

Game recommendation system.

Project Proposal

Mihail Kenarov

1. Introduction

What is the goal of our recommendation system?

The goal of this project is to create a recommendation system that allows a user to find similar video games to the one they have chosen. This will be determined by some of the features of the choice they have made and show other games that they may find interesting to get as well.

Why would we create such a system?

Let us start with this statement:

“Recommender systems are highly useful as they help users discover products and services they might otherwise have not found on their own.” ([What is a Recommendation System? | Data Science – NVIDIA](#))

But let us dive deeper into the “why” are they so useful and “how” is that a useful system that we use and helps us in our everyday lives.

Recommendation systems are everywhere around us and we use them in our everyday lives. They allow us to find more options of the similar things we like. This makes us, as the consumer, find more enjoyable products that can help us in our lives or show us services similar to the one we have chosen, which we can find useful. Most of the time when we buy a product we combine it with another one. For example when we buy a t-shirt. When we take on this action the recommendation system pops up and shows us a pair of jeans for example, that would go along well with it. This way, it helps us not only just buy a random product, but a pair that we would enjoy to wear.

Not only that but when it comes to the businesses that implement these kinds of systems it can bring in more customers, because of the joy from being multiple quality products all together, while also bringing in more revenue. This helps a company grow and make more either products or services for the customer, while also making sure to have enough resources to have increase the quality of their products.

Who would be interested in this project?

People who enjoy playing various kinds of video games may find this project useful, for future recommendations based on their interests. It goes not only for people who spend a lot of their time playing video games, but also for people who are interested in trying some new games in their busy schedules. With the creation of such a system, they can quickly find another game, similar to the one they like – quickly.

When will the project be constructed?

- Phase 1 – Research and planning
 - o Research about the project will be conducted in the first couple of weeks of the semester, to get a better understanding of recommendation systems and what would be the most suitable approach in our case.
- Phase 2 – Development
 - o During that time the data will be edited and structured properly for the selected algorithm, which will also be chosen during this time, based on what currently matches our expectations.
- Phase 3 – Delivery and evaluation
 - o After the creation and testing of the project it will be evaluated to see if it matches up to our expectations.

How will the project be created?

We will first get a suitable dataset that would give us the necessary information to help us create such a system. Afterwards we will see which recommendation algorithm works best in our case and find a suitable model to train. Afterwards inferencing will be done as a way of testing out the project and seeing if the recommendations work as expected.

2.Domain understanding

We all know what video games are, but not many of us know what exactly a recommendation system is and what it consists of, which brought me to create research of my own.

Research questions:

What is a recommendation system?

A recommendation system is a sophisticated tool designed to analyse user behaviour, preferences, and interactions with items to predict what the user might like or need. These systems are integral to many online platforms, including e-commerce websites, streaming services, and social media, where they help users discover new content, products, or services that align with their interests. The goal of a recommendation system is to enhance user experience by providing personalized suggestions, thereby increasing engagement, satisfaction, and potentially sales or conversions.

What about some history of the recommendation systems?

Recommendation systems have been with us since the creation of human time. It started exactly from us – the humans, spreading general ideas while talking to friends, family, or people we just enjoy being with, about things we would say go well together or we would like the people close to us to experience. These were the first ever recommendations that were ever given out and we still use them to this day.

With the evolution of technology, we even received even the first recommendation system, which was made by humans and operated on its own – “Grundy.” It was a system for the recommendation of books based on the users’ inputs. With time it started being criticized as all things in our world, especially in technology.

'The first known such solution was the computer librarian Grundy, which first interviewed users about their preferences and then recommended books to them considering this information. Based on the information collected the system allocated the user into a stereotype group using a rather primitive method, thus recommending the same books to all persons in the same group.

....

However, Grundy's solution quickly gained a lot of critics in the scientific world. Nisbett and Wilson state that "people are very weak in the study and description of their own cognitive processes". According to their studies, people often highlight their attributes which make them stand out from the rest of a particular group, making stereotyping efforts more difficult. Of course, it could happen that we simply want to create a different image of ourselves.' ([History of recommender systems - Onespire Ltd., Sándor Apáthy](#))

Today, we are surrounded by recommendation systems everywhere around us -from the online shops that we visit, to the talks we have with friends. We could say that they have "integrated" into our lives, considering the fact that they with us, no matter what we do – having a discussion with a friend, buying clothing online or just ordering food for home.

How does a recommendation system work?

