project work, there is no manual cost to spend for the proposed system. Also all the resources are already available, it give an indication of the system is economically possible for development.

2.4.2 TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

2.4.3 BEHAVIORAL FEASIBILITY

This includes the following questions:

- Is there sufficient support for the users?
- Will the proposed system cause harm?

The project would be beneficial because it satisfies the objectives when developed and installed. All behavioural aspects are considered carefully and conclude that the project is behaviourally feasible.

2.5 HARDWARE & SOFTWARE REQUIREMENTS

2.5.1 HARDWARE REQUIREMENTS:

Hardware interfaces specifies the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements.

- System : Intel I3
- Hard Disk : 120 GB.
- Monitor : 15” LED
- Input Devices : Keyboard, Mouse
- Ram : 1GB.

2.5.2 SOFTWARE REQUIREMENTS

Software Requirements specifies the logical characteristics of each interface and software components of the system. The following are some software requirements.

- Operating system : Windows 10
- Coding Language : Python
- Tool : Anaconda

3. ARCHITECTURE

3.1 PROJECT ARCHITECTURE

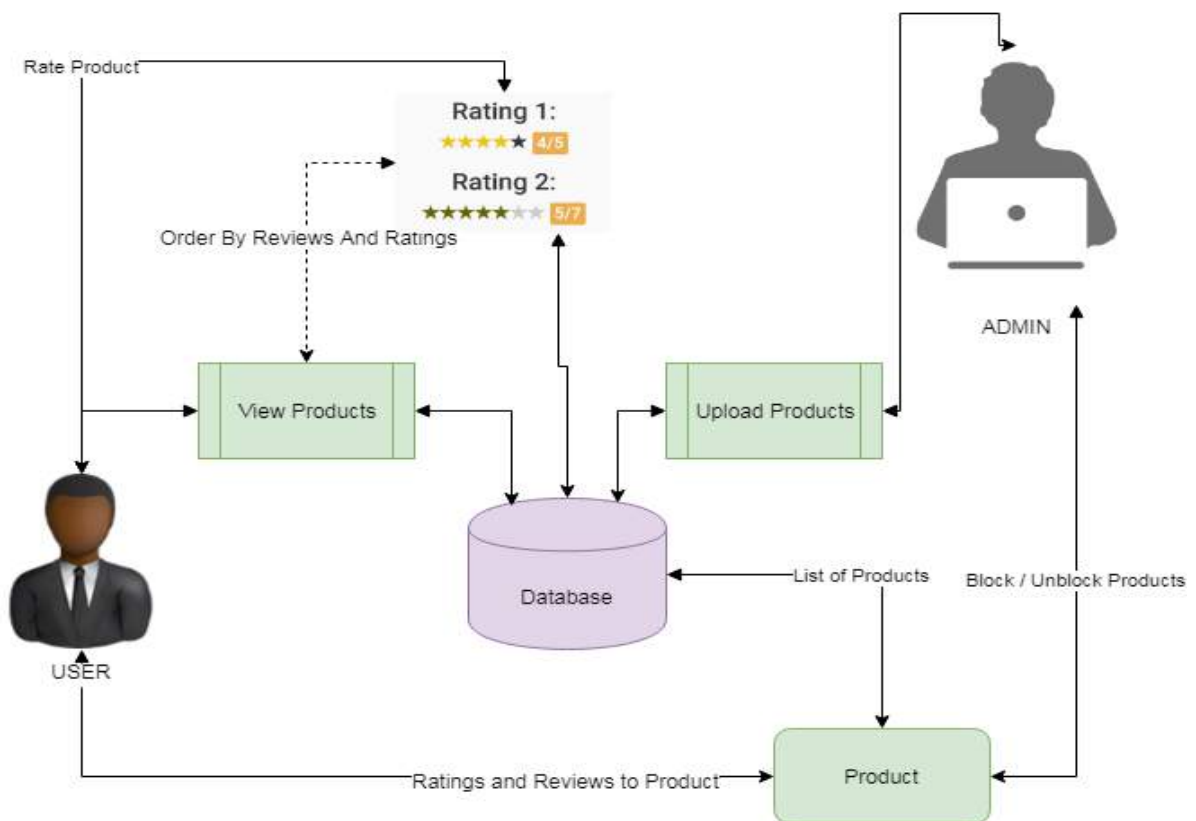


Fig 3.1 Project Architecture

3.2 DATA FLOW DIAGRAM

- The DFD is also called a bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

- The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
- DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.
- DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

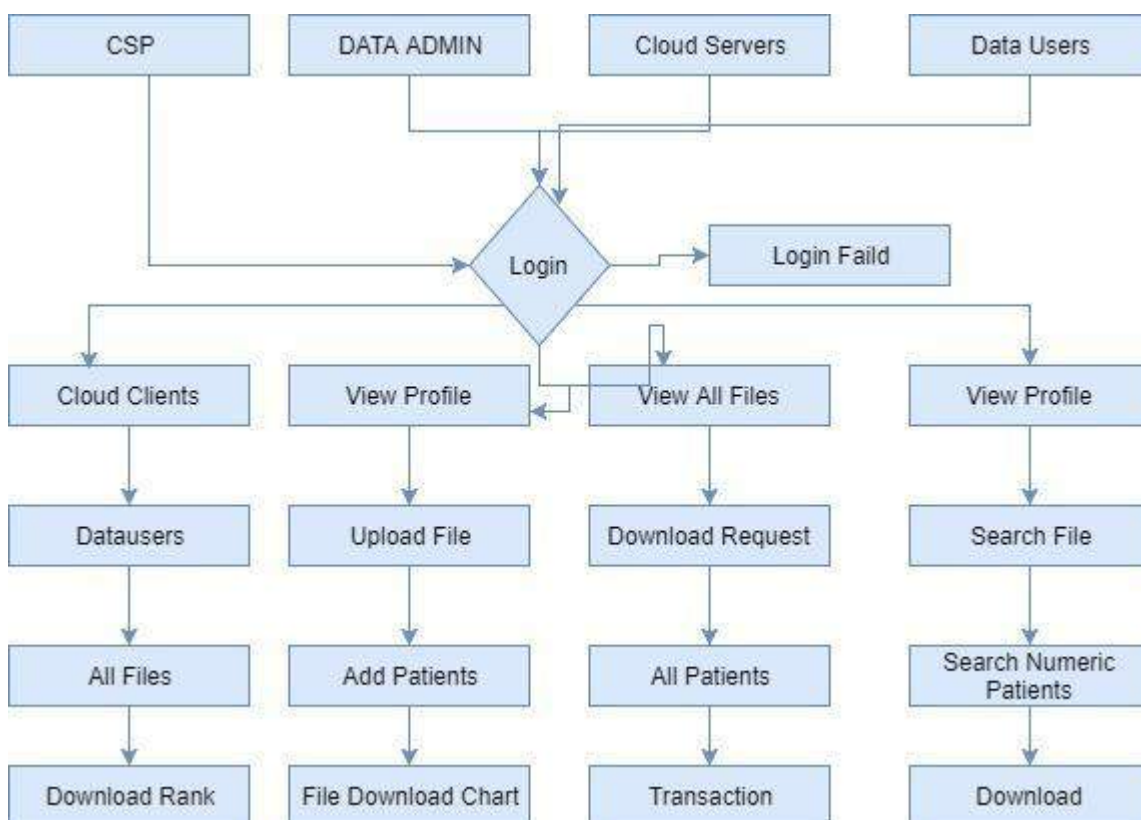


Fig 3.2 Data Flow Diagram

3.3 MODULES DESCRIPTION

Modules

- Upload Products
- Product Review Based Order
- Ratings And Reviews
- Data Analysis

3.3.1 UPLOAD PRODUCTS

Uploading the products is done by admin. Authorized person is uploading the new arrivals to systems that are listed to users. Product can be uploaded with its attributes such as brand, color, and all other details of warranty. The uploaded products are able to be blocked or unblocked by users.

3.3.2 PRODUCT REVIEW BASED ORDER

The suggestion to the user's view of products is listed based on the review by user and rating to particular item. Naïve Bayes algorithm is used in this project to develop whether the sentiment of a given review is positive or negative. Based on the output of the algorithm, a suggestion to users is given. The algorithm is applied and lists the products on the user side based on the positive and negative.

