DOI: 10.6977/IJoSI.202309_7(7).0006
O. Stitini, S. Kaloun, etc./Int. J. Systematic Innovation, 7(7),89-111(2023)

Towards a robust solution to mitigate all content-based filtering drawbacks within a recommendation system

Oumaima Stitini[1*], Soulaimane Kaloun[2], Omar Bencharef [3]

[1,2,3] Computer and system engineering laboratory, Cadi Ayyad University, Marrakesh, Morocco, oumaima.stitini@ced.uca.ma, so.kaloun@uca.ac.ma, o.bencharef@uca.ma

*Corresponding author Email: oumaima.stitini@ced.uca.ma

(Received 17 January 2023; Final version received 25 June 2023; Accepted 15 June 2023.)

Abstract

Recommendation systems deliver a method to simplify the user's desire. Recommendation systems are now commonly used on the Internet. It helps suggest items in various categories, including e-commerce, medical, education, tourism, and industrial. The electronic commerce sector has taken a big place in our daily lives as an active research tool, which helps people find what they are looking for. This paper presents a new contribution based on the combination of different algorithms to find a suitable solution to all the drawbacks of content-based recommender systems. The main contribution of this research lies in how to solve each problem and move on to the next. This paper describes an Ideal Solution Mitigating Content Disadvantages based on Three Phases called $ISMCD_{3P}$. Experiments show that the algorithm can propose an appropriate solution to solve all the problems of content-based filtering. Experimentations operating on real datasets are used to estimate the efficacy of our strategy.

Keywords: Content-based filtering drawbacks, Over-specialization, Limited content, Serendipity, Sparsity.

1. Introduction

Recommender systems are strategies and application solutions that offer consumers tailored recommendations concerning a selection of objects, such as goods, videos, music, or other resources (Oumaima et al., 2020). Recommender systems are beneficial when there is an information overload or when the user finds it difficult to navigate and make selections from the catalog due to the overwhelming number of options (Kumar & Thakur, 2018). Numerous online stores and multimedia services, such as Spotify, Netflix, YouTube, and Amazon, as well as social networking sites like Facebook and Twitter, have succeeded in increasing user satisfaction and revenue through the personalized assistance that recommender systems provide in the exploration and discovery of content. The last 10 years have seen an increase in interest in recommender systems, which is still an active area of study.

Recommender systems are computer programs that may examine a user's past actions and make recommendations for present problems. Consequently, to process data that might be quite vast in volume effectively. We frequently consult others while making decisions in life, whether it be choosing which shampoo or book to buy, music to listen to, a movie to watch (Saraswat et al., 2020), or an article to read (Stitini et al., 2021) on the Internet, among others. We may consult friends, family members, or more frequently these days online product reviewers. This has made the recommender's path more accessible.

Consumers have embraced e-commerce quickly thanks to the rise in the number of products available for purchase online and the accessibility of information on the features and functionality of these products. Online purchasing environments have become crowded due to the availability of information and options. Due to the overwhelming amount of options available to them and the abundance of information available for each, consumers are less motivated to filter and assess items (Papneja, 2018) as mentioned in Table 1. Numerous internet retailers provide thousands of unique items, including music, movies, books, and services. A store with a large selection of products is likely to have a product that suits customer preferences. Although more options are accessible, this does not always result in satisfied customers since it can occasionally be challenging to choose a product from the vast selection (Saraswat, 2022). Online vendors are using recommender systems more frequently to deal with the problem of information overload. These systems can be categorized into two types: non-personalized and personalized as mentioned in Fig. 1. Non-personalized algorithms offer the same recommendations to all users, based on factors such as item popularity. Conversely, personalized algorithms generate customized suggestions based on the individual user, resulting in different recommendations for different users.

Fig. 1. Main Recommender System Types.

We concentrate on personalized recommendations in this study. Personalized recommender systems often offer recommendations based on user profiles. Any information about a person, such as an ID, age, gender, and historical activities the user has taken with objects, may be included in a user profile. In this essay, "user profile" refers to a collection of things the user has rated. This opened the door for recommender systems, which are now well-liked platforms for e-commerce. Systems that recommend user items based on their interests and preferences have been created. These solutions simplify things for customers to make decisions by managing information overload, cutting the cost of searches, and enabling users to make wiser, more informed decisions. Recommender systems frequently serve as a user's sales assistant by assisting them. At the same time, they browse, putting the finishing touches on the list of items they have selected and, most importantly, providing personalization.

The key contributions of this paper are as follows: Section 2 contains our literature assessment and theoretical basis. We outline our research contribution in Section 3. We describe our suggested recommender system model in Section 4. Then, in Section 5, we will elaborate on experimental results. In Sections 6 and 7, we discuss our proposed approach $ISMCD_{3P}$ with other related works. Section 6 compares our proposed approach with existing ones. At the end of the work, we discuss and conclude all the work in Section 8.

2. Literature Review and Theoretical Background

Content-based filtering systems create a model by analyzing items and user preferences. This model considers a user's specific interests and tries to find content items that match the user's profile (Javed, 2021). In other words, the system examines the properties of the recommended content items and matches them with the user's preferences. This information is used to personalize the recommendations and suggest items that are likely to be of interest to the user. They also suffer from the drawback of needing sufficient data to create a solid classifier. Wrapper approaches break the characteristics down into smaller groups, analyze these groups, and then determine which of these groups appears to be the most promising. Heuristic techniques are employed by filtering algorithms to assess the content of features, which can be used regardless of the methods utilized. On the other hand, feature extraction is performed during the training stage of embedding techniques, which are integrated into the algorithm. As mentioned earlier, content-based filtering scrutinizes users' previous behavior and proposes items similar to their preferences, based on the characteristics considered. This purpose is to suggest movies to users based on related genres. Content-based filtering techniques use several attributes of an item to suggest other things with related qualities (Pérez-Almaguer et al., 2021). In terms of content-based filtering techniques, it aims to suggest to the active user products comparable to those rated favorably in the past (Sunandana et al., 2021). It is predicated on the idea that things with comparable features would receive comparable ratings. Text documents are the primary information source utilized by content-based filtering algorithms. The main emphasis of content-based recommendation is on modeling both person and item profiles using a single extracted feature strongly connected to item attributes. This section is divided into four parts the first 2.1 represents the procedure used during content-based filtering, the second 2.2 presents how content-based filtering is built, the third 2.3 describes the architecture of content-based filtering, and the last 2.4 shows the objectives and challenges of content-based filtering.

2.1 Recommendation Process

Content-based recommendation can be divided into four steps as mentioned in Fig.2:

- Analyze: Content-based recommender systems examine the item descriptions and a collection of papers that have already received user ratings to determine which things are of interest to the user.
- Develop: Based on the characteristics of the things that users have assessed, they construct a model or sketch out users' interests.
- Build: Then, using a machine learning model, recommendations are made based on user profiles.

- Recommend: The user profiles are compared to the content profiles, and the users are suggested material with comparable feature values.

Fig. 2. The main steps for the content-based recommendation process.

2.2 Content-based recommendation techniques representation.

The item is recommended by content-based recommender systems depending on how well the article's contents fit the user's profile. Content-based recommendation systems typically involve the following steps as mentioned in Fig. 3:

1. Item representation: This involves representing each item (e.g., movie, book, product) in the system using features such as textual data (e.g., title, description), metadata (e.g., genre, director), or other characteristics (e.g., price, release date).
2. User profile creation: This involves creating a profile for each user based on their past interactions with the system (e.g., items they have rated, viewed, or purchased) and their explicit feedback (e.g., ratings, likes, dislikes).
3. Content-based filtering: This involves using the item representations and the user profiles to recommend items that are similar to the items the user has already interacted with or expressed interest in. This can be done using various techniques such as cosine similarity, Euclidean distance, or clustering.

Content-based recommender systems use a user's preferences and interests to recommend items that match their profile. The system generates item recommendations based on the similarity between the content of the items and the user's interests. To do this, the system creates a representation of the item's content and the user's preferences. There are several techniques to represent the content of an item, including:

- Bag-of-words (BoW): This technique represents the content of an item as a set of words or terms that occur in the text. BoW does not take into account the order or context in which the words appear.
- TF-IDF: This technique is similar to BoW but assigns a weight to each word based on how often it appears in the document and how rare it is in the collection of documents. This helps to prioritize important words in the representation.
- Word embeddings: This technique represents words as dense vectors in a high-dimensional space, where words that have similar meanings are closer together. This allows the system to capture semantic relationships between words and to represent the content of an item as a combination of the embeddings of the words that appear in the text.

