

UNIVERSITY PARTNER



UNIVERSITY OF  
WOLVERHAMPTON



HERALD  
COLLEGE  
KATHMANDU

## Project and Professionalism (6CS007)

Academic Year	Module	Assessment Number	Assessment Type
S21	Project and Professionalism	A1	Individual Report

Student Name : Upshot Awal  
Student Id : 2051891  
Course : BSc (Hons) Computer Science  
Supervisor : Erin Shakya  
Reader : Sachin Kafle  
Submitted On : 2021/11/16

## Contents

<b>Statement of Project Details</b> .....	3
<b>Academic Questions:</b> .....	3
<b>AIMS</b> .....	3
<b>OBJECTIVE</b> .....	3
<b>Artefact</b> .....	3
<b>Introduction</b> .....	4
<b>Initial research into source of information</b> .....	4
<b>Overview of recommender pipeline</b> .....	4
<b>Artefact Elaboration</b> .....	6
<b>Tools and Technologies</b> .....	7
<b>Methodology</b> .....	8
<b>Testing</b> .....	8
<b>Plan and Schedule</b> .....	9
Work Breakdown Structure .....	9
Gantt chart.....	10
<b>Conclusion</b> .....	10
<b>Additional Information</b> .....	10
<b>References</b> .....	11
 Figure 1 Artefact of Recommendation System .....	6
Figure 2 Gantt chart for schedule .....	10

## Student Details

**Full Name (Surname, Forename):** Upshot Awal

**University Student Id:** 2051891

**College Email Address:** [np03a190209@heraldcollege.edu.np](mailto:np03a190209@heraldcollege.edu.np)

**Supervisor:** Erin shakya

**Date of Submission:** 2021/11/16

## Statement of Project Details

**Project Title:** Product Recommendation System

### Academic Questions:

- What tools and technologies will be used building this web application?
- Which AI is most preferable for running a smooth recommender system?
- How does AI help to get better recommendation?

### AIMS

- To help small store to grow business online
- Help user navigate through products more easily
- To train the dataset according to the user preference

### OBJECTIVE

- It will take user information and give them similar recommendation
- It will predict the possibility that a user would favor an item.
- It helps reduce the data cluster in the vast web and make searching easy.

### Artefact

Recommendation systems are a sort of information filtering that forecasts the rating and preference of an item or social element. The AI recommends a product to the user, sparing them from having to navigate through the entire website. The web application will help filter out the products for buyer and recommend the product that is most likely to be bought by the user. We can divide it into three stages as well. First is information collecting phase where we collect all the information and next is learning phase where we process the data and build the model to train the data and finally the prediction/ recommendation phase where we predict the result and give prediction.

## Introduction

My final year project is based to help small retailer who want their business to grow out from a small local area to other parts of the world. By taking the business online we can do just that and more. With the help of my product recommendation system they can easily get a better head start in the leading business industry. Nowadays people tend to have less time so they want everything done in short amount of time. My AI recommends the product for the user so they don't have to scroll through the whole website. Recommendation system is a subclass for information filtering that seeks to predict rating and preference of an item or social element by using a model built around the feature of item or the buyer's social environment. (Singh, 2021)

Recommendation algorithms have changed the way websites and users connect in recent years. The recommendation engine sifts through massive amounts of data to locate users' areas of interest and makes information retrieval easier. The CHARM algorithm is a prominent pattern-finding tool that can handle enormous datasets, unlike previous association mining algorithms that could not. It is coupled with pattern discovery algorithms such as association rule mining and clustering. (Thomas, 2015)

## Initial research into source of information

The internet contains a vast collection of unstructured data which makes it very hard to search a valuable information. Recommendation system can be further classified into two groups. Firstly, content-based filtering focuses solely on the features of the objects and offer suggestions based on their similarity. Here we have data of what the user previously like and watched/ listened. The system comprises a big database called Item Profile, which contains the items to be recommended as well as their attributes. The contents of the item, as well as the ratings assigned to it, are used as training data in content-based filtering, which is a regression modeling issue or user-specific classification. Based on the user's prior selection behavior, the training data supplied for each user corresponds to the contents of the item. (Thomas Hofmann, 2004)

Collaborative filtering helps filtering for information or patterns using techniques that require collaboration across diverse players, perspectives, and data sources. Here we don't take past data or preference of group or an individual user. In collaborative filtering there may include problems that forces us of predicting unrated items, for such similarities between items and users are calculated using different approach. Basically collaborating filtering method is to make similarity assumption between users or between products, according to their past selection behavior of item or past ratings. (Thomas Hofmann, 2004)

(F.O.IsinkayeaY.O.FolajimibB.A.Ojokohc, 20th August, 2015)

## Overview of recommender pipeline

- Pre-processing

In pre-processing first we normalize the data and the pick a model (matrix factorization) suitable for the data type then pick evaluation metric. (Cates, 2019)

- Hyper parameter tuning

We can use Grid search or random search for hyper parameter tuning for example. (sklerarn.model\_selection.GridSearchCV) (Cates, 2019)

Sequential Model-Based Optimization is an alternative smarter way for hyper tuning parameter.