3.3.3 RATINGS AND REVIEWS

Ratings and reviews are the main concept of the project in order to find effective product marketing. The main aim of the project is to get the user reviews based on how they purchased or whether they purchased or not. The major find out of the project is when they give the ratings and how effective it is. And this will be helpful for the users who are willing to buy the same kind of product.

3.3.4 DATA ANALYSIS

The main part of the project is to analyze the ratings and reviews that are given by the user. The products can be analyzed based on the numbers which are given by the user. The user data analysis of the data can be done by charts format. The graphs may vary like pie charts, bar charts or some other charts.

3.4 UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group. The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML consists of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML. The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software systems, as well as for business modeling and other non-software systems. The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems. The UML is a very important part of developing object oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

3.5 USE CASE DIAGRAMS

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case

the diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

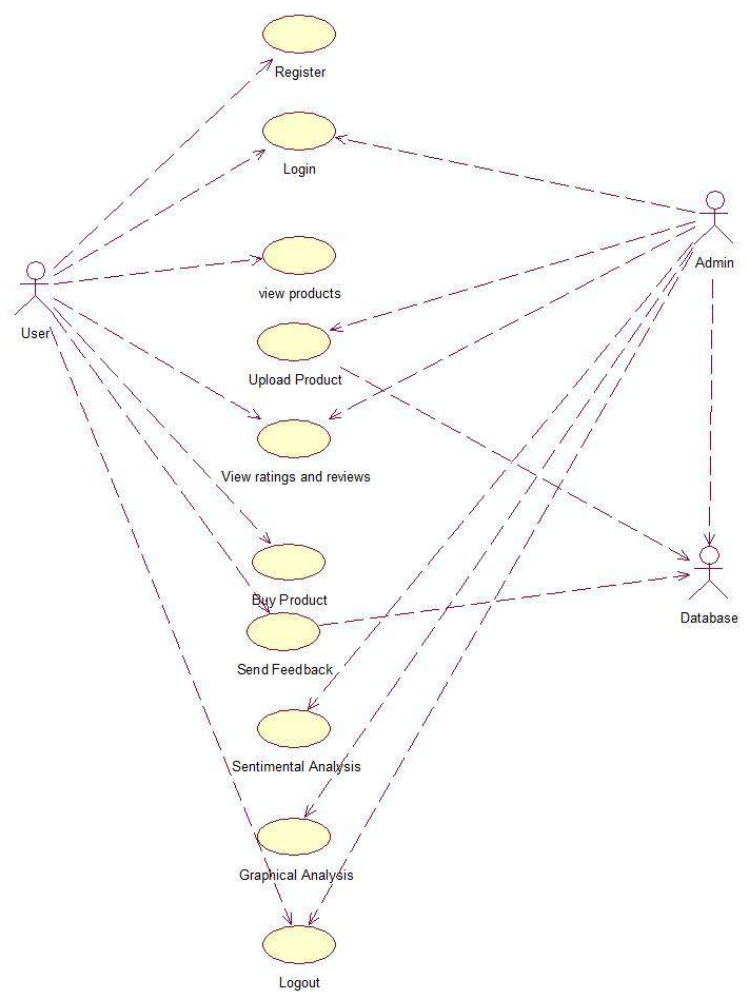


fig 3.5.1 use case diagram for Predicting Early reviewers

3.6 CLASS DIAGRAM

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

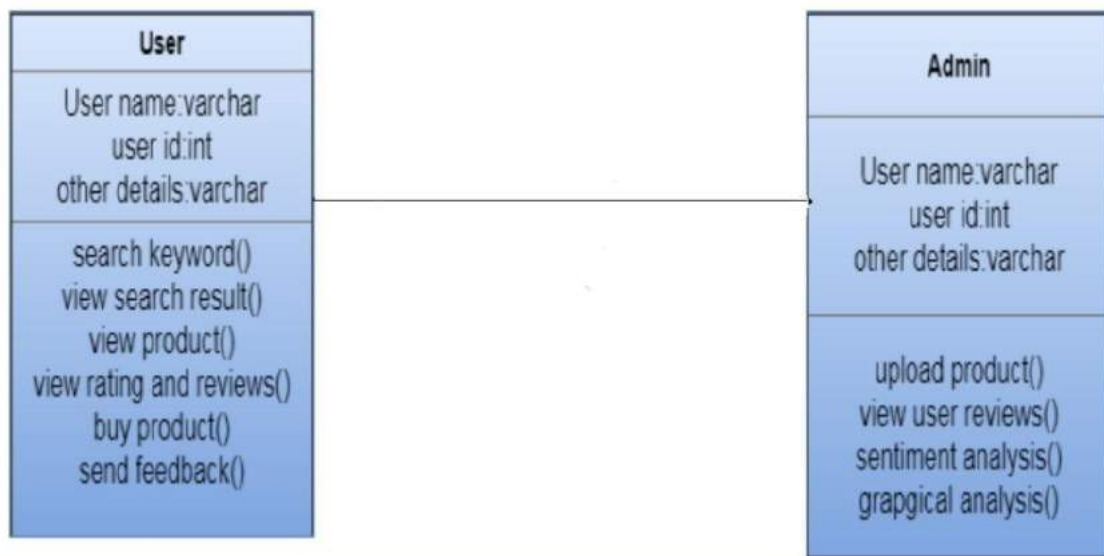


fig 3.6 Class Diagram

3.7 SEQUENCE DIAGRAM

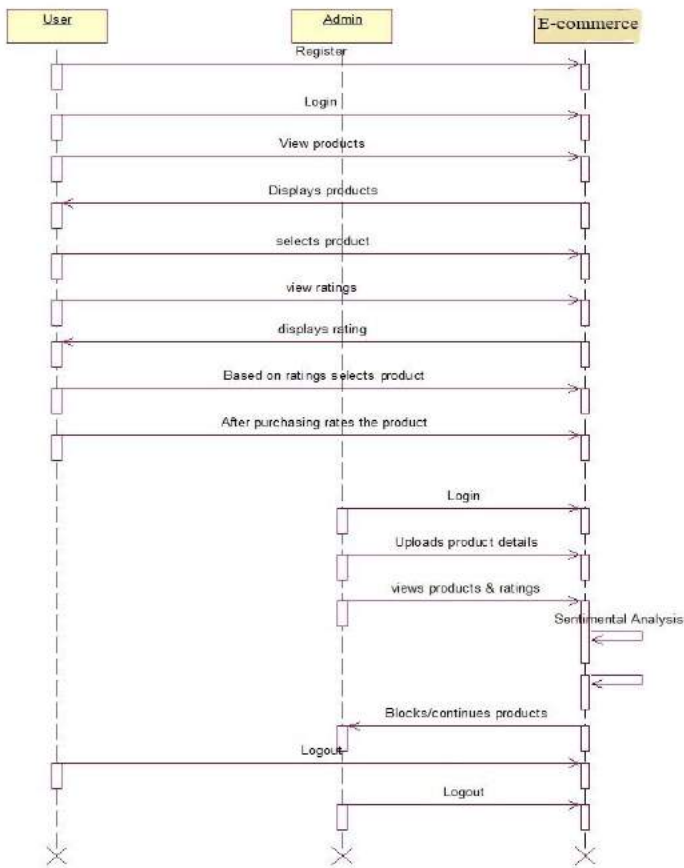


fig 3.6 Sequence diagram

3.8 Activity Diagram

Activity diagram is basically a flowchart to represent the flow from one activity to another activity. The activity can be described as an operation of the system.