Recommendation systems operate on the principle of predicting user preferences based on historical data. This data can include user ratings, browsing history, purchase history, and interactions with other users. The system analyses this data to identify patterns and correlations that can be used to predict future preferences. Not only that but it also tracks the similarities between products that a user has been interested in.

What types of recommendation systems are there?

1. **Content-Based Filtering** - This type of recommendation system suggests items to users based on the characteristics or attributes of the items themselves. It analyses the features of items that a user has interacted with in the past and recommends similar items. For example, if a user has

watched action movies in the past, a content-based filtering system might recommend other action movies with similar themes or actors.

2. **Collaborative Filtering** - Collaborative filtering recommends items to users based on the preferences or behaviour of similar users. It identifies users who have similar tastes or interests and suggests items that those users have liked or interacted with, but the current user has not. Collaborative filtering can be further divided into two subtypes:
 - 2.1. **User-based collaborative filtering**: It finds users who are similar to the target user based on their interactions with items and recommends items that those similar users have liked.
 - 2.2. **Item-based collaborative filtering**: It identifies items that are similar to the ones the user has interacted with and recommends those similar items.
3. **Hybrid Recommendation Systems** - Hybrid recommendation systems combine multiple recommendation techniques to provide more accurate and diverse recommendations. By leveraging both content-based and collaborative filtering methods, hybrid systems can overcome the limitations of individual approaches and offer improved recommendation quality.
4. **Matrix Factorization** - Matrix factorization techniques model the relationship between users and items by decomposing the user-item interaction matrix into lower-dimensional matrices. These methods are particularly useful when dealing with sparse data, such as user-item interactions in large-scale recommendation systems.
5. **Context-Aware Recommendation** - Context-aware recommendation systems take into account additional contextual information, such as time, location, device, and user activity, to provide more personalized recommendations. By considering the context in which recommendations are made, these systems can offer more relevant and timely suggestions to users.
6. **Demographic-Based Recommendation** - Demographic-based recommendation systems recommend items to users based on demographic information such as age, gender, location, and occupation. These systems can tailor recommendations to specific user segments or demographics, enhancing the personalization of the recommendation process.

What are some machine learning algorithms used in recommendation systems?

1. Collaborative Filtering Algorithms:

- 1.1. K-Nearest Neighbours (KNN): This algorithm finds users or items that are most similar to the target user or item and recommends items or users that those similar entities have liked or interacted with. ([Collaborative filtering, Wikipedia](#))
- 1.2. Matrix Factorization: Techniques like Singular Value Decomposition (SVD) and Alternating Least Squares (ALS) are used to decompose the user-item interaction matrix into lower-dimensional matrices. These decompositions can then be used to predict missing ratings or to recommend items. ([Matrix Factorization For Recommendation Systems, Melih Kacaman](#))

2. Content-Based Filtering Algorithms

- 2.1. Cosine Similarity: This algorithm measures the cosine of the angle between two vectors to determine how similar they are. In the context of recommendation systems, it can be used to compare the content of items to a user's profile to recommend similar items. ([A Guide to Content-Based Filtering In Recommender Systems](#))
- 2.2. TF-IDF (Term Frequency-Inverse Document Frequency): This technique is used to reflect how important a word is to a document in a collection or corpus. It can be used to represent items in a way that highlights their most notable features, which can then be used to recommend similar items.

3. Hybrid Algorithms

Hybrid Recommender Systems: These systems combine collaborative filtering and content-based filtering to leverage the strengths of both approaches. They can provide more accurate recommendations by considering both the similarity between users and the similarity between items.

4. Deep Learning Algorithms

- 4.1. Neural Networks: Deep learning-based recommendation systems, such as those using Convolutional Neural Networks (CNNs) for image-based recommendations or Recurrent Neural Networks (RNNs) for sequence-based recommendations, have shown great promise in capturing complex patterns in data.
- 4.2. Autoencoders: These are used for dimensionality reduction and can be particularly effective in recommendation systems where the data is high-dimensional. They can learn to represent items in a lower-dimensional space, making it easier to identify similarities.

5. Ensemble Methods

Ensemble Recommendation Systems: These systems combine predictions from multiple recommendation algorithms to improve the overall recommendation quality. Techniques like bagging and boosting can be used to combine the strengths of different algorithms.

6. Graph-Based Algorithms

Graph Neural Networks (GNNs): These are used in recommendation systems to model the relationships between users and items as a graph. GNNs can capture the complex relationships between entities, making them suitable for recommendation tasks.