- Model training and prediction

Using the value given by the optimal hyper parameter we can train our model to get predicted ratings so that we get optimal output or prediction in this case.

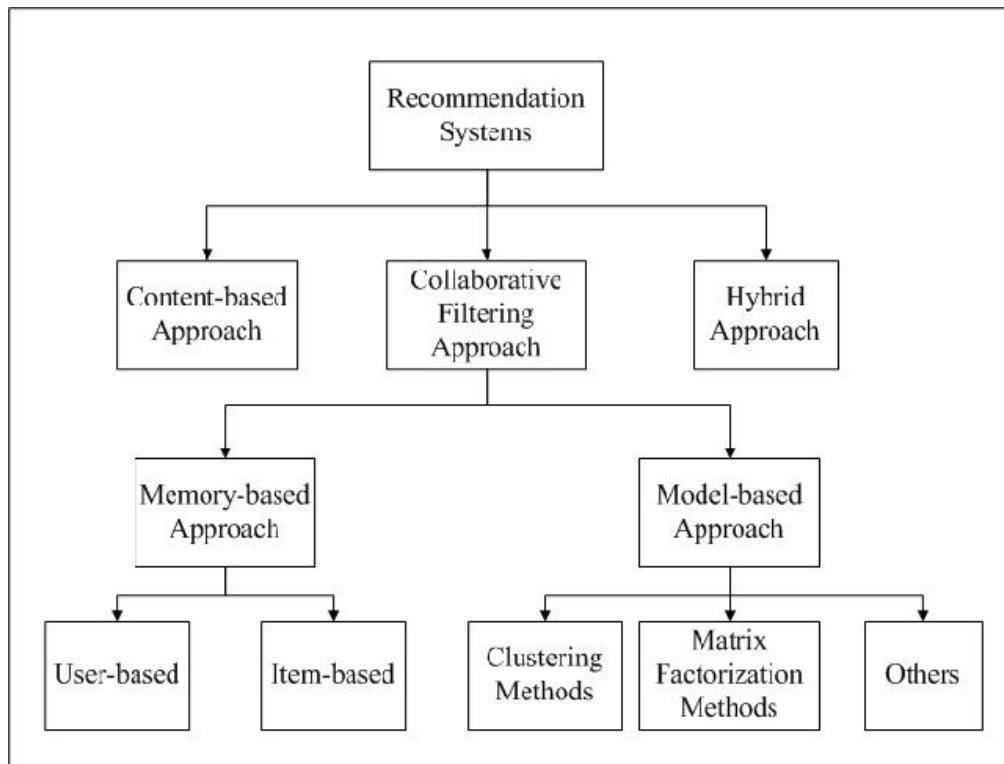
- Post- processing

In post-processing we will be filtering out the products or items that the user have purchased or watched/interacted.

- Evaluation

In traditional ML we split the data into half training set and validation set but doesn't work in recommendation system because the model won't work if we train all our data on a separate user population than the validation set. So we compare the predicted value with the actual value we have to evaluate the data. (Cates, 2019)

## Artefact Elaboration



*Figure 1 Artefact of Recommendation System*

Artificial intelligence does help in creating a better recommendation system. Because of the increased need to get a better understanding of users' preferences, recommender systems have advanced beyond simple user-item filtering to include many components for analyzing and integrating large datasets. For example we have OpenRec which is a Python framework for adaptable and extensible research in recommender systems. Each recommender is represented as a computational graph composed of an ordered ensemble of reusable modules linked by a set of well-defined interfaces and provides adaptability and flexibility while maintaining training efficiency and recommendation accuracy. (Deborah Estrin, 2018). As we know there are many approaches to building a good recommendation system. Machine learning is used to develop the core of AI recommendation system but in present day even deep learning approach are made to recommendation system as the market and competition is growing, companies demand a further advance way to read their clients behavior and patterns. Deep learning's ability to capture non-linear and non-trivial interactions between consumers and things, as well as its ability to integrate large amounts of data, makes it limitless in terms of generating high-quality recommendations. DNNs and CNNs are some of the deep learning algorithms used for recommendation system. (Rozhavsky) There are many approach to giving recommendation accurately. Content based, collaborative or hybrid, we have to process the dataset no matter the approach and by using Hadoop we can do just that.