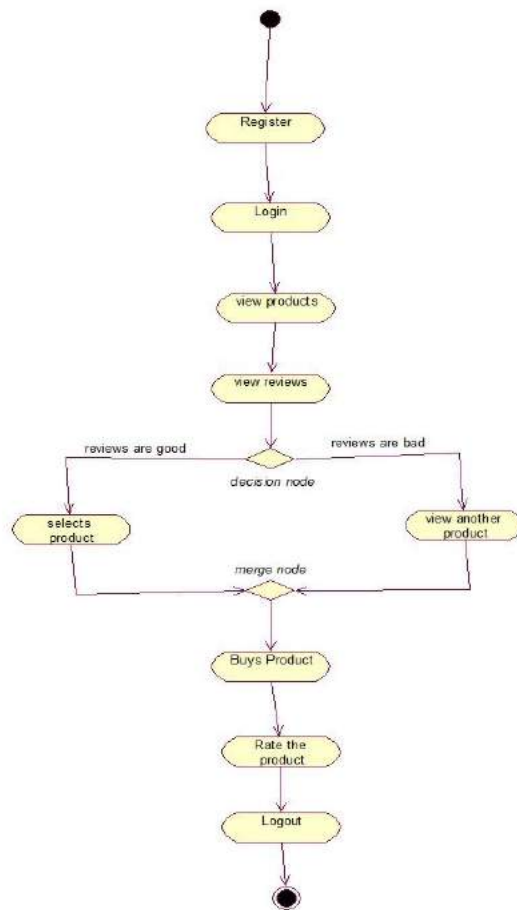


fig 3.8.1 Activity diagram for User

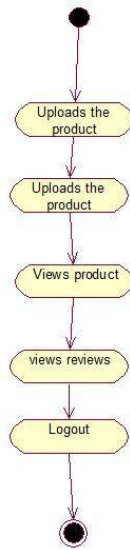


fig 3.8.2 Activity diagram for Admin

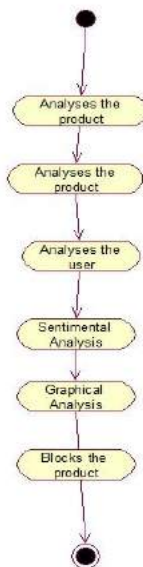


fig 3.8.3 Activity diagram for E-commerce

4. IMPLEMENTATION

4.1 SAMPLE CODE

```

from django.db.models import Count, Avg
from django.shortcuts import render, redirect, get_object_or_404

from admins.forms import UploadForm
from admins.models import Prodcuts
from user.models import Purchase, Feedback

def index(request):
    if request.method=="POST":
        username=request.POST.get('username','')
        password=request.POST.get('password','')
        if username=='admin' and password=='admin':
            request.session['userid']=1
            request.session['username']='admin'
            return redirect('admins:home')
    return render(request,'admins/index.html',)

def home(request):
    products=Prodcuts.objects.all()
    return render(request,'admins/home.html',{'products':products})

def uploadproducts(request):
    if request.method=="POST":
        forms=UploadForm(request.POST, request.FILES)
        if forms.is_valid():
            forms.save()
            return redirect('admins:home')
    else:
        forms = UploadForm()
    return render(request,'admins/uploadproducts.html',{'form':forms})

def charts(request,chart_type):
    d=None
    if chart_type=='all':
        d=Feedback.objects.values('product').annotate(dcount=Count('rating'))
    elif chart_type=='mobile':

```

```

                                d =
Feedback.objects.filter(product__product_name='mobile').values('product__vendor_name').a
nnotate(dcount=Count('rating'))
    elif chart_type=='laptop':
                                d =
Feedback.objects.filter(product__product_name='laptop').values('product__vendor_name').an
notate(
    dcount=Count('rating'))
    elif chart_type=='mobileaccessories':
                                d = Feedback.objects.filter(product__product_name='mobile
accessories').values('product__vendor_name').annotate(
    dcount=Count('rating'))
    elif chart_type=='watches':
                                d =
Feedback.objects.filter(product__product_name='watches').values('product__vendor_name').
annotate(
    dcount=Count('rating'))
    elif chart_type=='shoes':
                                d =
Feedback.objects.filter(product__product_name='shoes').values('product__vendor_name').an
notate(
    dcount=Count('rating'))
    return render(request,'admins/charts.html',{'chart_type':chart_type,'d':d})

def charts1(request,chart_type):
    d = Feedback.objects.values('user__profession').annotate(dcount=Count('rating'))
    return render(request,'admins/charts1.html',{'chart_type':chart_type,'d':d})

def charts2(request,chart_type):
    d=Feedback.objects.values('user__location').annotate(dcount=Count('rating'))
    return render(request,'admins/charts2.html',{'chart_type':chart_type,'d':d})

def charts3(request,chart_type):
    d1=Feedback.objects.filter(sentiment='positive').values('product__product_name').annotate(d
count=Count('sentiment'))

    d2=Feedback.objects.filter(sentiment='negative').values('product__product_name').annotate(
dcount=Count('sentiment'))
    return render(request,'admins/charts3.html',{'chart_type':chart_type,'d1':d1,'d2':d2})

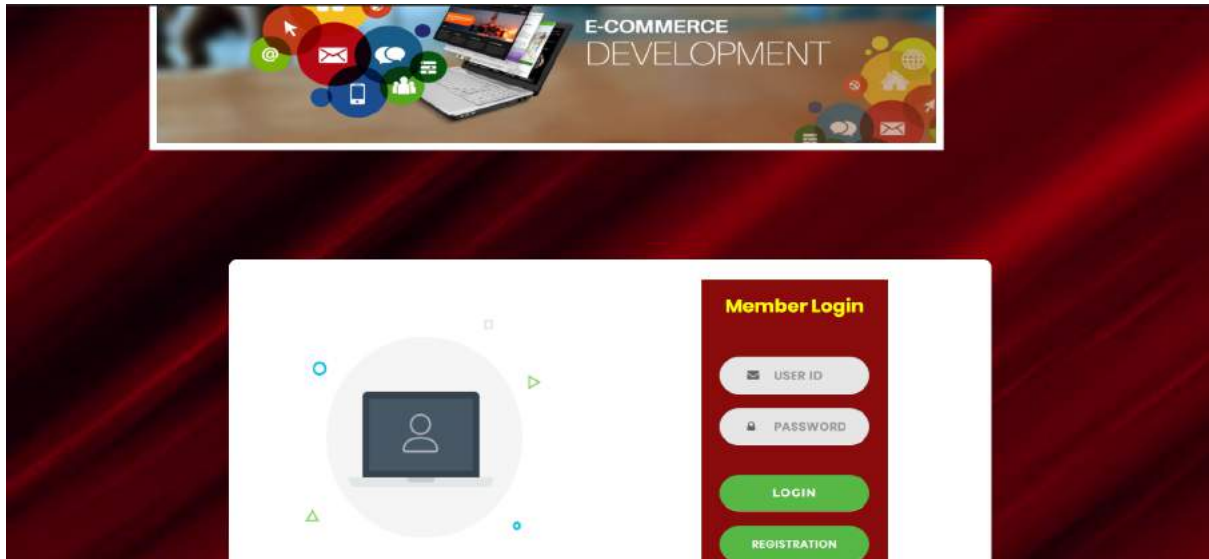
def logout(request):

    return redirect('admins:index')

```

5. SCREENSHOTS

5.1 USER LOGIN PAGE



5.2 VENDOR PRODUCT UPLOAD PAGE



5.3 USER/ADMIN PRODUCTS VIEW PAGE

CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON E-COMMERCE WEBSITES

home cart logout



Product Name: shirt
Vendor Name :Van Heusen
Color :green
Price :2000.0
Features :good product



Product Name: shirt
Vendor Name :Xiami Peter England
Color :white & Blue
Price :1870.0
Features :All good features available




Product Name: T-shirt
Vendor Name :Parx
Color :blue
Price :10000.0
Features :it is good cloth

5.4 USER/ADMIN PRODUCT'S REVIEW & RATING PAGE

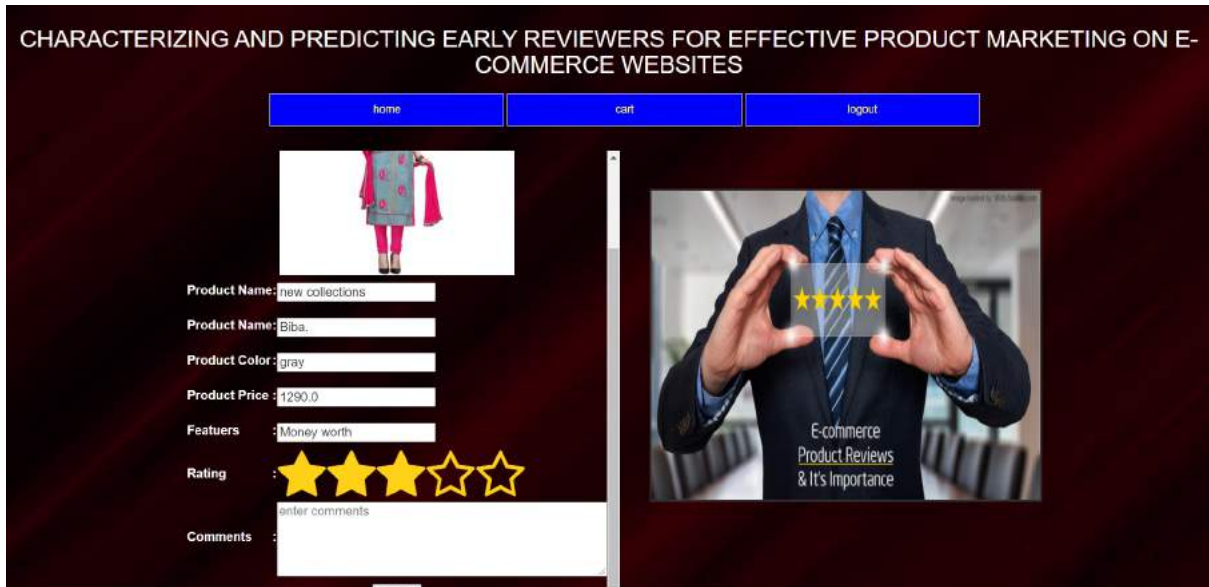
CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON E-COMMERCE WEBSITES