[What about pros and cons, limitations, and possible ethical problems?](#)

Advantages:

1. Revenue Boost - One of the critical benefits of recommendation systems is their potential to drive revenue growth using data filtering tools. You can increase cross-selling and upselling opportunities by presenting personalized and relevant recommendations. That is why customers are more likely to discover and purchase additional products or services that align with their interests, leading to increased sales and revenue generation. ([Recommendation Systems: Benefits, Types & Examples](#))
2. Enhanced Customer Satisfaction - Recommender systems can elevate customer satisfaction levels, leading to increased customer retention. Recommendations that resonate with a particular user's preferences may change their perception of the platform as attentive and

responsive to their needs. This, in turn, may result in heightened satisfaction and a positive overall user experience. ([Recommendation Systems: Benefits, Types & Examples](#))

3. Personalisation at Scale - Recommendation systems excel at delivering personalised experiences. That is because the recommendation engine processes data to create individual profiles and offer tailored recommendations to each customer. This level of personalisation increases the likelihood of conversions because customers may appreciate the platform's ability to curate products specifically for them. ([Recommendation Systems: Benefits, Types & Examples](#))
4. Controlled Retailing - Controlled retailing enables you to showcase new or underexposed items, manage inventory, and drive sales in trending directions. You can influence purchasing behaviour and promote specific offerings by strategically guiding customers towards particular products or services. ([Recommendation Systems: Benefits, Types & Examples](#))

Disadvantages:

1. Lack of Data - The biggest issue facing recommender systems is that they need a lot of data to effectively make recommendations. It is no coincidence that the companies most identified with having excellent recommendations are those with a lot of consumer user data: Google, Amazon, Netflix, Last.fm. The more item and user data a recommender system must work with, the stronger the chances of getting good recommendations. But it can be a chicken and egg problem – to get good recommendations, you need a lot of users, so you can get a lot of data for the recommendations. ([5 Problems of Recommender Systems - ReadWrite](#))
2. Changing Data - Recommendation systems face challenges in adapting to changing data, particularly in industries like fashion where trends evolve rapidly. As highlighted by Paul Edmunds and David Reinke, relying solely on past behaviour or algorithmic approaches may not suffice due to the dynamic nature of preferences and attributes. Social recommenders offer a promising avenue for addressing these complexities, emphasizing the need for adaptive and context-aware recommendation strategies. ([5 Problems of Recommender Systems - ReadWrite](#))
3. Changing User Preferences - Paul Edmunds highlights the challenge of changing user preferences in recommendation systems, where browsing intentions can vary from day to day. For instance, a user might search for personal items one day and look for gifts the next. These issues underscore the importance of dynamic recommendation algorithms that adapt to evolving user preferences. ([5 Problems of Recommender Systems - ReadWrite](#))
4. Information variability - Recommendation engines rely on historical or current data, which can become outdated quickly, especially in rapidly changing industries like media, online gaming, and marketing. This reliance on "old data" can result in less relevant or even

irrelevant suggestions. One solution is to continuously retrain the neural network with fresh data after each period, ensuring recommendations stay up-to-date and accurate.

([Recommendation Systems: Benefits And Development Process Issues](#))

5. Unpredictable Items – “In our post on the Netflix Prize, about the \$1 Million prize offered by Netflix for a third party to deliver a collaborative filtering algorithm that will improve Netflix’s own recommendations algorithm by 10%, we noted that there was an issue with eccentric movies. The type of movie that people either love or hate, such as Napoleon Dynamite. These type of items are difficult to make recommendations on, because the user reaction to them tends to be diverse and unpredictable.” ([5 Problems of Recommender Systems - ReadWrite](#))

Ethical Problems:

- After reviewing some of the problems previously mentioned, we can see that data plays a crucial role in the creation of a recommendation systems and how it functions. While gathering all of the data it is possible that because of the “slow start”, we are given products or services which are not to our liking and not only that, but it is highly possible that because of data that is not clean, a bias develops.
- While talking about data, it is also important to mention some of the privacy concerns:

“The more the algorithm knows about the customer, the more accurate its recommendations will be. However, many customers are hesitant to hand over personal information, especially given several high-profile cases of customer data leaks in recent years. However, without this customer data, the recommendation engine cannot function effectively. Therefore, building trust between the business and customers is key.