Once the item is represented in a suitable way, the system can compare it to the user's profile and recommend most similar items. This is typically done using a similarity measure such as cosine similarity, which calculates the cosine of the angle between the two vectors representing the item and the user's profile. The higher the cosine similarity score, the more similar the item is to the user's profile and the more likely it is to be recommended.

Fig. 3. Content-based recommendation representation.

2.3 Architecture of content-based filtering

Content-based techniques build user profiles based on the features and descriptions of the products the user evaluates rather than drawing on the preferences of other users when generating suggestions (Kunde et al., 2022). Content-based have several benefits; on the one hand, strategies over collaborative filtering algorithms are their capacity to address the issue of new products or the potential for encouraging new things for which there is no user input (Stitini et al., 2022). On the other hand, Contrary to collaborative filtering approaches, which may be used everywhere, content-based algorithms heavily rely on the recommendation domain. They also rely on the availability of trustworthy information about the characteristics of the items, which might be difficult to get at times. Additionally, content-based strategies could though not usually be prone to over-specialization, which is the tendency for them to propose goods that are excessively similar to ones the customer has already assessed. Algorithms from several domains, including Information Retrieval, Semantic Web, and Machine Learning, are included in proposals for content-based recommendation systems. For instance, early concepts for Web recommendations, news recommendations, and, more recently, social tagging systems incorporated term-weighting models from information retrieval. For content-based recommendations, such as news recommendations or movie and music recommendations utilizing Linked Open Data, methods utilizing Semantic Web technologies have also been proposed. The architecture of a content-based filtering system typically involves the following steps as mentioned in Fig. 4:

1. **Data Collection:** Collect data on items to be recommended, which can include textual descriptions, tags, or metadata.
2. **Content Analyzer:** Use a Content Analyzer component to extract features from the item data. The Content Analyzer analyzes the text and metadata to identify important features of the item, such as keywords, topics, and categories.
3. **Profile Learner:** Create user profiles based on their historical behavior, such as items they have viewed, rated, or purchased. The Profile Learner builds a profile for each user based on their preferences and behavior.
4. **Filter Component:** The Filter Component then takes the user profiles and the features of the items and calculates the similarity between the user profiles and the items. The Filter Component recommends items that are most similar to the user profile, based on the features of the items.
5. **Evaluation:** It is important to evaluate the performance of the recommendation system to ensure its effectiveness. Common evaluation metrics include accuracy, precision, recall, and F1-score.

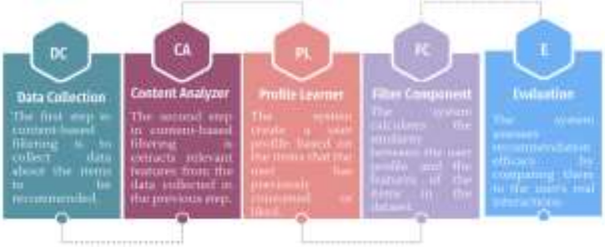

Fig. 4. Global architecture of content-based recommendation system.

2.3.1 Content Analyzer

Examining texts to find pertinent ideas that best describe the content enables the introduction of semantics into the recommendation process. With this method, the appropriate meaning or senses for each ambiguous word are chosen based on the context in which they are used (Stitini et al., 2022). To combat the issues caused by ambiguity in natural language, ideas rather than keywords are used to express texts in this fashion. A repository of disambiguated documents is the end product of the preprocessing procedure. This semantic indexing primarily draws on the linguistic information in the WordNet lexical ontology and is based on natural language processing techniques (Felfernig et al., 2014).

2.3.2 Profile Learner

The system learns the user's preferences and creates a user profile based on the content they have interacted with in the past. To create a user profile, the system typically uses machine learning algorithms to analyze the user's historical behavior such as what items they clicked on, how long they spent on each item, and what items they purchased or rated positively. These behaviors are used to identify patterns and characteristics that can be used to make recommendations in the future. The user profile created in this step typically contains information such as the user's preferred genres, authors, actors, and other attributes that are relevant to the type of content being recommended. This information is then used in the next step of the recommendation process, which is content filtering. Overall, the profile learner is an essential component of a content-based recommendation system as it helps the system to personalize recommendations based on a user's individual preferences and behavior.

2.3.3 Filter Component

The Filter Component is responsible for selecting and filtering the items that will be recommended to the user based on their preferences and interests. This component uses a set of filtering techniques to identify the items that are most relevant to the user. The filtering techniques can include content-based filtering, collaborative filtering, or a hybrid of both. Content-based filtering involves recommending items that are similar to items the user has interacted with in the past. Collaborative filtering involves recommending items that other users with similar interests and preferences have interacted with. The Filter Component takes into account the user's profile, which includes their past interactions, demographic information, and explicit feedback. The user's profile is compared to the item profiles, which include attributes such as genre, artist, director, and rating. The Filter Component uses algorithms to match the user profile with the item profiles to determine the relevance of each item to the user. Once the Filter Component has identified the most relevant items, they are passed onto the next component, which is the Recommender Component. The Recommender Component is responsible for generating a ranked list of items that will be recommended to the user. The ranked list takes into account the relevance of each item and the user's preferences and interests.

2.4 Objectives and Challenges of content-based filtering

Recently, recommender systems have become more prevalent in several commercial applications to assist customers in finding their favorite goods. The accuracy of predictions and the relevance of suggestions have typically been the emphasis of recommender system research. Other suggestion quality criteria, on the other hand, may significantly influence a recommender systems overall performance and user satisfaction (Sharma et al., 2011). As a result, the focus of researchers on this subject has lately expanded to incorporate other recommender system goals. Fig. 5 shows the challenges of content-based filtering.

Fig. 5. Content-based filtering drawbacks.

Recommender systems have recently gained popularity on several online platforms that offer many things, resulting in information overload for consumers. The primary purpose of an RS is to give customers a list of possibly desired things to assist them in shopping online, which benefits both the company and the client. The goal of recommender systems is to make product suggestions relevant to the user's interests. In the recent decade, the objective of RS research has been to give suggestions that are relevant to the user's preferences. However, in real-world circumstances, this is insufficient to captivate customers and persuade them to check out and buy different things. Research on RS has recently switched its focus to integrating relevance with other goals diversity, novelty, coverage, and serendipity.

A recommender system helps users find tailored items, documents, friends, places, and services while saving time. Furthermore, the recommender system addresses the problem of information overload that has plagued the internet in the twenty-first century. Simultaneously, various settings or technologies cloud, mobile, and social networks have grown popular in recent years and are confronted with the challenge of massive amounts of data. As a result, the researchers believe the recommender system is an appropriate answer to this problem in specific settings.

A million digital goods like documents, merchandise, music, and books are uploaded to the internet every day. This enhances the information available on the Internet and gives consumers more options. Users will find looking for and locating the target documents, items, music, and books difficult and time consuming.

When the Recommender system RS is implemented, the quantity of digital products that a user will search for is reduced to the user's most likely favored set of objects. To put it another way, the RS aims to assist online users in finding content that is relevant to their preferences rather than sifting through a mass of undifferentiated data.

2.4.1 Problem 1: over-specialization issue

A crucial method for assessing anything unusual is absent from a content-based recommendation system. The algorithm may offer ideas with a higher score when analogized to the user profile (Jain et al., 2015). It stands likewise called the serendipity problem since it highlights the limitations of content-based recommendations. The amount of fresh content produced by the ”ideal” content-based method would be small, which would limit its potential uses.

Because they only propose goods similar to those that consumers have already assessed, the content-based approach suffers from the over-specialization challenge. One solution to this problem might be incorporating any randomness. Not all content-based algorithms that cannot suggest things that are unrelated to what the consumer has already viewed fall under the category of over-specialization (Mohamed et al., 2019). Items that are highly similar to what the user has previously seen, such as many news stories documenting the same occurrence, should not be made available in some cases.

The inability of the program to suggest to the user objects that are distinct from those he has already observed is only one aspect of the over-specialization problem but also that it must not recommend items that are too close to those he has enjoyed in the past. As a result, certain recommendation algorithms exclude not just things that differ from the user profile, but also those that are highly similar to those already followed by the present user. The diversity of the recommendations is a criterion for evaluating the quality of the recommendations. The user must receive diversified and not homogeneous relevant recommendations. For example, it is not wise to recommend all of Henri Garetta’s books to a user who has enjoyed one of his books.

2.4.2 Problem 2: serendipity issue

Serendipity refers to the potential for obtaining an unanticipated and lucky article (Iaquinta et al., 2008). It is a technique for increasing the variety of suggestions. Due to over-specialization, content-based algorithms lack a crucial method of providing spontaneous ideas, even though people rely on happenstance and experimentation to uncover new items they did not realize they wanted (Shrivastava et al., 2022). Operational serendipity tactics significantly increase the functionality of content-based recommender systems and lessen the problem of over-specialization. Table 2 summarizes contributions regarding the serendipity problem.