home cart logout

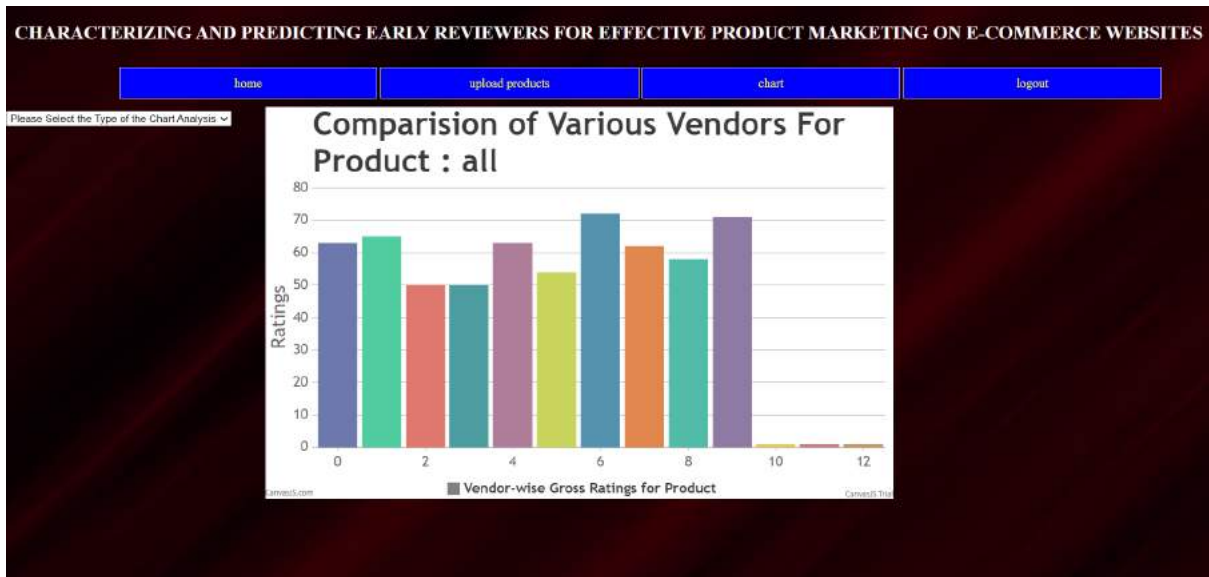
Customer Name	Product	Purchasing Status	Rating	Review
priyanka	sudithar-new collections-Biba.	purchased	4	great purchase
navaneetha	sudithar-new collections-Biba.	not purchased	3	good
vinay	sudithar-new collections-Biba.	purchased	2	It is worst
vinay	sudithar-new collections-Biba.	purchased	1	not good
vinay	sudithar-new collections-Biba.	purchased	3	poor purchase
vinay	sudithar-new collections-Biba.	not purchased	4	super
vinay	sudithar-new collections-Biba.	purchased	2	poor purchase