Many businesses are thriving thanks to recommendation engines. While they do bring enormous opportunities, it is vital to be aware of the many challenges inherent to the technology in order to utilize it to the fullest. We would not recommend anything less.” ([7 Critical Challenges of Recommendation Engines](#))

- Sometimes another problem is that the human factor of creativity and change of interest can occur, but the system could not recognise such changes or it finds it difficult to adapt to them. This will make the algorithm irrelevant or it might the possibility of a user to “broaden their horizon”. ([7 Critical Challenges of Recommendation Engines](#))

Who is most affected by the usage and creation of recommendations systems?

Users

Users are the primary beneficiaries of recommendation systems. They receive personalized recommendations that help them discover new products, content, or services that match their preferences and interests. By receiving relevant suggestions, users can save time, make informed decisions, and enhance their overall experience on platforms such as e-commerce websites, streaming services, social media platforms, and more.

Businesses

Businesses that implement recommendation systems benefit from increased user engagement, retention, and conversion rates. By providing personalized recommendations, businesses can improve customer satisfaction, loyalty, and revenue generation. Recommendation systems also enable businesses to gather valuable insights into user behaviour and preferences, which can inform marketing strategies, product development, and inventory management.

Advertisers

Advertisers benefit from recommendation systems through targeted advertising opportunities. By analysing user data and preferences, recommendation systems can deliver personalized ads to users who are more likely to be interested in the advertised products or services. This targeted approach can improve ad relevance, click-through rates, and return on investment for advertisers.

At first it does seem kind of random from where we might have gotten the information from, but this is because we needed to ask a more... “core” question to begin with. And this is the one:

Where are recommendation systems used?

There are many different platforms and places where the recommendation systems are used for a variety of purposes, having a large affect the market. They are used in the online sphere to help customers find things that are of similar interest to the things they like.

After asking an AI chatbot called “[Phind](#)”, this is what we get:

E-commerce and Retail

- Online Shopping Platforms: Amazon, eBay, and Alibaba use recommendation systems to suggest products based on a user's browsing history, past purchases, and the purchases of similar users.
- Clothing and Fashion Retailers: Stores like ASOS, Zara, and Nordstrom use recommendation systems to personalize clothing and accessory suggestions for their customers.
- Electronics Retailers: Companies like Best Buy and Newegg recommend products based on customer reviews, product specifications, and customer purchase history.

Streaming Services

- Music Streaming Platforms: Services like Spotify, Apple Music, and Pandora use recommendation systems to suggest songs, playlists, and artists based on a user's listening history and preferences.
- Video Streaming Platforms: Netflix, Hulu, and Disney+ recommend movies, TV shows, and series based on a user's viewing history, ratings, and the viewing habits of similar users.

Social Media

- Facebook and Instagram: These platforms use recommendation systems to suggest friends, groups, and content that a user might find interesting based on their interactions and the interactions of similar users.
- LinkedIn: Recommends job postings, companies, and networking opportunities based on a user's professional background, skills, and interests.

Online Advertising

- Google Ads: Google uses recommendation systems to suggest ads to users based on their search history, website visits, and the content of the websites they visit.
- Amazon Advertising: Amazon uses recommendation systems to suggest products to advertisers based on the products that similar users have viewed or purchased.

Travel and Hospitality

- Airbnb: Recommends accommodations based on a user's search history, past bookings, and the preferences of similar users.
- Booking.com: Suggests hotels, flights, and car rentals based on a user's search history, past bookings, and the preferences of similar users.

News Aggregators

- Google News: Recommends news articles based on a user's reading history and the reading habits of similar users.
- Reddit: Recommends subreddits and content based on a user's browsing history and the browsing habits of similar users.

Education and Learning Platforms

- Coursera, Udemy, and Khan Academy: Recommend courses, tutorials, and learning paths based on a user's past enrolments, ratings, and the enrolments of similar users.

Healthcare and Wellness

- Fitbit and MyFitnessPal: Recommend workout plans, diet plans, and health tips based on a user's activity data, dietary habits, and the habits of similar users.

Although asking a Chatbot is quite possible to give us a bias answer, this does give us many examples of where and how the recommendation systems are used. However, since it is possible that the information may be subjective it would be a good idea to see some other sources and see if any of these statements are supported elsewhere.