Serendipitous recommendations are suggestions that may interest users, even if they aren't directly related to their previous behavior or preferences. By introducing unexpected recommendations, users are exposed to diverse content that they might not have discovered otherwise. This can lead to a more enjoyable user experience and increased engagement. In their work, (De Gemmis et al., 2015) explore how serendipitous recommendations can engage users in unique ways. They discuss various methods for generating such recommendations, including using contextual information or machine learning algorithms that consider diverse factors beyond just user behavior. Recommender systems have become widespread in various domains, but users are increasingly concerned about the privacy and security of their data, as well as the transparency and accountability of the algorithms used. (Rodríguez-García et al., 2019) review existing research on trustworthy recommender systems, including privacy-preserving recommendation algorithms, explainable AI techniques, and the design of transparent and accountable recommender systems.

A crucial mechanism for finding anything unexpected is absent from a content-based recommendation system. The user offers products comparable to previously rated things since products are suggested as reasons for the high score and compatibility with the user’s profile (Ziarani & Ravanmehr, 2021). The over-specialization problem, which describes the propensity of content-based systems to offer unoriginal ideas, is another name for this issue (Kotkov et al., 2020). An excellent content-based strategy is essential to discover some inventive and unbelievable recommendations.

Item-based recommendation systems face the serendipity issue because they only suggest things to users who have already previously loved them. For instance, a movie recommendation engine could only suggest films to a user if the genre or performers are comparable to ones they previously adored. On the other hand, a user-based suggestion can provide surprising advice by examining the users' close friends who have rated the same item as them and by examining their ratings of different products that they have not yet evaluated (Kotkov et al., 2017).

Due to the restricted content analysis, overspecialization results, with CB filtering selecting previously seen items over brand-new ones. By using evolutionary algorithms that offer diversity to suggestions, we may promote unique and coincidental items alongside well-known goods by adding additional hacks and noting unpredictability. There is no necessary method to find something surprising in a recommendation system that uses content.

The user offers products comparable to previously evaluated entities since products are recommended based on a high score and also fit the user’s profile. The tendency of content-based algorithms to provide ideas with minimal originality is sometimes referred to as the over-specialization problem.

2.4.3 Problem 3: limited content issue

Content-based approaches restrict the number and type of common qualities they may offer manually or automatically with objects. Expertise in the domain and taxonomies relevant to the field are also essential. The content-based recommender system could not provide appropriate suggestions if the examined content lacks sufficient information. According to restricted content analysis, the system can only give a tiny quantity of information about its consumers or the range of its products. On the other hand, the way content-based strategies advertise new items results in over-specialization (Adamopoulos & Tuzhilin, 2014). For instance, in a recommendation system for movies, the framework can suggest to a user a film with the same genre or cast as one they have already seen. As a result, the algorithm could overlook some items that the user finds intriguing. A natural limitation of content-based filtering is the need to have a varied and rich representation of item content, which is not always the case. The quantity of data the algorithm needs to distinguish between products the user likes and doesn't like affects how accurate the suggestions are. Table 3 summarizes the contribution regarding the limited content problem.

2.4.4 Problem 4: Scalability issue

For a recommendation machine to understand user preferences, a lot of ratings must be gathered. Because no past data is available, the algorithm cannot give reliable recommendations to new users. With little or limited information, accurate suggestions cannot be generated for new users/items. This is known as the ”cold start” issue. Before the algorithm can understand the user's tastes and offer pertinent recommendations, the user must rate some goods (Su et al., 2022). The user-cold start problem is the name given to this issue in the literature. Lack of consideration of the evolution of the user's interests. Table 4 summarizes the contribution regarding the limited scalability problem.

2.4.5 Problem 5: Synonym issue

Synonyms are two or more words expressing the same thing or concept. However, recommendation algorithms are unable to distinguish between these terms. A memory-based CF method, for example, determines between ”comedy movie” and ”comedy film”. Synonym overuse degrades the quality of a recommender system (Isinkaye et al., 2015). Table 5 summarizes contributions regarding the synonym problem

Table 1. Summarization of contribution regarding the over-specialization problem.

Contribution	Dataset	Proposed approach	Solution proposed	Metric
(Stitini et al., 2022b)	MovieLens	Genetic algorithm	Genetic algorithm & We made an effort to investigate a fresh strategy to address the issue of over-specialization in content-based recommender systems and develop novel things for the user. The genetic algorithmRRS_{GA} was employed in this work to carry out content-based filtering. RRS_{GA} employs a genetic algorithm approach to provide suggestions to the user. The main goal of this system is to find a list of fresh goods with a strong correlation to user preferences and a high likelihood of being selected (the proposed fitness function)	Novelty and diversity
(Adamopoulos & Tuzhilin, 2014)	MovieLens and MovieTweetings	Probabilistic neighborhood selection (PNS) algorithm	The authors argue that collaborative filtering systems can suffer from over-specialization, where users are recommended items that are too similar to the ones they have already consumed. This can limit user exploration and prevent them from discovering new items that they might enjoy. Additionally, CF systems can also exhibit concentration bias, where popular items receive a disproportionate amount of recommendations, while niche items are overlooked. To address these issues, the authors propose a probabilistic neighborhood selection (PNS) algorithm that selects neighbors based on the probability of their ratings being similar to the user's ratings. This helps to increase diversity in recommendations and reduce concentration bias.	Novelty and diversity

Table 2. Summarization of contribution regarding the serendipity problem.

Contribution	Definitions	Metrics	Solution proposed
(De Gemmis et al., 2015)	Serendipity is a representation of anything valuable, difficult to recognize, unexpected, and happening just once.	They define serendipity in the context of content-based recommender systems as: • Relevant: Items that still connect to or resemble the user profile in some way. • An enthusiastic response from users was unexpected. • novel: describing to users as new things.	No solution
(Grange et al., 2019)	Serendipity is the unanticipated occurrence of fortuitous circumstances, such as finding necessary knowledge.	Unexpectedness and informational value.	No solution
(Saat et al., 2018)	The term serendipity is a symbol of value, difficult to identify, unexpected, and only happens at first sight.	Relevant, unexpected, and novel	Linked Open Data.

Table 3. Summarization of contribution regarding the limited content problem.

Contribution	Definitions	Metrics	Solution proposed
(Beleveslis & Tjortjis, 2020)	They provide a feature-weighted heuristic technique for content-based filtering to foster suggestion diversity and streamline similarity computations.	Diversity	The hashing technique in the suggested method accelerates and streamlines the computation of product similarity compared to conventional methods
(Stitini et al., 2023)	A content-based recommendation system recommends items to users based on their preferences and past behavior. One of the limitations of these systems is that they can suffer from a limited content issue, where they only recommend items that are very similar to each other. This can result in a lack of diversity in the recommendations and lower overall user satisfaction.	Novelty and Diversity	They suggest novel, unpredictable, and startling objects that may be loved by consumers and may help make up for the lack of content.

Table 4. Summarization of contribution regarding the scalability problem.

Contribution	Definitions	Metrics	Solution proposed
(Ishtiaq et al., 2016)	Cold start users and insufficient data definition.	Accuracy	This introduces a novel method for generating recommendations that is both accurate and scalable. The algorithm employs various techniques for scalability to minimize processing demands and produce recommendations based on a vast quantity of data.
(Su et al., 2022b)	The authors evaluate their approach on several real-world datasets and show that it can significantly reduce the computational cost of distance-based link prediction algorithms while maintaining high prediction accuracy. They also compare their approach to other similarity selection methods and demonstrate its superiority in terms of both efficiency and effectiveness.	Accuracy	The authors propose a new method for similarity selection called Cluster-based Similarity Selection (CSS), which partitions the items into clusters and selects a representative item from each cluster. The similarities are then calculated only between the representative items, rather than between all pairs of items.

Table 5. Summarization of contribution regarding the scalability problem.

Contribution	Definitions	Metrics	Solution proposed
(Kim et al., 2017)	The authors aim to improve recommendation systems by analyzing the correlation between data collected from different types of content, specifically movies and music, which were gathered simultaneously.	Novelty and diversity	Despite their differences in content, folksonomy tags for music are considered by them to be associated data with movie genres.
(Rodríguez-García et al., 2019)	The article presents a novel approach to the problem of providing personalized dating recommendations using ontology-based modeling and context-aware techniques. The BlindDate Recommender platform has the potential to improve the online dating experience for users by providing them with more relevant and personalized recommendations.	-	The platform is designed to provide users with personalized dating recommendations based on their preferences and interests. The platform uses an ontology-based approach to model the domain of dating and to represent user preferences and interests. The article describes the architecture of the BlindDate Recommender platform and the various components that make up the platform. The article also discusses the evaluation of the platform using a dataset of real user profiles. The evaluation results show that the BlindDate Recommender platform provides accurate and effective dating recommendations to users.