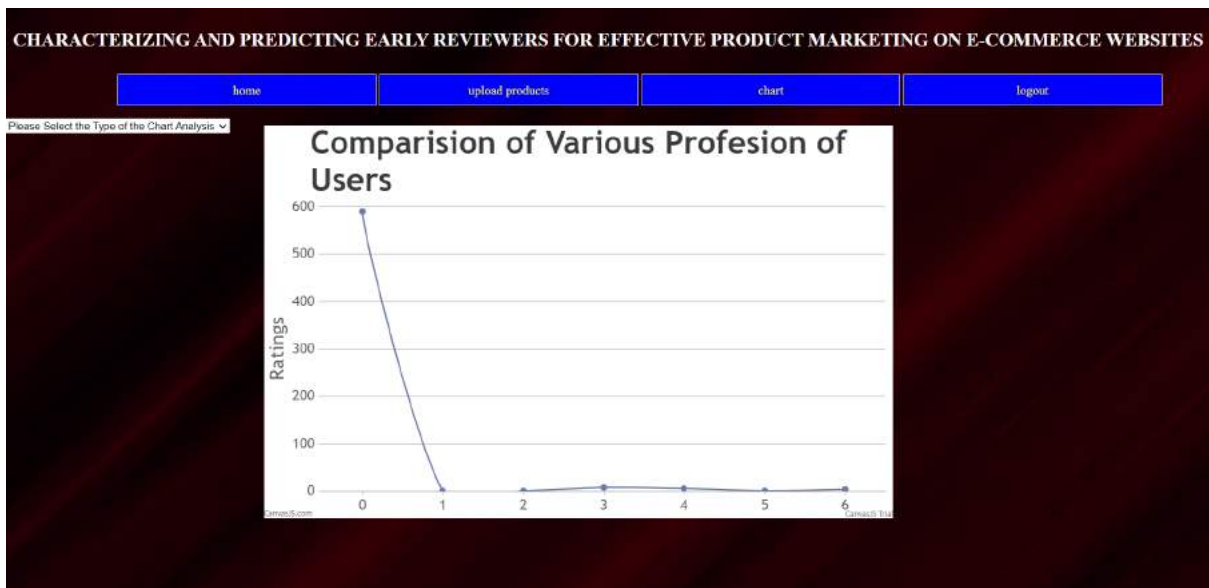
5.5 USER PRODUCT RATING PAGE



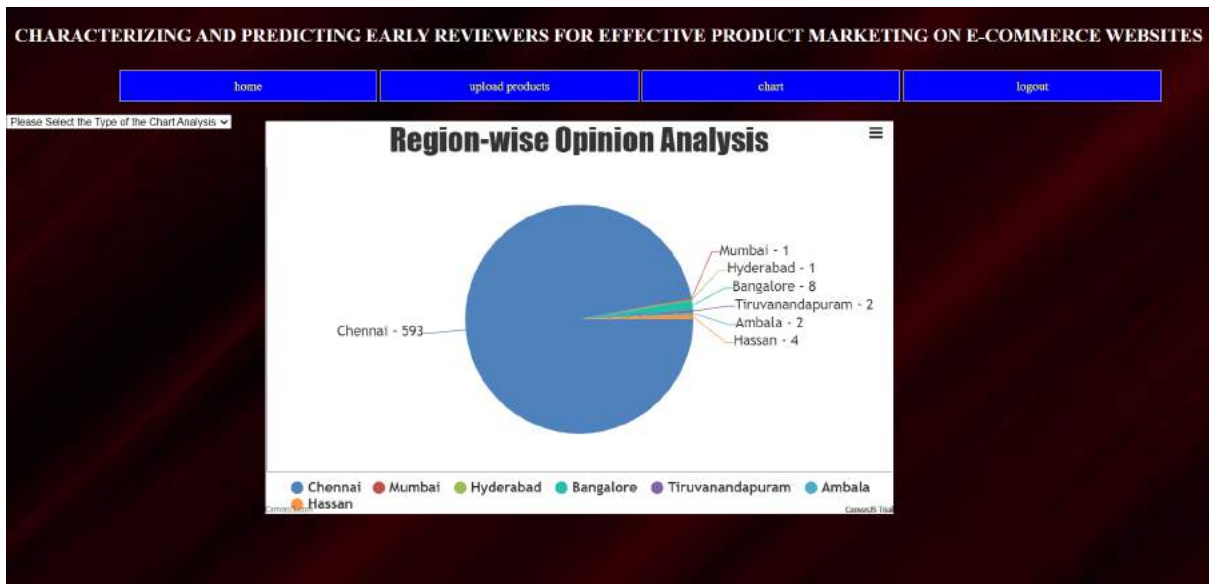
5.6 SENTIMENTAL ANALYSIS PAGE



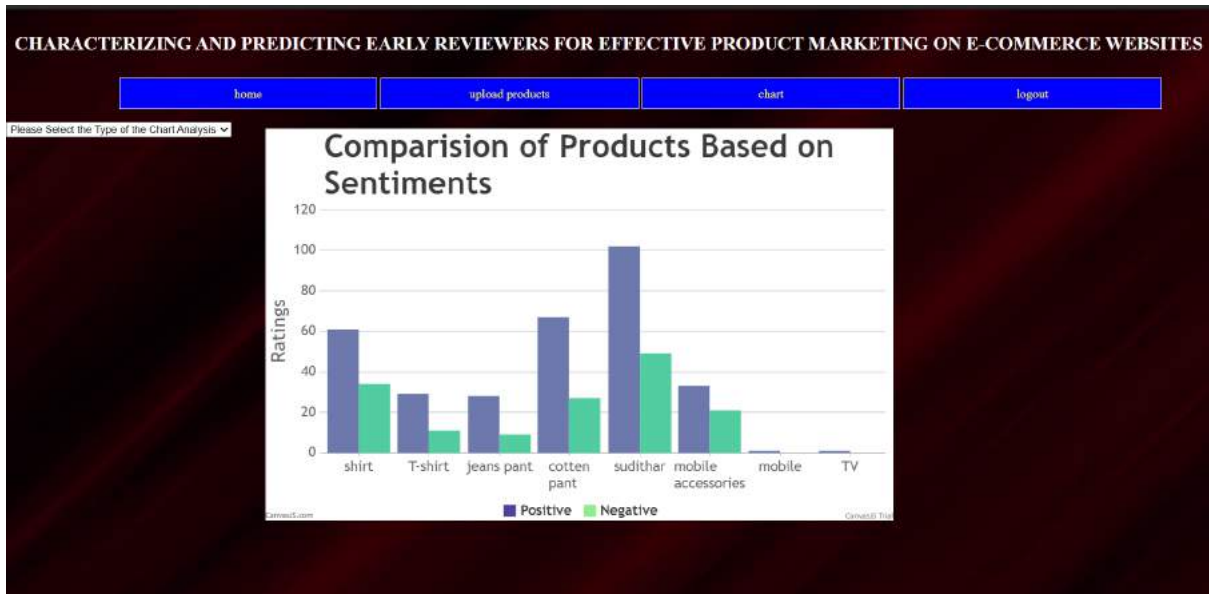
5.7 GRAPHICAL ANALYSIS PAGE



5.8 SENTIMENTAL ANALYSIS PAGE



5.9 GRAPHICAL ANALYSIS PAGE



6. TESTING

6.1 INTRODUCTION TO TESTING

Software testing is the process used to assess the quality of computer software. Software testing is an empirical technical investigation conducted to provide stakeholders with information about the quality of the product or service under test, with respective content of in which it is intended to operate. This includes, but is not limited to, the process of executing a program or application with the intent of finding software bugs. Testing can never completely establish the correctness of arbitrary computer software; testing furnishes a criticism or comparison that compares the state and the behavior of the product against a specification. An important point is that software testing should be distinguished from the separate discipline of Software Quality Assurance (S Q A), which encompasses all business process areas, not just testing. Software testing is a critical element of the software quality assurance and represents the ultimate review of specification, design and coding. Testing is the exposure of the system to trial input to see whether it produces correct output or not. Once the source code is generated, software must be tested to uncover as many errors as possible before delivery to the customer. So the main goal of testing is to design a series of test cases that have a high likelihood of finding errors. Except for small computer programs, systems should not be tested as a single monolithic unit. Large systems are built out of subsystems which are built out of modules, which are composed of procedures and functions. The testing process should therefore proceed in stages where testing is carried out incrementally in conjunction with system implementation.

SOFTWARE TESTING is defined as an activity to check whether the actual results match the expected results and to ensure that the software system is Defect free. It involves execution of a software component or system component to evaluate one or more properties of interest. Software testing also helps to identify errors, gaps or missing requirements in contrary to the actual requirements. It can be either done manually or using automated tools.

6.2 TYPES OF TESTING

6.2.1 UNIT TESTING

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

6.2.2 INTEGRATION TESTING

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfied, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

6.2.3 FUNCTIONAL TESTING

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals. Functional testing is centred on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.

Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to Business process flows; data fields, predefined processes.

6.2.4 Module Testing

A module is a collection of dependent components such as an object or some looser collection of procedures and functions. A module encapsulates related components so that they can be tested without other system modules.

6.2.5 Sub-system Testing

This phase involves testing collections of modules which have been integrated into sub-systems. may be independently designed and implemented and most common problems which arise in large software systems or sub-system interface mismatches.

6.2.6 System Testing

Subsystems are integrated to make up the entire system. The testing process is concerned with finding errors which normally result from an unanticipated interface between sub-systems and components. It is also concerned with validating that the system meets its functional and non-functional requirements .

6.2.7 Acceptance Testing

This is the final stage in the testing process before the system is accepted for operational use. It involves testing the system with the data applied by the system prowler rather than simulated data developed as part of the testing process. Acceptance testing of an reveals errors and emissions in the system requirements definition. The requirement may not reflect the actual facilities and performance required by the user and testing may demonstrate that the system doesn't exhibit the anticipated performance and functionality .

6.3 TEST CASES

6.3.1 TABLE: PRODUCT UPLOAD

S. No	Input	Expected output	Results
1	Product upload	Product is uploaded	Uploaded product is added

6.3.2 TABLE: User Rating Page

S. No	Input	Expected Output	Results
1	User gives the review to product	Rating is shown	Rating is uploaded

6.3.3 TABLE: Analysis page

S. No	Input	Expected Output	Results
1	It should Analyse the rating & reviews	Analysis Graphs are shown	Sentimental Analysis & Graphical Analysis is done

7. CONCLUSION

In this project, we have studied the novel task of early reviewer characterization and prediction on two real-world online review datasets. Our empirical analysis strengthens a series of theoretical conclusions from sociology and economics. We found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our experiments also indicate that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity at a later stage. We have adopted a competition-based viewpoint to model the review posting process, and developed a margin based embedding ranking model (MERM) for predicting early reviewers in a cold-start setting. In our current work, the review content is not considered. In the future, we will explore effective ways in incorporating review content into our early reviewer prediction model. Also, we have not studied the communication channel and social network structure in diffusion of innovations partly due to the difficulty in obtaining the relevant information from our review data. We will look into other sources of data such as Flixster in which social networks can be extracted and carry out more insightful analysis. Currently, we focus on the analysis and prediction of early reviewers, while there remains an important issue to address, i.e., how to improve product marketing with the identified early reviewers. We will investigate this task with real e-commerce cases in collaboration with e-commerce companies in the future. Our work is also related to the studies on mining review data [63], [64]. However, we focus on characterizing early reviews and detecting early reviewers, which is different from the existing works on extracting opinions or identifying opinion targets (or holders) from review data. To our knowledge, it is the first time that the task of early reviewer analysis and detection has been investigated on the real world e-commerce review datasets, i.e., Amazon and Yelp. We propose a novel margin-based embedding ranking model in a competition-based framework, which has never been adopted in early adopter detection. In addition, we extend the original competition-based framework by incorporating important side information about products. We also use a distributed representation approach to address the cold start problem. Our empirical analysis has confirmed a series of theoretical conclusions from the sociology and economics

8. BIBLIOGRAPHY

8.1 REFERENCES

- [1] J. McAuley and A. Yang, “Addressing complex and subjective product-related queries with customer reviews,” in WWW, 2016, pp. 625–635
- [2] N. V. Nielsen, “E-commerce: Evolution or revolution in the fastmoving consumer goods world,” nngroup. com, 2014.
- [3] W. D. J. Salganik M J, Dodds P S, “Experimental study of inequality and unpredictability in an artificial cultural market,” in ASONAM, 2016, pp. 529–532.
- [4] R. Peres, E. Muller, and V. Mahajan, “Innovation diffusion and new product growth models: A critical review and research directions,” *International Journal of Research in Marketing*, vol. 27, no. 2, pp. 91 – 106, 2010.
- [5] L. A. Fourt and J. W. Woodlock, “Early prediction of market success for new grocery products.” *Journal of Marketing*, vol. 25, no. 2, pp. 31 – 38, 1960.
- [6] B. W. O, “Reference group influence on product and brand purchase decisions,” *Journal of Consumer Research*, vol. 9, pp. 183–194, 1982.
- [7] .E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, “Detecting product review spammers using rating behaviors,” in CIKM, 2010, pp. 939–948.
- [8] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, “Liblinear: A library for large linear classification,” *Machine Learning Research*, vol. 9, pp. 1871–1874, 2008.
- [9] E.-P. Lim, V.-A. Nguyen, N. Jindal, B. Liu, and H. W. Lauw, “Detecting product review spammers using rating behaviors,” in CIKM, 2010, pp. 939–948