We do see however see other sources that do align with what we were given:

- **Support for “*The statement of ‘Phind’*” A**

“Recommendation systems have a wide range of use cases across several industries, including:

eCommerce: Recommendation systems are widely used in eCommerce to provide personalized product recommendations to customers based on their past behaviours and preferences. Learn more about recommendation systems in eCommerce (coming soon)

Entertainment: Recommendation systems are used in entertainment, such as music and video streaming services, to recommend content that is likely to be of interest to users.

News and media: Recommendation systems are used in news and media platforms to recommend articles, videos, and other content that is relevant to a user's interests.

Social media: Recommendation systems are used in social media to recommend friends, groups, or posts that are likely to be of interest to users.

Healthcare: Recommendation systems are used in healthcare to provide personalized recommendations for treatments, medications, and other healthcare services.

Finance: Recommendation systems are used in finance to provide personalized recommendations for investments, credit products, and other financial services.

Advertising: Recommendation systems are used in advertising to provide personalized recommendations for ads and offers that are likely to be of interest to users.

Education: Recommendation systems are used in education to provide personalized recommendations for courses, programs, and other educational content.”

[\(What are recommender systems? Use cases, types, and techniques, Maor Michael\)](#)

- **Support for “The statement of ‘Phind’” B**

“What are the use Cases and applications of recommendation system

E-Commerce

A recommendation system is very helpful for E-Commerce platforms, it helps provide relevant suggestions to users based on their previous purchases. Recommendation systems help provide personalized offers, product recommendations and recommendations for users with similar tastes.

A good recommender system can give a 22.66% lift in conversion rates, on an average.

Entertainment

Recommendation Models can analyse and understand consumer behaviour to detect patterns that can help provide content suggestions to the users. This way a recommendation system is very likely to provide suggestions that will match the user’s needs.

This is what Netflix does, by analysing the user's tastes and preferences it helps come up with more recommendations for the user.

Social Media

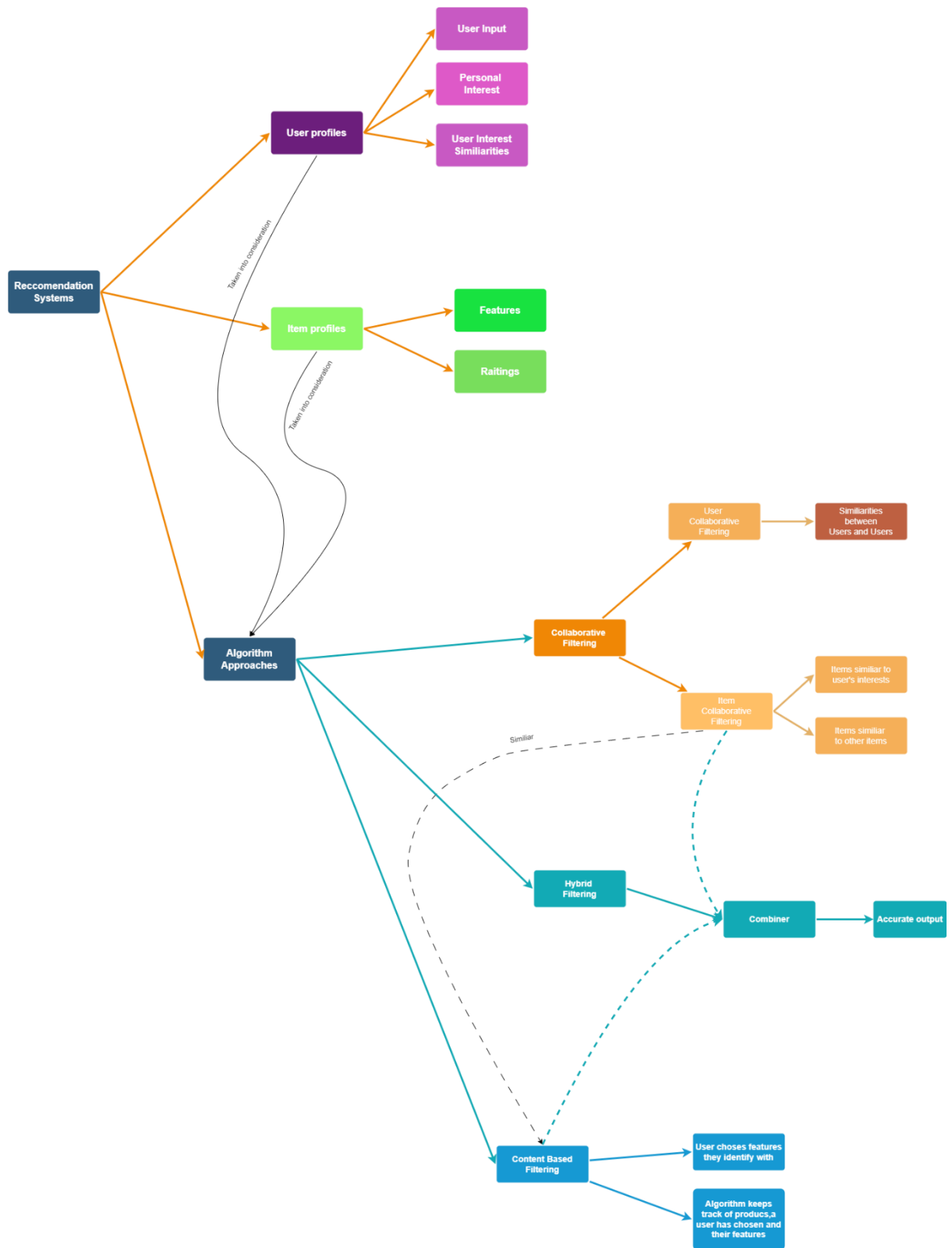
Social media platforms that have been growing rapidly with millions of active users also use recommendation systems. Social media platforms use recommendation systems to understand the user's interests, and analyse their data, to suggest other users with similar interests.