3. Research Contribution

3.1 Motivation

Relying solely on accuracy for evaluating a recommendation system may result in the system suggesting redundant options to the user. The reason is that a system solely focused on accuracy would give priority to recommending items similar to those the user has already consumed, instead of offering novel and varied options.

Integrating metrics like novelty, diversity, unexpectedness, utility, usefulness, relevance, and popularity into recommendation systems can potentially improve user satisfaction and involvement.

- Novelty: Recommending new and unique items to users can enhance their experience by exposing them to a wider range of content.
- Diversity: Recommending a diverse set of items can broaden users' horizons and prevent them from being trapped in a filter bubble, where they only see content that reinforces their existing beliefs or preferences.
- Unexpectedness: Recommending items that are unexpected but still relevant to users' interests can provide a pleasant surprise and increase their engagement with the system.
- Utility and usefulness: Recommending items that are relevant to users' needs and preferences can increase the likelihood that they will find the recommendations useful and continue to use the system.
- Relevance: Recommending items that are closely related to users' interests can improve the relevance of the recommendations and increase their satisfaction with the system.
- Popularity: Recommending popular items can increase users' trust in the system and provide a social validation effect, where users are more likely to engage with items that others have enjoyed.

Serendipity has a significant role in improving recommendation systems by preventing redundancy and enhancing user experience. If a recommendation system relies solely on accuracy, it may suggest items that the user has already encountered, resulting in a monotonous and tedious experience. Conversely, the inclusion of serendipity into a recommendation system can stimulate the exploration of new and surprising options, maintaining the user's interest in the system's recommendations. Consequently, this can elevate user satisfaction and involvement in the recommendation system. To summarize, although accuracy is crucial for recommendation systems, integrating serendipity into the assessment process can offer users a more varied and captivating experience while mitigating the issue of recommendation redundancy.

3.2 Demonstration of the Choice of One Solution

Recommendation systems play a crucial role in helping users navigate the vast amounts of content available online, but striking a balance between personalized recommendations and unexpected discoveries can be tricky. Serendipity is an important aspect of recommendation systems, as it allows users to encounter content that lies beyond their typical preferences but is still enjoyable and relevant to them. The idea is to introduce users to new and captivating items while keeping recommendations fresh and engaging.

A serendipitous recommendation can broaden a user's horizon and expose them to new interests. However, implementing serendipitous recommendations while maintaining personalized recommendations that cater to the user's interests can be challenging for recommendation system designers. Despite the challenge, achieving a balance between these two types of recommendations can lead to a more engaging and enriching user experience.

Content-based filtering has some limitations, such as synonym sparsity, which refers to the situation where the system fails to find similar items due to a lack of synonyms or similar terms in the features or attributes used for recommendation. An ideal solution to mitigate this disadvantage would be to use a combination of techniques that overcome the limitations of content-based filtering.

3.3 Contribution

To sum up, the article describes the following key contributions:

- A precise explanation of all content-based filtering drawbacks.
- A new solution that combines different approaches to generate unexpected recommendations is designed and tested.
- We named our proposed approach an **I**deal **S**olution **M**itigating **C**ontent-based filtering **D**rawbacks$ISMCD_{3P}$, which uses three phases that best evaluate the recommender systems rather than precision and diminish the monotony.

- We conducted a comparison between our proposed model and other advanced serendipity recommender systems and exhibited the practicality, technical accuracy, and consistent performance of our model.
- We assessed the effectiveness of our recommender system in movie Lens application scenarios and displayed that using the established criteria, the recommendation process can significantly enhance the recommendation quality, not just limited to precision.

4. The proposed recommender system model

Our proposed model describes an Ideal Solution Mitigating Content Disadvantages based on Three Phases called$ISMCD_{3P}$. This section is divided into two sections the first 3.1 represents the aim of the study, and the second 3.2 produces detailed steps in our suggested methodology.

4.1 Aim of the study

Our interest is to find a general solution for dealing with all content-based filtering drawbacks. Our goal is to provide a general model to follow by categorizing by phase each solution that can contain. Fig. 6 describes the suitable solution for each issue.

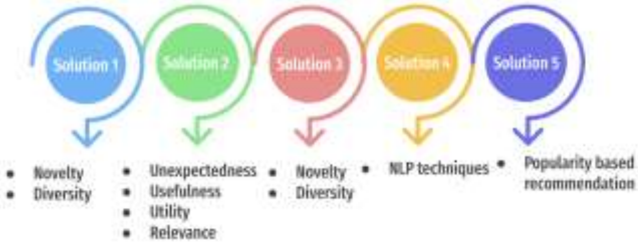

Fig. 6. Content-based filtering drawbacks.

4.2 The Proposed Architecture

Instead of selecting individual products to create a list of recommendations, $\boldsymbol{ISMCD_{3P}}$ prioritizes the overall composition of the recommendation list. Its primary principle involves systematically evaluating the entire set of suggestions and presenting customers with new products that align with their interests. Algorithm 1 outlines the key steps of $\boldsymbol{ISMCD_{3P}}$ main procedure.

DOI: 10.6977/IJoSI.202309_7(7).0006
O. Stitini, S. Kaloun, etc./Int. J. Systematic Innovation, 7(7),89-111(2023)

Algorithm 1 The main procedure of $\boldsymbol{ISMCD_{3P}}$

Input: User preferences.
Output: Recommendation List.

1: Phase 1 NLP Techniques
1.a: Generate the initial population that contains cleaned text with Punctuation removal.
1.b: Convert all text to lowercase.
1.c: Remove common words (stop-words).
1.d: Deal with emojis by either removing them or converting them to textual representations.
1.e: Eliminate words that are not relevant to the context or topic.
1.f: Correct misspelled words to their appropriate spelling.
1.g: Reduce words to their base or root form by removing suffixes or prefixes.
2: Popularity
2.a: New user
2.b: Address the ability of our proposed solution to handle increasing amounts of data, users, or resources without compromising performance or functionality.
2.c: Deal with situations where data is sparse, meaning there are a large number of missing or empty values, which can pose challenges for analysis or modeling.
3: Metrics Applications
3.a: Measures the degree of novelty of the recommended list by assessing how different it is from user preferences given in the input.
3.b: Evaluate the variety or range of different elements, options, or perspectives within a set or system.
3.c: Assess the level of surprise or deviation from expectations that are in particular results.
3.d: Determine the degree of significance, applicability, or pertinence of a particular item, in relation to a specific context given in the input.

4.3 Methodology and overall approach

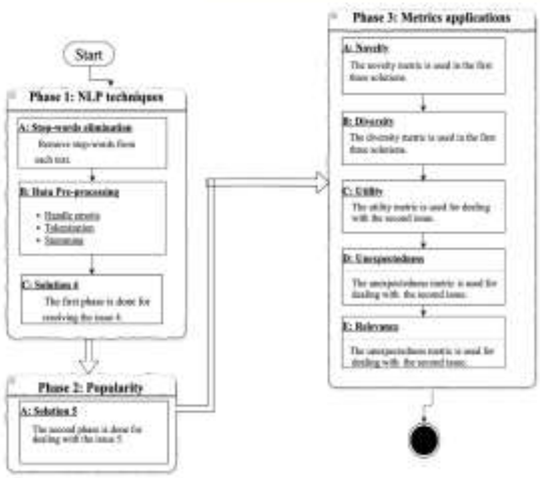

Fig. 7. Architecture of the proposed approach.

We started with the first phase which deals with the synonym drawback by applying NLP techniques, moving to the second phase which concentrates on resolving the problem of scalability or data sparsity, for example when we have a new user. The last phase solves the first three issues which converge in the same direction. Fig. 7 presents the architecture of our proposed architecture.

For that, our proposed technique $ISMCD_{3P}$ contains a three-phase process to deal with all content-based filtering challenges starting with NLP techniques, moving to the popularity-based recommendation, and finishing by applying new metrics.

1. Phase 1: NLP techniques

We utilized data preparation methods on our existing data to reduce its size. Texts of raw feelings are unstructured data sources containing noisy information. Before disabling the template's functionality, the raw text must first undergo pre-processing. The text may be transformed in various ways to make it modeling-ready. We then left off the punctuation.

The capital letter information was then converted into the text's lowercase. We have deleted the pointless stop words in a language and created noise when employed as text classification characteristics. The sentences are then returned to their original forms. We used the data preprocessing techniques on the news Twitter articles to reduce the amount of accurate data. Unstructured sources of information like raw news texts may include distracting material. The basic text must be preprocessed until the functionality of the model is eliminated.