- [10] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2007.
- [11] .Liu, *Sentiment Analysis - Mining Opinions, Sentiments, and Emotions*. Cambridge University Press, 2015.
- [12] R. S. Z. Ryan Kiros, Ruslan Salakhutdinov, “Multimodal neural language models,” in *ICML*, 2014, pp. 595–603.
- [13] C. D. M. A. Y. N. Richard Socher, Danqi Chen, “Reasoning with neural tensor networks for knowledge base completion,” in *NIPS*, 2013, pp. 926–934..
- [14] M. Cattelan, “Models for paired comparison data: A review with emphasis on dependent data,” *Statistical Science*, vol. 27, no. 3, pp. 412–433, 2012.

8.2 GITHUB LINK

<https://github.com/vineeth14-dot/early-review-cmrhc-cse-19>

8.3 PAPER PUBLICATION LINK

<http://www.ijaema.com/CURRENT-ISSUE/>

CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE

Mr. Vineeth : ,UG Scholar, Dept of Computer Science & Engineering, CMR Technical Campus, UGC Autonomous, Kandlakoya(V), Medchal Road, Hyderabad-501401, INDIA .

Email ID: vineethbathula1999@gmail.com

Mr. Pradeep: UG Scholar, Dept of Computer Science & Engineering, CMR Technical Campus, UGC Autonomous, Kandlakoya(V), Medchal Road, Hyderabad-501401, INDIA .

Email ID: pradeepkodavati007@gmail.com

Mr. Vishwa Teja: UG Scholar, Dept of Computer Science & Engineering, CMR Technical Campus, UGC Autonomous, Kandlakoya(V), Medchal Road, Hyderabad-501401, INDIA .

Email ID: vishwateja610@gmail.com

Mr. Kotha Mahesh: Assistant Professor, Dept of Computer Science & Engineering, CMR Technical Campus, UGC Autonomous, Kandlakoya(V), Medchal Road, Hyderabad-501401, INDIA .

Email ID: Mahesh.cse@cmrtc.ac.in

ABSTRACT :

Online reviews have become an important source of information for users before making an informed purchase decision. Early reviews of a product tend to have a high impact on the subsequent product sales. In this paper, we take the initiative to study the

behavior characteristics of early reviewers through their posted reviews on two real-world large e-commerce platforms, i.e., Amazon and Yelp. In specific, we divide product lifetime into three consecutive stages, namely early, majority and laggards. A user who has posted a review in the early

stage is considered as an early reviewer. We quantitatively characterize early reviewers based on their rating behaviors, the helpfulness scores received from others and the correlation of their reviews with product popularity. We have found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our analysis of product reviews also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. By viewing the review posting process as a multiplayer competition game, we propose a novel margin-based embedding model for early reviewer prediction. Extensive experiments on two different e-commerce datasets have shown that our proposed approach outperforms a number of competitive baselines.

Keywords : Early reviewer, Early review, Embedding model.

I INTRODUCTION

The emergence of e-commerce websites has enabled users to publish or share purchase experiences by posting product reviews, which usually contain useful opinions, comments and feedback towards a product.

As such, a majority of customers will read online reviews before making an informed purchase decision [1]. It has been reported about 71% of global online shoppers read online reviews before purchasing a product [2]. Product reviews, especially the early reviews (i.e., the reviews posted in the early stage of a product), have a high impact on subsequent product sales [3]. We call the users who posted the early reviews early reviewers. Although early reviewers contribute only a small proportion of reviews, their opinions can determine the success or failure of new products and services [4], [5]. It is important for companies to identify early reviewers since their feedbacks can help companies to adjust marketing strategies and improve product designs, which can eventually lead to the success of their new products.

For this reason, early reviewers become the emphasis to monitor and attract at the early promotion stage of a company. The pivotal role of early reviews has attracted extensive attention from marketing practitioners to induce consumer purchase intentions [6]. For example, Amazon, one of the largest e-commerce company in the world, has advocated the Early Reviewer Program¹, which helps to acquire early reviews on products that have few or no reviews. With

this program, Amazon shoppers can learn more about products and make smarter buying decisions. As another related program, Amazon Vine² invites the most trusted reviewers on Amazon to post opinions about new and prerelease items to help their fellow customers make informed purchase decisions. Based on the above discussions, we can see that early reviewers are extremely important for product marketing. Thus, in this paper, we take the initiative to study the behavior characteristics of early reviewers through their posted reviews on representative e-commerce platforms, e.g., Amazon and Yelp. We aim to conduct effective analysis and make accurate prediction on early reviewers. This problem is strongly related to the adoption of innovations. In a generalized view, review posting process can be considered as an adoption of innovations³, which is a theory that seeks to explain how, why, and at what rate new ideas and technology spread [8]. The analysis and detection of early adopters in the diffusion of innovations have attracted much attention from the research community. Three fundamental elements of a diffusion process have been studied: attributes of an innovation, communication channels, and social network structures [8].

II. LITERATURE SURVEY

Ting Bai, Jian-Yun Nie[1] provided a an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our analysis of product reviews also indicates that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity. In viewing review posting procedure as a multiplayer competition game, we propose a novel margin based embedding model for early reviewer forecast. Experimenting on two different e-commerce datasets have shown that our proposed system outperforms a number of competitive baselines.

Julian McAuley, Alex Yang[2] Provided a Online audits are regularly our first port of call while considering items and buys on the web. While assessing a potential buy, we may have a particular inquiry as a main priority. To answer such inquiries we should either swim through colossal volumes of buyer audits planning to discover one that is pertinent, or generally suggest our conversation starter straightforwardly to the network by means of a Q/A framework. In this paper we would like to meld these two ideal models: given a huge volume of beforehand addressed questions about items, we trust to consequently realize whether an

audit of an item significant to a given question. We define this as a machine learning issue utilizing a blend of-specialists compose system—here each audit is a 'specialist' that gets the opportunity to vote on the reaction to a specific question; all the while we take in an importance capacity with the end goal that 'applicable' audits are those that vote accurately. At test time this scholarly importance work enables us to surface audits that are important to new questions on request.

Matthew J. Salganik, Peter Sheridan Dodds, Duncan J. Watts [3] provided Collaborative filtering has proven to be valuable for recommending items in many different domains. Here, we explore the use of collaborative filtering to recommend research papers, using the citation web between papers to create the ratings matrix. We tested the ability of collaborative filtering to recommend citations that would be suitable for additional references to target a research paper. We analyzed six methods for selecting citations, evaluating this through offline demonstration against a database of over 186,000 research papers hold in Research Index. We also performed an online demonstrate with over 120 users to measure user opinion of the effectiveness of the algorithms and of the utility of such

recommendations for common research tasks. We came across large differences in the accuracy of the algorithms in the offline experiment, especially when balanced for coverage. In the online experiment, users felt they received quality recommendations, and were enthusiastic about the idea of receiving recommendations in this domain.

III SYSTEM ANALYSIS

EXISTING SYSTEM

Previous studies have highly emphasized the phenomenon that individuals are strongly influenced by the decisions of others, which can be explained by herd behaviour. The influence of early reviews on subsequent purchase can be understood as a special case of herding effect. Early reviews contain important product evaluations from previous adopters, which are valuable reference resources for subsequent purchase decisions. As shown in existing papers, when consumers use the product evaluations of others to estimate product quality on the Internet, herd behaviour occurs in the online shopping process.

Chen et al. have proposed to use multidimensional representations to capture

both intransitivity and context information for modeling pairwise comparison relations.

PROPOSED SYSTEM

To model the behaviors of early reviewers, we develop a principled way to characterize the adoption process in two real-world large review datasets. More specially, given a product, the reviewers are sorted according to their timestamps for publishing their reviews.