Hadoop is an open source cloud computing platform using distributed file system, it implements the MapReduce framework. The MapReduce framework enables users to divide a huge problem into many smaller problems, which are then handled by the Hadoop platform, enhancing processing performance. (Ming-sheng Shang, Jan, 2010)

## Tools and Technologies

Some of the tools that will help me build this web application are listed below:

1. PyCharm

It is an IDE for python programming which is free and used by professional developer.

2. Brave browser

It is a web browser where we can research and download useful datasets.

3. Python and Django

It is programming language we will be using for backend. We will be using Django framework. Python is an excellent programming language for beginners as well as advanced programmers who have worked with other programming languages such as C++ and Java. Both the front-end and back-end works of the web application can be done in Django which is designed to build application as quickly as possible. Django is a high-level Python web platform for building stable and maintainable websites quickly. (geeks, n.d.)

4. React.Js

It is a framework used in frontend development of this web application.

5. HTML/CSS

HTML is a markup language that is used to create electronic documents (also known as pages) that can be viewed through the Internet. Each page has a number of hyperlinks that connect it to other pages. Every web page you see on the Internet was built using HTML code in some fashion.

CSS, or Cascading Style Sheets, is a simple design language created to make the task of rendering web pages presentable easier. The look and sound of a web page is handled by CSS.

6. Git-Hub

It is a web hosting business that focuses on software development and GIT version control. It incorporates GIT distributed version control and source code management capabilities.

7. Open CV

We will be using jupyter notebook for handling and training datasets from open sources like kaggle and other websites. We can also use some of the tools like:

## 8. Python tools

Import surprise (@NicolasHug)

Import implicit (@benfred)

Import LightFm (@lyst)

Import pyspark.mllib.recommendation

## Methodology

I've broken this entire system into sprints to make it simpler to complete, and an iterative approach is one of the best sprint strategies for this project because it allows me to focus on past work throughout each sprint.

### Requirement stage

Requirement stage is one of the important stages for a project to be handled out swiftly. In this stage we do research and investigate the necessary information for the project to be done. We make decision on which technique, datasets and software/ hardware to be used.

### Design stage

After that we come to the design state. Here we look at the requirements and build a wire frame or a design prototype so that we have a brief idea on what the project may look like and work according to it and make improvements along the way.

### Execution and testing stage

After we have got the optimal design we need to complete the project so we start execution required and functioning features to the project and test them out. Backend and frontend development and any other aspect is done and tested on this stage.

### Deployment stage

Finally after everything is complete we can now finally deploy the service or project to user, client or the public. Or this all the previous stage s should be carried out swiftly so that we don't have any errors after the deployment so that the client does not have any problem on operating the system

## Testing

For the testing of the web application we can follow scrum methodology by having a test case and a bug report for the following test case. First we put what we are testing and



find if it is the actual result that we want if does not fill the required output then the test is failed and move to bug.

For example:

Task Type	Description	Test Step	Expected Result	Status
Functionality	Area should accommodate 20 characters	Input up to 20 characters	20 characters in the request should be appropriate.	Pass or Fail

### Plan and Schedule

As we are following scrum / agile method to build our web application. We have to separate our projects into different sprints and each sprint will have its own backlog and will be following the work breakdown structure. As for the scheduling we will be using Gantt chart for keeping track of the project and the deadline for each part of project.

### Work Breakdown Structure

Level 1	Level 2	Level 3
Recommendation System	1 Initiation	1 Project Registration Form (PRF) 2 Communication Plan 3 Initial Project Plan 4 Getting a PRF Approved 5 PRF registration
	2 Planning	1 Submit Project Proposal 2 Develop Scope Baseline 3 Understand the Requirements 4 Project Risks 5 Project Plan Approval 6 Discussion of Initial Design/Prototype
	3 Execution	1 Project Assessment 1 2 Project Review Meeting 3 Design Solution 4 Project Presentation 5 Build the System 6 Monitoring and Control 7 Test the System
	4 Monitoring and Control	1 Project Review Meeting 2 Respond to Changes 3 Artefact Progress Check (Testing) 4 Respond to Changes

	5 Closure	1 Draft submission of report 2 Document Feedback 3 Final Project Report and management 4 Archive Documents 5 Demonstrate Artefact
--	-----------	---