There are several ways to modify the text so it can be modeled. Fig. 7 shows a general outline of our first phase in the proposed model which contains eight steps: Punctuation removal, lower Casing, stop-words removal, handling emojis, removing irrelevant words, spelling correction, stemming, and tokenization. By using this step we can deal with the synonym issue, exactly the problem number 4 as it is mentioned in Fig. 5.

Fig. 8. The main steps in phase 1.

2. Phase 2: Popularity

The cold start problem, which complicates the system by not possessing any past rating history and covers three cases—recommending a new user, recommending a new product, and recommending a new product to a new user—is the most concerning of these difficulties. Content-based systems make an effort to provide suggestions based on the target user's ratings and the features connected to the specific item. Only things with a past rating history, or those that have been rated before, are eligible for this suggestion strategy. This method is impossible to produce an effective result without a prior rating history. By integrating the popularity which means the use of demographic information we can mitigate the new user issue, exactly the problem number 5 as it is mentioned in Fig. 5.

3. Phase 3: Metrics applications

Over-specialization, limited content, and serendipity all these issues converge to the same problem meaning. The accuracy of predictions and the applicability of suggestions have always been the main topics of research in recommender systems. However, the effectiveness of a recommender system as a whole and user happiness may be significantly impacted by other suggestion quality indicators. As a result, current research in this area has focused more on other recommender system goals. For that, we propose these three issues, especially problems 1,2, and 3 as it is mentioned in Fig. 5 in one solution summarized in Fig. 9.

Fig. 9. Proposed solution for the first third issues mentioned in Fig. 5.

5. Experimental results

Evaluating feedback systems includes addressing an issue and utilizing an assessment technique to determine how the problem has been resolved. For recommender systems to be useful, a problem must have a solution. The problem must be precisely explained to establish if the issue has been resolved. This section shows how the proposed strategy was tested. To compare the proposed recommender system to other recommendation methods, many experiments were run:

- Content-based filtering: Based on the cosine similarity, this recommendation method produced the suggestions.
- The use of clustering: To improve the suggestion, they employ clustering, particularly the k-means method.

Table 6. Comparison of our proposed approach with the recommender system approach.

Properties	**Content-Based Filtering**	**Clustering**	**Our approach**
Effect of over-specialization	High	Medium	Less
Effect of serendipity	High	Medium	Less
Effect of limited content	High	Medium	Less
Synonym	High	Medium	Less
Cold start	High	Medium	Less

Table 6 describes the comparison between our proposed method and other recommender system approaches in terms of over-specialization, serendipity, limited content, synonym, and cold start, which is high if we use the classic content-based filtering and becomes higher when using our proposed approach.

5.1 Phase 1: NLP techniques

5.1.1 Solution 4 Interpretation

Table 7 above shows the transformation of the raw dataset into an understandable format using the eight steps mentioned in Fig. 7.

Table 7. The transition steps in data pre-processing.

Before pre-processing	**After pre-processing**
A conseiller +++	a conseil
Excellent	excel
Excellent rapport qualité prix	excel rapport qualité prix

5.2 Phase 2: Popularity

5.2.1 Solution 5 interpretation

Table 8. Popular recommendation example.

Domain	**New user**	**Top-N popular recommendations**
Movie recommendation	New user 1	["The Godfather", "The Dark Knight", "Fig. . ht Club"]
Book recommendation	New user 2	["The Hunger Games", "Harry Potter", "A Fairy Story"]
Hotel recommendation	New user 3	["Hotel Ekta", "Le Domaine de La Reserve","Château de Roncourt"]

Table 8 shows an example of the second phase in our proposed approach.

5.3 Phase 3: metrics applications

5.3.1 Interpretation of Solutions 1,2, and 3

Table 9. Novelty results of the recommendation methods.

Method Recommendation	**K =1**	**K =3**	**K=5**	**K=7**	**K=9**	**K=11**
Content-based filtering	0.285	0.289	0.297	0.316	0.328	0.410
Our third phase proposition	**0.845**	**0.740**	**0.726**	**0.6329**	**0.602**	**0.533**

Table 9 demonstrates the obtained novelty findings. The reader may see that, when compared to previous recommendation systems, our third phase proposal methodology exhibits notable improvements. The term "K" represents the range of recommendations, starting from one recommendation (k=1) and extending up to eleven recommendations (k=11). In Top 1 and 3, the uniqueness of the third phase's recommended method reaches its peak, which starts to decline. The outcomes show how practical the suggested approach is. Otherwise, the third phase recommended technique outperformed the content-based recommendation method in terms of originality by an average of 56%.

Table 10. Comparison of precision results of both methods.

Method Recommendation	**K=1**	**K=3**	**K=5**	**K=7**	**K=9**	**K=11**
Using CB filtering	0.641	0.640	0.643	0.636	0.615	0.643
Our third phase proposition	0.740	0.740	0.739	0.738	0.736	0.735

Table 11. Comparison of Recall results of both methods.

Method Recommendation	**K=1**	**K=3**	**K=5**	**K=7**	**K=9**	**K=11**
Using CB filtering	0.238	0.235	0.245	0.236	0.231	0.239
Our third phase proposition	0.673	0.671	0.670	0.670	0.698	0.698

Tables 10 and 11 show the outcomes of the recommendation techniques recall and accuracy. The authors conclude from these data that all recommendation algorithms performed better as the number of Top-N recommendations increased. This is because recall shows the proportion of the collection's favorite suggested things out of all choices. As a result, the likelihood of proposing goods to users that they will find exciting rises as the number of recommended items increases.

6. Distinctions

Table 12 shows a comparative study between the novel proposed approach and other existing ones.

Table 12. Works Comparison.

Works	**Over-specialization**	**Limited content**	**Serendipity**	**Synonym**	**Scalability**
(Stitini et al., 2022i)	✓				
(De Gemmis et al., 2015e)			✓		
(Saat et al., 2018b)			✓		
(Kotkov et al., 2020b)			✓		
(Adamopoulos & Tuzhilin, 2014b)		✓			
(Su et al., 2022c)					✓
(Ishtiaq et al., 2016)					✓
(Isinkaye et al., 2015)				✓	
(Kim et al., 2017)				✓	
$ISMCD_{3P}$	✓	✓	✓	✓	✓

7. Discussion

This study's primary goal is to evaluate content-based filtering problems and provide a single fix that eliminates all of their downsides. Therefore, our suggested method looks for a recommendation list that matches three essential criteria:

- Lack of synonymous words in a recommendation list.
- The suggestions list contains novel and serendipitous items.

Lack of data sparsity by integrating popularity into the recommendation list.

Our proposed method for generating a list that satisfies those criteria involves a combination of various techniques and algorithms, including natural language processing and cutting-edge algorithms, to ensure diversity in the recommendation list. The authors suggest that the performance of the method is influenced by the size of the dataset and the number of items suggested (i.e., the size of the individual). To achieve optimal results, the Top N should be selected carefully and based on empirical evidence, using a substantial dataset.

The key obstacle affecting content-based filtering is over-specialization or the limited content, or in other words serendipity issues. As a result, our proposed methodology $ISMCD3P$ intends to address all content-based filtering issues to increase suggestion quality and recommendation accuracy. The suggested methodology was tested on the MovieLens dataset.

8.Conclusion and future work

The rise of information overload has emphasized the importance of recommender systems, leading to an investigation into a new strategy for addressing the limitations of content-based recommenders. The goal of this research study was to develop a solution that would overcome these drawbacks by proposing a multi-task model for content-based filtering. This model utilizes a range of techniques to provide recommendations to users based on their interests. The main idea behind this approach is to create a list of highly connected items that are semantically related, while also introducing new and diverse content, taking into account potential new user and synonym words.

Although this research study presents a promising solution, it is important to note that it is limited by its use of only one dataset (MovieLens). However, future studies could address this limitation by incorporating additional datasets to expand the scope of the proposed multi-task model. Overall, this research provides a significant contribution to the development of recommender systems and paves the way for further advancements in this field.

References

Mistry, P., Maes, P., & Chang, L. (2009). WUW-wear Ur world: a wearable gestural interface. In CHI'09 extended abstracts on Human factors in computing systems (pp. 4111-4116).

Wang, R. Y., & Popović, J. (2009). Real-time hand-tracking with a color glove. ACM transactions on graphics (TOG), 28(3), 1-8.

Rautaray, S. S. (2012). Real time hand gesture recognition system for dynamic applications. International Journal of ubicomp (IJU), 3(1).

Saxe, D., & Foulds, R. (1996, September). Toward robust skin identification in video images. In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition (pp. 379-384). IEEE.