In our work here, we mainly focus on two tasks, the first task is to analyze the overall characteristics of early reviewers compared with the majority and laggard reviewers. We characterize their rating behaviors and the helpfulness scores received from others and the correlation of their reviews with product popularity. The second task is to learn a prediction model which predicts early reviewers given a product.

IV IMPLIMENTATION

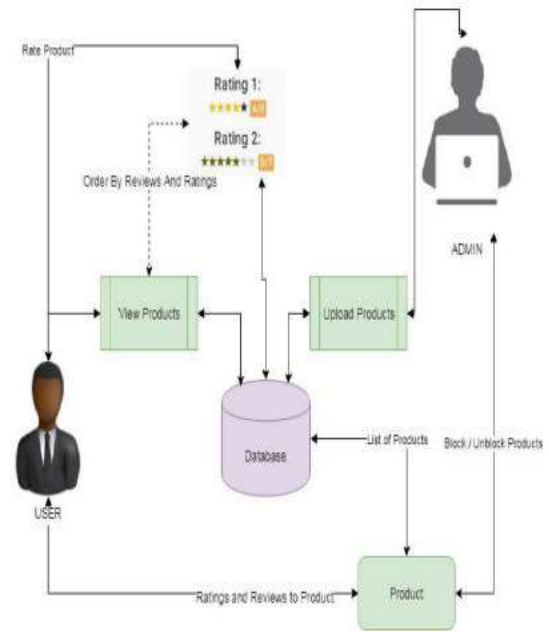


Figure 1: System Architecture

MODULES:

There are three modules can be divided here for this project they are listed as below

- Upload products
- Product Review Based Order
- Rating and Reviews
- Data Analysis

From the above three modules, project is implemented. Bag of discriminative words are achieved

Upload Products

Uploading the products is done by admin. Authorized person is uploading the new arrivals to system that are listed to users. Product can be uploaded with its attributes such as brand, color, and all other details of warranty. The uploaded products are able to block or unblock by users.

Product Review Based Order

The suggestion to user's view of products is listed based on the review by user and rating to particular item. Naïve bayes algorithm is used in this project to develop the whether the sentiment of given review is positive or negative. Based on the output of algorithm suggestion to users is given. The algorithm is applied and lists the products in user side based on the positive and negative.

Ratings And Reviews

Ratings and reviews are main concept of the project in order to find effective product marketing. The main aim of the project is to get the user reviews based on how they purchased or whether they purchased or not. The major find out of the project is when they give the ratings and how effective it is. And this will helpful for the users who are willing to buy the same kind of product.

Data Analysis

The main part of the project is to analysis the ratings and reviews that are given by the user. The products can be analysis based on the numbers which are given by user. The user data analysis of the data can be done by charts format. The graphs may vary like pie chart, bar chart or some other charts.