Travel & hospitality

Recommendation systems are also used to enhance the travel and hospitality platform. It provides the users with personalized travel options and hotel recommendations. Along with that it also helps provide destination suggestions, travel packages and itineraries”.

([Recommendation systems | Engati](#))

As we can see, from the small examination of the given spheres of interests in the world, we can comfortably say that recommendation systems are big factor in the way that businesses expand and also the users are experiencing used to seeing them everywhere – no matter is if it is a digital or an actual holdable product.



3. Data Sourcing

Objective

Considering that the mission of the project is to create a recommendation system for video games, which will help gamers find more alternatives that they find interesting.

Data Requirements

The dataset that has been selected is full of text data as well as some numerical data that will be useful to our project. Also, all the columns have been labelled which will make it easier for us to understand it.

Data Sources

Publicly available dataset - <https://www.kaggle.com/datasets/asaniczka/video-game-sales-2024>

Actual source of the dataset - <https://www.vgchartz.com/>

The site seems to specialize in showing different kinds of data that would be interesting for gamers. We are shown variety of charts and articles about different topics. It varies from hardware to software as well. It does seem quite trustworthy and shows a lot of information that can be useful to the users. When it comes to the data – it has been collected in 2024, so we can comfortably say that it is recent and quite relevant.

Data Legality and Ethics

It is publicly available, and we have also been given a source from where it has been scraped from.

Data Diversity

- (0) Title: The name of the game - Text
- (1) Console: The console on which the game is played on- Text
- (2) Genre: The genre of a game - Text
- (3) Publisher: The game publishers which could be considered 'The big names' of the industry - Text
- (4) Developer: The studio that worked on the creation of the game - Text
- (5) Critic score: The score that is given to a game from a certain agency (like IGN) - Number
- (6) Total sales: The number of times the game has been sold worldwide - Number in millions
- (7) NA sales: The number of times a game has been sold in North America - Number in millions

(8) Japan sales: The number of times a game has been sold in Japan - Number in millions

Version Control

There are some missing fields like the ratings of a game on a specific console, but it can be fixed with some after some data cleaning. This means that some changes do need to be made to the chosen dataset.

A history of the project and the processing of the data will be kept on a Git repository in case of an incident and as a way for a version control.

The link to the git repository:

- <https://git.fhict.nl/I509460/video-game-reommendation.git>

Iterative Process

The model will be check for its accuracy continuously and depending on the results, more processing of the data is going to be done.

4. Analytic Approach

The target is to give the user the titles of multiple games which can interest them for them to buy. This will be done by finding similar games to the one they have selected and based on the features of the game, find other ones that will be similar to it.

The problem is that the model is a bit of a hybrid between a classification issue (currently the genre) as well as some would say it is a regression issue (because of the rating).

The chosen model for the task currently is Nearest Neighbours (kNN) and we will try to determine the success of it based on the accuracy. We will be using kNN as a starting point, since it is a great algorithm for the creation of a system that is of the classification type, while also giving us the desired effect. This will be a good choice since we are currently going for a content-based filtering system.