Chai, D., & Ngan, K. N. (1998, April). Locating facial region of a head-and-shoulders color image. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 124-129). IEEE.

Yang, J., Lu, W., & Waibel, A. (1997). Skin-color modeling and adaptation. In Computer Vision—ACCV'98: Third Asian Conference on Computer Vision Hong Kong, China, January 8–10, 1998 Proceedings, Volume II 3 (pp. 687-694). Springer Berlin Heidelberg.

Yu, C., Wang, X., Huang, H., Shen, J., & Wu, K. (2010, October). Vision-based hand gesture recognition using combinational features. In 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (pp. 543-546). IEEE.

Rekha, J., Bhattacharya, J., & Majumder, S. (2011, December). Shape, texture, and local movement hand gesture features for Indian sign language recognition. In 3rd international conference on trends in information sciences & computing (TISC2011) (pp. 30-35). IEEE.

Panwar, M., & Mehra, P. S. (2011, November). Hand gesture recognition for human computer interaction. In 2011 International Conference on Image Information Processing (pp. 1-7). IEEE.

Malima, A. K., Özgür, E., & Çetin, M. (2006). A fast algorithm for vision-based hand gesture recognition for robot control.

Manigandan, M., & Jackin, I. M. (2010, June). Wireless vision based mobile robot control using hand gesture recognition through perceptual color space. In 2010 International Conference on Advances in Computer Engineering (pp. 95-99). IEEE.

Fang, Y., Wang, K., Cheng, J., & Lu, H. (2007, July). A real-time hand gesture recognition method. In 2007

IEEE International Conference on Multimedia and Expo (pp. 995-998). IEEE.

Dardas, N. H., & Georganas, N. D. (2011). Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques. IEEE Transactions on Instrumentation and Measurement, 60(11), 3592-3607.

Rehg, J. M., & Kanade, T. (1994, November). Digiteyes: Vision-based hand tracking for human-computer interaction. In Proceedings of 1994 IEEE workshop on motion of non-rigid and articulated objects (pp. 16-22). IEEE.

Hsieh, C. T., Yeh, C. H., Hung, K. M., Chen, L. M., & Ke, C. Y. (2012, June). A real time hand gesture recognition system based on DFT and SVM. In 2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012) (Vol. 3, pp. 490-494). IEEE.

Tewari, D., & Srivastava, S. K. (2012). A visual recognition of static hand gestures in Indian sign language based on Kohonen self-organizing map algorithm. International Journal of Engineering and Advanced Technology (IJEAT), 2(2), 165-170.

Ramli, S. (2012, July). GMT feature extraction for representation of BIM sign language. In 2012 IEEE Control and System Graduate Research Colloquium (pp. 43-48). IEEE.

Utsumi, A., & Ohya, J. (1998, June). Image segmentation for human tracking using sequential-image-based hierarchical adaptation. In Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231) (pp. 911-916). IEEE..

Blake, A., North, B., & Isard, M. (1998). Learning multi-class dynamics. Advances in neural information processing systems, 11.

Yang, M. H., & Ahuja, N. (1998, October). Detecting human faces in color images. In Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269) (Vol. 1, pp. 127-130). IEEE.

Sigal, L., Sclaroff, S., & Athitsos, V. (2004). Skin color-based video segmentation under time-varying illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence, 26(7), 862-877.

Yoon, H. S., Soh, J., Bae, Y. J., & Yang, H. S. (2001). Hand gesture recognition using combined features of location, angle, and velocity. Pattern recognition, 34(7), 1491-1501.

Rahman, M. A., Purnama, I. K. E., & Purnomo, M. H. (2014, August). Simple method of human skin detection using HSV and YCbCr color spaces. In 2014 International Conference on Intelligent Autonomous Agents, Networks and Systems (pp. 58-61). IEEE.

Yu, Y., Bi, S., Mo, Y., & Qiu, W. (2016, June). Real-time gesture recognition system based on Camshift algorithm and Haar-like feature. In 2016 IEEE International Conference on cyber technology in Automation, control, and Intelligent Systems (CYBER) (pp. 337-342). IEEE.

Kaur, S., & Nair, N. (2018, January). Electronic Device Control Using Hand Gesture Recognition System for Differently Abled. In 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 371-375). IEEE.

Mahmood, M. R., Abdulazeez, A. M., & Orman, Z. (2018, October). Dynamic hand gesture recognition system for Kurdish sign language using two lines of features. In 2018 International Conference on Advanced Science and Engineering (ICOASE) (pp. 42-47). IEEE.

Heenaye-Mamode Khan, M., Ittoo, N., & Coder, B. K. (2019). Hand Gestures Categorisation and Recognition. In Information Systems Design and Intelligent Applications: Proceedings of Fifth International Conference INDIA 2018 Volume 2 (pp. 295-304). Springer Singapore.

Rautaray, S. S., & Agrawal, A. (2015). Vision based hand gesture recognition for human computer interaction: a survey. Artificial intelligence review, 43, 1-54.

Rehg, J. M., & Kanade, T. (1995, June). Model-based tracking of self-occluding articulated objects. In Proceedings of IEEE International Conference on Computer Vision (pp. 612-617). IEEE.

Shimada, N., Shirai, Y., Kuno, Y., & Miura, J. (1998, April). Hand gesture estimation and model refinement using monocular camera-ambiguity limitation by inequality constraints. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 268-273). IEEE.

Wu, Y., & Huang, T. S. (1999, September). Capturing articulated human hand motion: A divide-and-conquer approach. In Proceedings of the seventh IEEE International Conference on computer vision (Vol. 1, pp. 606-611). IEEE.

Wu, Y., Lin, J. Y., & Huang, T. S. (2001, July). Capturing natural hand articulation. In Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001 (Vol. 2, pp. 426-432). IEEE.

Lee, J., & Kunii, T. L. (1995). Model-based analysis of hand posture. IEEE Computer Graphics and Applications, 15(5), 77-86.

Heap, T., & Hogg, D. (1996, October). Towards 3D hand tracking using a deformable model. In Proceedings of the Second International Conference on Automatic Face and Gesture Recognition (pp. 140-145). IEEE.

Cutler, R., & Turk, M. (1998, April). View-based interpretation of real-time optical flow for gesture recognition. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 416-421). IEEE.

Martin, J., Devin, V., & Crowley, J. L. (1998, April). Active hand tracking. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 573-578). IEEE.

Yuan, Q., Sclaroff, S., & Athitsos, V. (2005, January). Automatic 2D hand tracking in video sequences. In 2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)-Volume 1 (Vol. 1, pp. 250-256). IEEE.

Yang, Q. (2010, June). Chinese sign language recognition based on video sequence appearance modeling. In 2010 5th IEEE Conference on Industrial Electronics and Applications (pp. 1537-1542). IEEE.

Kim, I. C., & Chien, S. I. (2001). Analysis of 3D hand trajectory gestures using stroke-based composite hidden Markov models. Applied Intelligence, 15, 131-143.

Chaudhurya, S., Banerjeeb, S., Ramamoorthya, A., & Vaswania, N. (2000). Recognition of dynamic hand gestures. The Journal of The Pattern Recognition Society. Department of Electrical Engineering and Department of Computer Science Engineering, IIT Delhi.

Bhuyan, M. K., Ghoah, D., & Bora, P. K. (2006, September). A framework for hand gesture recognition with applications to sign language. In 2006 Annual IEEE India Conference (pp. 1-6). IEEE.

Chen, Q., & Zhu, F. (2018, July). Long term hand tracking with proposal selection. In 2018 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) (pp. 1-6). IEEE.

Guo, J. M., Liu, Y. F., Chang, C. H., & Nguyen, H. S. (2011). Improved hand tracking system. IEEE Transactions on Circuits and Systems for Video Technology, 22(5), 693-701.

Koh, E., Won, J., & Bae, C. (2009, May). On-premise skin color modeling method for vision-based hand tracking. In 2009 IEEE 13th International Symposium on consumer electronics (pp. 908-909). IEEE.

Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking. IEEE Transactions on pattern analysis and machine intelligence, 25(5), 564-577.

Jepson, A. D., Fleet, D. J., & El-Maraghi, T. F. (2003). Robust online appearance models for visual tracking. IEEE transactions on pattern analysis and machine intelligence, 25(10), 1296-1311.

Zhou, S. K., Chellappa, R., & Moghaddam, B. (2004). Visual tracking and recognition using appearance-adaptive models in particle filters. IEEE Transactions on Image Processing, 13(11), 1491-1506.

McKenna, S. J., Raja, Y., & Gong, S. (1999). Tracking colour objects using adaptive mixture models. Image and vision computing, 17(3-4), 225-231.