V RESULT AND DISCUSSION

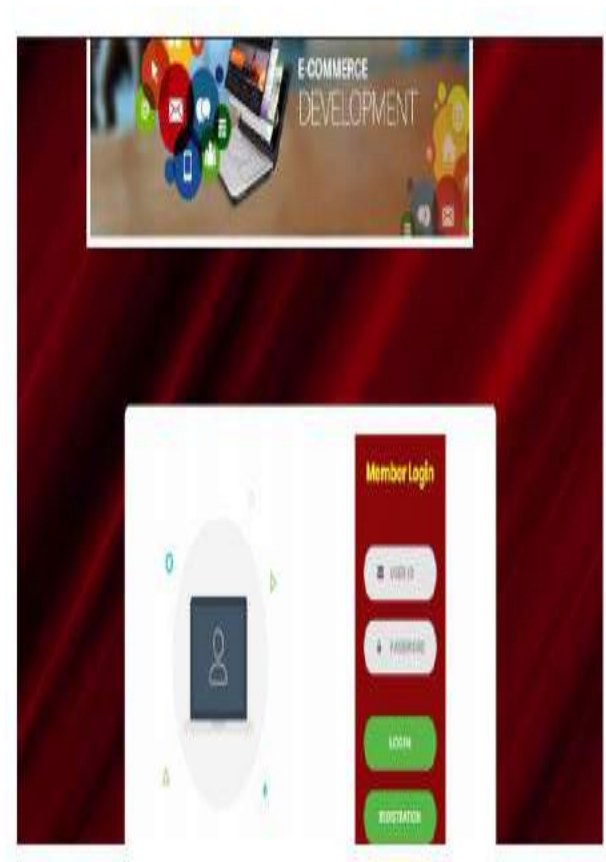


Figure 2: User Login Page



Figure 3: Vender Product Upload Page



Figure 5: User/ Admin Product's Review & Rating Page



Figure 4: User/Admin Products View Page



Figure 6: User Product Rating Page

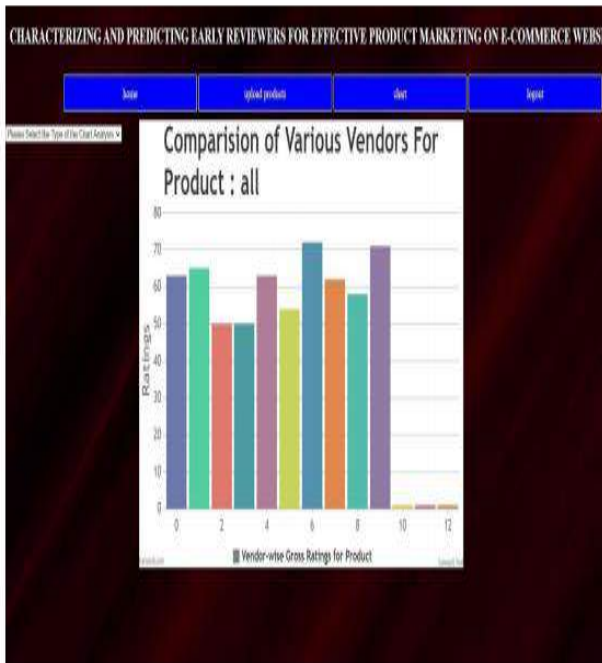


Figure 7: Sentimental Analysis Page

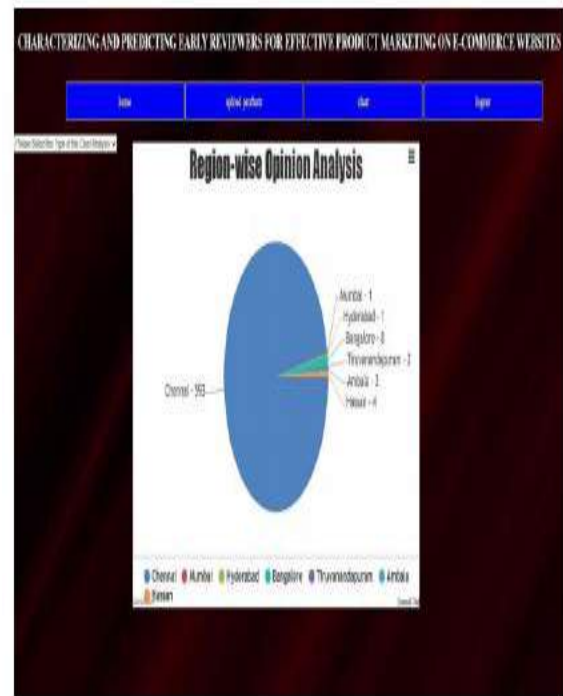


Figure 9: Sentimental Analysis Page

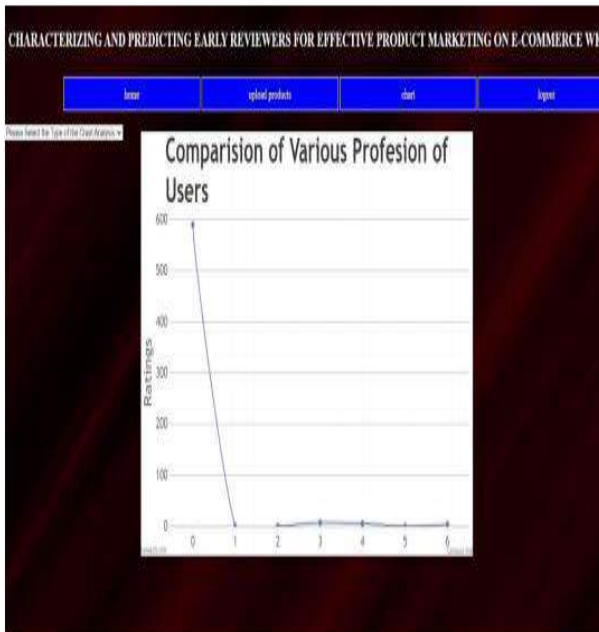


Figure 8: Grapical Analysis Page

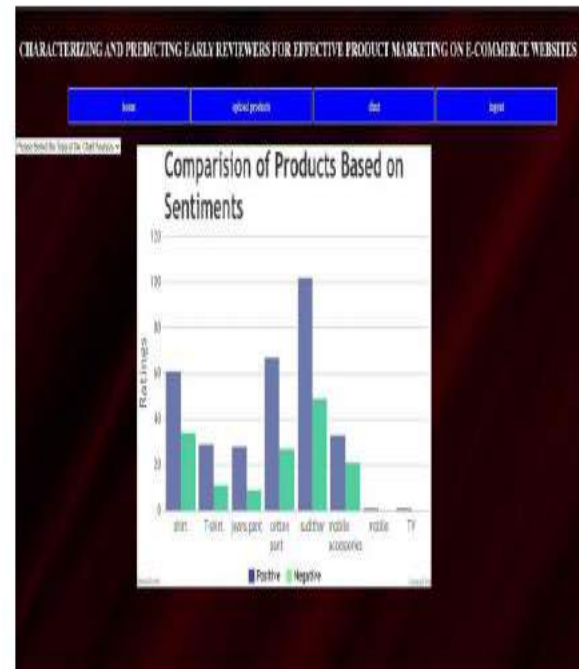


Figure 10: Grapical Analysis Page

VI CONCLUSION AND FUTURE WORK

In this paper, we have studied the novel task of early reviewer characterization and prediction on two real-world online review datasets. Our empirical analysis strengthens a series of theoretical conclusions from sociology and economics. We found that (1) an early reviewer tends to assign a higher average rating score; and (2) an early reviewer tends to post more helpful reviews. Our experiments also indicate that early reviewers' ratings and their received helpfulness scores are likely to influence product popularity at a later stage. We have adopted a competition-based viewpoint to model the review posting process, and developed a marginbased embedding ranking model (MERM) for predicting early reviewers in a cold-start setting. In our current work, the review content is not considered. In the future, we will explore effective ways in incorporating review content into our early reviewer prediction model. Also, we have not studied the communication channel and social network structure in diffusion of innovations partly due to the difficulty in obtaining the relevant information from our review data. We will

look into other sources of data such as Flixster in which social networks can be extracted and carry out more insightful analysis. Currently, we focus on the analysis and prediction of early reviewers, while there remains an important issue to address, i.e., how to improve product marketing with the identified early reviewers. We will investigate this task with real e-commerce cases in collaboration with e-commerce companies in the future.

VII REFERENCES

- [1] J. McAuley and A. Yang, "Addressing complex and subjective product-related queries with customer reviews," in WWW, 2016, pp. 625–635.
- [2] N. V. Nielsen, "E-commerce: Evolution or revolution in the fastmoving consumer goods world," nngroup.com, 2014.
- [3] W. D. J. Salganik M J, Dodds P S, "Experimental study of inequality and unpredictability in an artificial cultural market," in ASONAM, 2016, pp. 529–532.
- [4] R. Peres, E. Muller, and V. Mahajan, "Innovation diffusion and new product growth models: A critical review and research directions," International Journal of Research in Marketing, vol. 27, no. 2, pp. 91 – 106, 2010.

- [5] L. A. Fourt and J. W. Woodlock, "Early prediction of market success for new grocery products." *Journal of Marketing*, vol. 25, no. 2, pp. 31 – 38, 1960.
- [6] B. W. O, "Reference group influence on product and brand purchase decisions," *Journal of Consumer Research*, vol. 9, pp. 183–194, 1982.
- [7] J. J. McAuley, C. Targett, Q. Shi, and A. van den Hengel, "Imagebased recommendations on styles and substitutes," in *SIGIR*, 2015, pp. 43–52.
- [8] E. M. Rogers, *Diffusion of Innovations*. New York: The Rise of HighTechnology Culture, 1983.
- [9] K. Sarkar and H. Sundaram, "How do we find early adopters who will guide a resource constrained network towards a desired distribution of behaviors?" in *CoRR*, 2013, p. 1303.
- [10] D. Imamori and K. Tajima, "Predicting popularity of twitter accounts through the discovery of link-propagating early adopters," in *CoRR*, 2015, p. 1512.
- [11] X. Rong and Q. Mei, "Diffusion of innovations revisited: from social network to innovation network," in *CIKM*, 2013, pp. 499– 508.
- [12] I. Mele, F. Bonchi, and A. Gionis, "The early-adopter graph and its application to web-page recommendation," in *CIKM*, 2012, pp. 1682–1686.
- [13] Y.-F. Chen, "Herd behavior in purchasing books online," *Computers in Human Behavior*, vol. 24(5), pp. 1977–1992, 2008.
- [14] Banerjee, "A simple model of herd behaviour," *Quarterly Journal of Economics*, vol. 107, pp. 797–817, 1992.
- [15] A. S. E, "Studies of independence and conformity: I. a minority of one against a unanimous majority," *Psychological monographs: General and applied*, vol. 70(9), p. 1, 1956.
- [16] T. Mikolov, K. Chen, G. S. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," in *ICLR*, 2013.
- [17] A. Bordes, N. Usunier, A. Garc'ia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multirelational data," in *NIPS*, 2013, pp. 2787–2795.
- [18] A. S. E, "Studies of independence and conformity: I. a minority of one against a unanimous majority," *Psychological monographs: General and applied*, vol. 70(9), p. 1, 1956.
- [19] M. L. S. D. X. W. L. S. Mingliang Chen, Qingguo Ma, "The neural and psychological basis of herding in purchasing books online:

an event-related potential study,”
Cyberpsychology, Behavior, and Social
Networking, vol. 13(3), pp. 321–328, 2010.

[20] V. G. D. W. Shih-Lun Tseng, Shuya Lu,
“The effect of herding behavior on online
review voting participation,” in AMCIS,
2011.

The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - A Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : <http://ijaema.com> / e-mail: submitijaema@gmail.com



Certificate of Publication

This is to certify that the paper entitled

“CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE”

Authored by :

B.VINEETH

From

CMR TECHNICAL CAMPUS UGC AUTONOMOUS Kandlakoya (V), Medchal Road, Hyderabad

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE VI, JUNE- 2021



T.A. Olszewski

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL



<http://ijaema.com/>

The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - A Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : <http://ijaema.com> / e-mail: submitijaema@gmail.com



Certificate of Publication

This is to certify that the paper entitled

“CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE”

Authored by :

K. DURGA PRADEEP

From

CMR TECHNICAL CAMPUS UGC AUTONOMOUS Kandlakoya (V), Medchal Road, Hyderabad

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE VI, JUNE- 2021



T.A.O.

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL



<http://ijaema.com/>

The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - A Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : <http://ijaema.com> / e-mail: submitijaema@gmail.com



Certificate of Publication

This is to certify that the paper entitled

“CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE”

Authored by :

N B VISHWA TEJA

From

CMR TECHNICAL CAMPUS UGC AUTONOMOUS Kandlakoya (V), Medchal Road, Hyderabad

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE VI, JUNE- 2021



T.A. Olszewski

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL



<http://ijaema.com/>

The International Journal of Analytical and Experimental Modal analysis

An UGC-CARE Approved Group - A Journal

An ISO : 7021 - 2008 Certified Journal

ISSN NO: 0886-9367 / web : <http://ijaema.com> / e-mail: submitijaema@gmail.com



Certificate of Publication

This is to certify that the paper entitled

“CHARACTERIZING AND PREDICTING EARLY REVIEWERS FOR EFFECTIVE PRODUCT MARKETING ON ECOMMERCE WEBSITE”

Authored by :

Mr. KOTHA MAHESH (Assistant Professor)

From

CMR TECHNICAL CAMPUS UGC AUTONOMOUS Kandlakoya (V), Medchal Road, Hyderabad

Has been published in

IJAEMA JOURNAL, VOLUME XIII, ISSUE VI, JUNE- 2021



T.A. Olszewski

Michal A. Olszewski Editor-In-Chief
IJAEMA JOURNAL



<http://ijaema.com/>