## Gantt chart

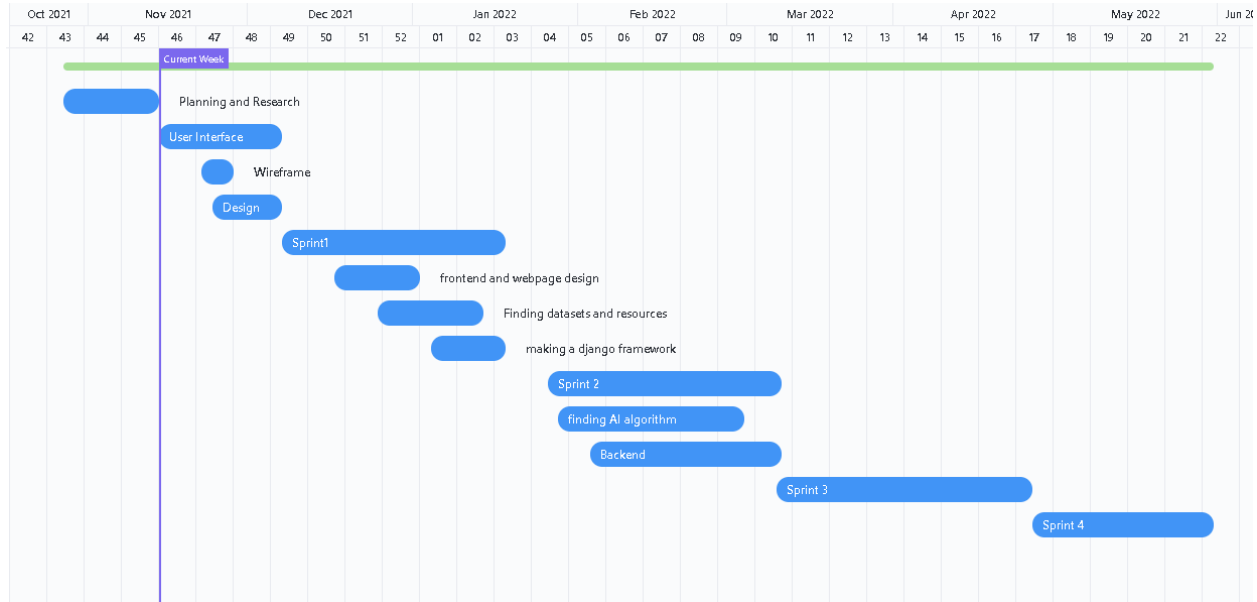


Figure 2 Gantt chart for schedule

## Conclusion

In conclusion, this is a web application for ecommerce which uses artificial intelligence such as recommendation system to interact with the user and help them enhance the user interaction with the website all sell their products more efficiently. A recommender system's primary principle is to evaluate a user's previous interactions with products, because previous item selections may impact the user's future choices, allowing comparable items to be recommended to the user. . (Kwan, 2012)

## Additional Information

Resources: I will be using a normal computer and laptop to research and code. Internet viability is must.

Client: I would like my supervisor to guide and check the web application as I keep developing it. So my supervisor Erin Shakya will be the client of this project.

## References

- Cates, J. (2019, March 15). How to Design and Build a Recommendation System Pipeline in Python .
- Deborah Estrin, L. Y. (2018, February 02 ). OpenRec: A Modular Framework for Extensible and Adaptable Recommendation Algorithms. pp. 664–672.
- F.O.IsinkayeaY.O.FolajimibB.A.Ojokohc. (20th August, 2015). Recommendation systems: Principles, methods and evaluation.
- geeks, g. f. (n.d.). Retrieved from <https://www.geeksforgeeks.org/>
- Kwan, S. R. (2012). *Book Recommendation System using Data Mining for the*. Creative Commons.
- Ming-sheng Shang, Z.-D. Z. (Jan, 2010). *User-Based Collaborative-Filtering Recommendation Algorithms on Hadoop*. Institute of Electrical and Electronic Engineers.
- Rozhavsky, V. (n.d.). The rise of deep learning recommender systems. *Product Recommendations Course*.
- Singh, L. (2021, November 13). *Research*. Retrieved from [https://d1wqtxts1xzle7.cloudfront.net/38584474/IJETT-V4I5P132\\_1\\_-with-cover-page-v2.pdf?Expires=1636795268&Signature=gC7GAVNaAZRIBwbAVhw-s~7pZ0Uynm98gshyPaOrXdrIrfd3JL9f5xywgc dy5iylAIBn0OZc6oE~kehgu86PVTbQHnzW4TkDZXV2j7b3bg8o5wAOHLSuktHLvhVx1qWkRRVjmr5iSP](https://d1wqtxts1xzle7.cloudfront.net/38584474/IJETT-V4I5P132_1_-with-cover-page-v2.pdf?Expires=1636795268&Signature=gC7GAVNaAZRIBwbAVhw-s~7pZ0Uynm98gshyPaOrXdrIrfd3JL9f5xywgc dy5iylAIBn0OZc6oE~kehgu86PVTbQHnzW4TkDZXV2j7b3bg8o5wAOHLSuktHLvhVx1qWkRRVjmr5iSP)
- Thomas Hofmann, J. B. (2004, July 04). Unifying collaborative and content-based filtering.
- Thomas, P. N. (2015). *Survey on recommendation system methods*. IEEE.