Binh, N. D., Shuichi, E., & Ejima, T. (2005). Real-time hand tracking and gesture recognition system. Proc. GVIP, 19-21.

Imagawa, K., Lu, S., & Igi, S. (1998, April). Color-based hands tracking system for sign language recognition. In Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 462-467). IEEE.

Isard, M., & Blake, A. (1998). CONDENSATION--conditional density propagation for visual tracking. International journal of computer vision, 29(1), 5.

Shan, C., Tan, T., & Wei, Y. (2007). Real-time hand tracking using a mean shift embedded particle filter. Pattern recognition, 40(7), 1958-1970.

Weng, S. K., Kuo, C. M., & Tu, S. K. (2006). Video object tracking using adaptive Kalman filter. Journal of

Visual Communication and Image Representation, 17(6), 1190-1208.

Zhang, Q. Y., Zhang, M. Y., & Hu, J. Q. (2009). Hand Gesture Contour Tracking Based on Skin Color Probability and State Estimation Model. Journal of Multimedia, 4(6).

Zheng, W., & Bhandarkar, S. M. (2009). Face detection and tracking using a boosted adaptive particle filter. Journal of Visual Communication and Image Representation, 20(1), 9-27.

Utsumi, A., & Ohya, J. (1999, June). Multiple-hand-gesture tracking using multiple cameras. In Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149) (Vol. 1, pp. 473-478). IEEE.

Peterfreund, N. (1999). Robust tracking of position and velocity with Kalman snakes. IEEE transactions on pattern analysis and machine intelligence, 21(6), 564-569.

Mohd Asaari, M. S., Rosdi, B. A., & Suandi, S. A. (2015). Adaptive Kalman Filter Incorporated Eigenhand (AKFIE) for real-time hand tracking system. Multimedia Tools and Applications, 74, 9231-9257.

Isard, M., & Blake, A. (1998, January). A mixed-state condensation tracker with automatic model-switching. In Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271) (pp. 107-112). IEEE.

Mammen, J. P., Chaudhuri, S., & Agrawal, T. (2001, September). Simultaneous Tracking of Both Hands by Estimation of Erroneous Observations. In BMVC (pp. 1-10).

Pérez, P., Hue, C., Vermaak, J., & Gangnet, M. (2002). Color-based probabilistic tracking. In Computer Vision—ECCV 2002: 7th European Conference on Computer Vision Copenhagen, Denmark, May 28–31, 2002 Proceedings, Part I 7 (pp. 661-675). Springer Berlin Heidelberg.

Bhuyan, M. K., Ghosh, D., & Bora, P. K. (2006, June). Feature extraction from 2D gesture trajectory in dynamic hand gesture recognition. In 2006 IEEE Conference on Cybernetics and Intelligent Systems (pp. 1-6). IEEE.

Nadgeri, S. M., Sawarkar, S. D., & Gawande, A. D. (2010, November). Hand gesture recognition using CAMSHIFT algorithm. In 2010 3rd International Conference on Emerging Trends in Engineering and Technology (pp. 37-41). IEEE.

Bradski, G. R. (1998, October). Real time face and object tracking as a component of a perceptual user interface. In Proceedings Fourth IEEE Workshop on Applications of Computer Vision. WACV'98 (Cat. No. 98EX201) (pp. 214-219). IEEE.

Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. " O'Reilly Media, Inc.".

Wang, X., & Li, X. (2010, December). The study of MovingTarget tracking based on Kalman-CamShift in the video. In The 2nd International Conference on Information Science and Engineering (pp. 1-4). IEEE.

Huang, S., & Hong, J. (2011, April). Moving object tracking system based on camshift and Kalman filter. In 2011 International Conference on Consumer Electronics, Communications and Networks (CECNet) (pp. 1423-1426). IEEE.

Shi, J. (1994, June). Good features to track. In 1994 Proceedings of IEEE conference on computer vision and pattern recognition (pp. 593-600). IEEE.

Kolsch, M., & Turk, M. (2004, June). Fast 2D hand tracking with flocks of features and multi-cue integration. In 2004 Conference on Computer Vision and Pattern Recognition Workshop (pp. 158-158). IEEE.

Porikli, F., Tuzel, O., & Meer, P. (2006, June). Covariance tracking using model update based on means on Riemannian manifolds. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition (Vol. 1, pp. 728-735).

Tuzel, O., Porikli, F., & Meer, P. (2006). Region covariance: A fast descriptor for detection and classification. In Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006. Proceedings, Part II 9 (pp. 589-600). Springer Berlin Heidelberg.

Black, M. J., & Jepson, A. D. (1998). Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. International Journal of Computer Vision, 26, 63-84.

Xiu, C., Su, X., & Pan, X. (2018, June). Improved target tracking algorithm based on Camshift. In 2018 Chinese Control and decision conference (CCDC) (pp. 4449-4454). IEEE.

Elmezain, M., Al-Hamadi, A., Appenrodt, J., & Michaelis, B. (2008, December). A hidden Markov model-based continuous gesture recognition system for hand motion trajectory. In 2008 19th International Conference on pattern recognition (pp. 1-4). IEEE.

Kao, C. Y., & Fahn, C. S. (2011). A human-machine interaction technique: hand gesture recognition based on hidden Markov models with trajectory of hand motion. Procedia Engineering, 15, 3739-3743.

Xu, D., Wu, X., Chen, Y. L., & Xu, Y. (2015). Online dynamic gesture recognition for human robot interaction. Journal of Intelligent & Robotic Systems, 77(3-4), 583-596.

Elmezain, M., Al-Hamadi, A., & Michaelis, B. (2009). Hand gesture recognition based on combined features extraction. International Journal of Electrical and Computer Engineering, 3(12), 2389-2394.

Bhuyan, M. K., Bora, P. K., & Ghosh, D. (2008). Trajectory guided recognition of hand gestures having only global motions. World Academy of science, engineering, and technology, 21, 753-764.

Bhuyan, M. K., Ajay Kumar, D., MacDorman, K. F., & Iwahori, Y. (2014). A novel set of features for continuous hand gesture recognition. Journal on Multimodal User Interfaces, 8, 333-343.

Li, K., Zhou, Z., & Lee, C. H. (2016). Sign transition modeling and a scalable solution to continuous sign language recognition for real-world applications. ACM Transactions on Accessible Computing (TACCESS), 8(2), 1-23.

Rubine, D. (1991). Specifying gestures by example. ACM SIGGRAPH computer graphics, 25(4), 329-337.

Signer, B., Norrie, M. C., & Kurmann, U. (2011). iGesture: A Java framework for the development and deployment of stoke-based online Gesture recognition algorithms. Technical Report/ETH Zurich, Department of Computer Science, 561.

Premaratne, P., Yang, S., Vial, P., & Ifthikar, Z. (2017). Centroid tracking based dynamic hand gesture recognition using discrete Hidden Markov Models. Neurocomputing, 228, 79-83.

Cen, M., & Jung, C. (2017). Complex form of local orientation plane for visual object tracking. IEEE Access, 5, 21597-21604.

Tang, J., Cheng, H., Zhao, Y., & Guo, H. (2018). Structured dynamic time warping for continuous hand trajectory gesture recognition. Pattern Recognition, 80, 21-31.

Singla, A., Roy, P. P., & Dogra, D. P. (2019). Visual rendering of shapes on 2D display devices guided by hand gestures. Displays, 57, 18-33.

Misra, S., & Laskar, R. H. (2019). Development of a hierarchical dynamic keyboard character recognition system using trajectory features and scale-invariant holistic modeling of characters. Journal of Ambient Intelligence and Humanized Computing, 10(12), 4901-4923.

Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. IEEE assp magazine, 3(1), 4-16.

Charniak, E. (1993). Statistical language learning MIT Press. Google Scholar Google Scholar Digital Library Digital Library.

Chen, F. S., Fu, C. M., & Huang, C. L. (2003). Hand gesture recognition using a real-time tracking method and hidden Markov models. Image and vision computing, 21(8), 745-758.

Lee, H. K., & Kim, J. H. (1999). An HMM-based threshold model approach for gesture recognition. IEEE Transactions on pattern analysis and machine intelligence, 21(10), 961-973.

Marcel, S., Bernier, O., Viallet, J. E., & Collobert, D. (2000, March). Hand gesture recognition using input-output hidden Markov models. In proceedings fourth IEEE International Conference on automatic face and gesture recognition (Cat. No. PR00580) (pp. 456-461). IEEE.

Just, A., & Marcel, S. (2009). A comparative study of two state-of-the-art sequence processing techniques for hand gesture recognition. Computer Vision and Image Understanding, 113(4), 532-543.

Yang, R., & Sarkar, S. (2006, August). Detecting coarticulation in sign language using conditional random fields. In 18th International conference on pattern recognition (ICPR'06) (Vol. 2, pp. 108-112). IEEE.

Beh, J., Han, D., & Ko, H. (2014). Rule-based trajectory segmentation for modeling hand motion trajectory. Pattern Recognition, 47(4), 1586-1601.

Siddiqi, S. M., Gordon, G. J., & Moore, A. W. (2007, March). Fast state discovery for HMM model selection and learning. In Artificial Intelligence and Statistics (pp. 492-499). PMLR.

Ulas, A., & Yildiz, O. T. (2009, December). An incremental model selection algorithm based on cross-validation for finding the architecture of a hidden Markov model on hand gesture data sets. In 2009 International Conference on Machine Learning and Applications (pp. 170-177). IEEE.

Bishop, C. M. (1995). Neural networks for pattern recognition. Oxford University Press.

Haykin, S. (2009). Neural networks and learning machines, 3/E. Pearson Education India.

Bamwenda, J., & Özerdem, M. S. (2019). Static hand gesture recognition system using artificial neural networks and support vector machine. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 10(2), 561-568.

Yang, M. H., Ahuja, N., & Tabb, M. (2002). Extraction of 2D motion trajectories and its application to hand gesture recognition. IEEE Transactions on pattern analysis and machine intelligence, 24(8), 1061-1074.

Yang, M. H., & Ahuja, N. (1998, June). Extraction and classification of visual motion patterns for hand gesture recognition. In Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231) (pp. 892-897). IEEE.

Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multiclass support vector machines. IEEE Transactions on Neural Networks, 13(2), 415-425.

Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. Neural processing letters, 9, 293-300.

Gopalan, R., & Dariush, B. (2009, October). Toward a vision based hand gesture interface for robotic grasping. In 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 1452-1459). IEEE.

Dominio, F., Donadeo, M., & Zanuttigh, P. (2014). Combining multiple depth-based descriptors for hand gesture recognition. Pattern Recognition Letters, 50, 101-111.

Thirumuruganathan, S. (2010). A detailed introduction to K-nearest neighbor (k-NN) algorithm. Retrieved March, 20, 2012.

Ge, S. S., Yang, Y., & Lee, T. H. (2008). Hand gesture recognition and tracking based on distributed locally linear embedding. Image and Vision Computing, 26(12), 1607-1620.

Oka, K., Sato, Y., & Koike, H. (2002). Real-time fingertip tracking and gesture recognition. IEEE Computer Graphics and Applications, 22(6), 64-71.

Davis, J., & Shah, M. (1994). Recognizing hand gestures. In Computer Vision—ECCV'94: Third European Conference on Computer Vision Stockholm, Sweden, May 2–6, 1994 Proceedings, Volume I 3 (pp. 331-340). Springer Berlin Heidelberg.

Yeasin, M., & Chaudhuri, S. (2000). Visual understanding of dynamic hand gestures. Pattern Recognition, 33(11), 1805-1817.

Hong, P., Turk, M., & Huang, T. S. (2000, March). Gesture modeling and recognition using finite state machines. In Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580) (pp. 410-415). IEEE.

Suk, H. I., Sin, B. K., & Lee, S. W. (2010). Hand gesture recognition based on dynamic Bayesian network framework. Pattern recognition, 43(9), 3059-3072.

Thai, L. H., Hai, T. S., & Thuy, N. T. (2012). Image classification using support vector machine and artificial neural network. International Journal of Information Technology and Computer Science, 4(5), 32-38.

Kang, S., & Park, S. (2009). A fusion neural network classifier for image classification. Pattern Recognition Letters, 30(9), 789-793.

Dinh, T. B., Dang, V. B., Duong, D. A., Nguyen, T. T., & Le, D. D. (2006, February). Hand gesture classification using boosted cascade of classifiers. In 2006 International Conference on Research, Innovation, and Vision for the Future (pp. 139-144). IEEE.

Burger, T., Aran, O., Urankar, A., Caplier, A., & Akarun, L. (2008). A Dempster-Shafer theory based combination of classifiers for hand gesture recognition. In Computer Vision and Computer Graphics. Theory and Applications: International Conference VISIGRAPP 2007, Barcelona, Spain, March 8-11, 2007. Revised Selected Papers (pp. 137-150). Springer Berlin Heidelberg.

Ng, C. W., & Ranganath, S. (2002). Real-time gesture recognition system and application. Image and Vision Computing, 20(13-14), 993-1007.

Wang, G. W., Zhang, C., & Zhuang, J. (2012). An application of classifier combination methods in hand gesture recognition. Mathematical Problems in Engineering, 2012.

Corradini, A. (2002, May). Real-time gesture recognition by means of hybrid recognizers. In Gesture and Sign Language in Human-Computer Interaction: International Gesture Workshop, GW 2001 London, UK, April 18–20, 2001 Revised Papers (pp. 34-47). Berlin, Heidelberg: Springer Berlin Heidelberg.

Sridevi, P., Islam, T., Debnath, U., Nazia, N. A., Chakraborty, R., & Shahnaz, C. (2018, December). Sign Language recognition for Speech and Hearing Impaired by Image processing in Matlab. In 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC) (pp. 1-4). IEEE.

Singha, J., Misra, S., & Laskar, R. H. (2016). Effect of variation in gesticulation pattern in dynamic hand gesture recognition system. Neurocomputing, 208, 269-280.

Mohammed, A. A., Minhas, R., Wu, Q. J., & Sid-Ahmed, M. A. (2011). Human face recognition based on multidimensional PCA and extreme learning machines. Pattern recognition, 44(10-11), 2588-2597.

Liu, H., Yu, L., Wang, W., & Sun, F. (2016). Extreme learning machine for time sequence classification. Neurocomputing, 174, 322-330.

Chen, X., & Koskela, M. (2015). Skeleton-based action recognition with extreme learning machines. Neurocomputing, 149, 387-396.

Yang, W., Liu, Y., Zhang, Q., & Zheng, Y. (2019). Comparative object similarity learning-based robust visual tracking. IEEE Access, 7, 50466-50475.

Wu, X. Y. (2020). A hand gesture recognition algorithm based on DC-CNN. Multimedia Tools and Applications, 79(13-14), 9193-9205.

Singha, J., & Laskar, R. H. (2016). Recognition of global hand gestures using self co-articulation information and classifier fusion. Journal on Multimodal User Interfaces, 10(1), 77-93.

Benitez-Garcia, G., Prudente-Tixteco, L., Castro-Madrid, L. C., Toscano-Medina, R., Olivares-Mercado, J., Sanchez-Perez, G., & Villalba, L. J. G. (2021). Improving real-time hand gesture recognition with semantic segmentation. Sensors, 21(2), 356.

Shanthakumar, V. A., Peng, C., Hansberger, J., Cao, L., Meacham, S., & Blakely, V. (2020). Design and evaluation of a hand gesture recognition approach for real-time interactions. Multimedia Tools and Applications, 79, 17707-17730.

Yadav, K. S., Misra, S., Khan, T., Bhuyan, M. K., & Laskar, R. H. (2020). Segregation of meaningful strokes, a pre-requisite for self co-articulation removal in isolated dynamic gestures. IET Image Processing, 15(5), 1166-1178.

Choudhury, A., Talukdar, A. K., Sarma, K. K., & Bhuyan, M. K. (2021). An adaptive thresholding-based movement epenthesis detection technique using hybrid feature set for continuous fingerspelling recognition. SN Computer Science, 2, 1-21.

AUTHOR BIOGRAPHIES

Oumaima STITINI She is currently a temporary Professor at the Cadi Ayyad University, Faculty of Science and Technology, and the Private University of Marrakech. She received his Ph.D. in Computer Science Engineering, especially in Robust optimization and implementation of recommendation systems based on artificial intelligence from the Faculty of Science and Technology in 2022. His main research interests include artificial intelligence, recommender systems, and IoT all of these applied to different fields like medicine, education, and entertainment.

Soulaimane KALOUN He is currently holding the position of a Permanent Associate Professor at the Faculty of Science and Technology located in Marrakech, Morocco. He earned his doctorate in Computer Science and is presently serving as a professor at the same institution. Moreover, he has also received an HDR in data science. Soulaimane's principal areas of research revolve around Big Data, machine learning, multiagent systems, and text mining.

BENCHAREF Omar He is currently a Permanent Professor at the Faculty of Science and Technology, Marrakech, Morocco. Omar received his Ph.D. in Computer Science and is currently a professor at the Faculty of Science and Technology, Marrakech, Morocco. He’s also HDR in data science. His main research interests include artificial intelligence (AI), data science, machine learning, multiagent systems, and text mining.