UNIVERSITY OF CAPE TOWN

MASTER'S DISSERTATION

A Hybrid Multi-Modal Recommender System using Neural Collaborative Filtering and Content Based Filtering

---

Author:
Pavan SINGH

Supervisor:
Assoc Professor Ian DURBACH
Dr. Allan E CLARK

A dissertation presented for the degree of
Master of Science Advanced Analytics

from the

Department of Statistical Sciences

August 25, 2023

Contents

Chapter 1

Introduction

In an age where there is an exponentially increasing volume of data being produced, the ability for a user to get the information they seek has become ever more challenging (SINTEF, 2013). The abundance of information which gives an overwhelming number of options to a user when making a decision is known as the information overload problem (Bawden and Robinson, 2020). With the vast amount of information available, it can be challenging for users to find items or products that meet their preferences or needs. Recommender systems are tools which directly aim to address the challenge of information overload (O'Donovan and Smyth, 2005). The general idea behind a recommender system is to narrow down the perceived available options and present to a user a limited set of personalised choices (recommendations) based on said user's preferences, behaviour and or other relevant factors (O'Donovan and Smyth, 2005). In this light, a recommender system can be seen as a filtration tool which greatly washes out undesirable results and brings forth content desired or more relevant to a user's current interests and needs. In this way they are able to help to reduce information overload and make the decision-making process easier and more efficient. A lucrative byproduct of this is an improved overall user experience, increased engagement and user satisfaction. Given these outcomes, recommender systems have become a highly researched area over recent years (Seth and Sharaff, 2022). So much so that they have become the backbone of many big technology companies particularly those in e-commerce. Recommender systems are a crucial tool in the arsenal of e-commerce platforms which try to help consumers navigate through an abundance of product options. To this end, recommender systems are capable of enhancing the customer experience by showing products that they are inclined to want and thus also boost sales of these product. The apparent effectiveness of these systems has led to this surge in research in this domain. In this thesis, we aim to directly contribute to this area of research by developing a hybrid recommender system that incorporates data from multiple modalities and investigates the potential impact of incorporating product review text and sentiment in improving the accuracy of recommendations.

The specific type of recommender systems used in this thesis are known as neural collaborative filtering and content based filtering. This chapter will introduce collaborative filtering and content based filtering recommender system paradigms in Section 1.1 and some other technologies which are gaining a lot of attention in the study area. In Section 1.2 we draw light to our primary research questions and objectives. The implications and significance of enhancing and incorporating text into recommenders for e-commerce applications are explained in Section 1.3 in detail before concluding the chapter by explaining the structure of the remaining sections of this paper in Section 1.4.

1.1 Recommender Systems: definition and types

A recommender system is a type of information filtering system that seeks to predict user preferences or interests by generating *recommended* items such as products, services or content that the user is likely to be interested in (Seth and Sharaff, 2022). Instead of sifting through irrelevant information and products, users are presented with content and products that are more likely to be of interest to them. Recommender Systems have quickly become a necessity, given that users cannot search through millions of content to connect with products, services or knowledge (i.e., items) that is important to them.

Recommender systems deal with two types of information: characteristic information, which includes details about the objects or products, such as keywords or categorisation; and user-item interactions, which encompass data like user ratings or likes and so on. With respect to the types of recommender

systems available, the data also plays an important role in determining which RS would be effective. In the field of recommender systems, there exist three main branches: collaborative filtering, content based filtering and hybrid systems (Thorat et al., 2015).
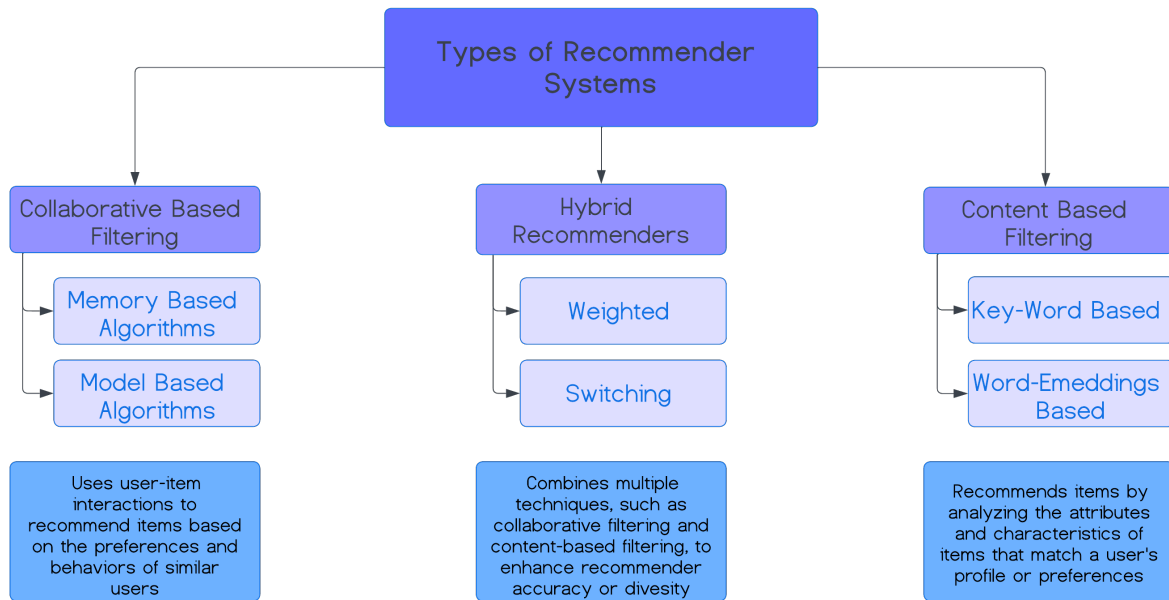

Figure 1.1: The types of recommender systems at the highest level with some popular sub-categories for each type.

Collaborative filtering, probably the most extensively implemented type of recommender, is based on user-item interactions data and helps you find what you like by looking for users, or items, who are similar to you (Thorat et al., 2015). So the algorithm is based on using a similarity measure to determine how much a user will like an item. At a high level, there are two types of collaborative filtering algorithms which have been proposed in the literature: memory-based algorithms, which rely on computing the similarity between users or items in order to make recommendations, and model-based algorithms, which rely on building a model or mathematical representation of the data to make recommendations. Both memory-based and model-based algorithms have their strengths and weaknesses which we discuss further later on.

Content based filtering, by contrast, is based on characteristic information and works by understanding the underlying features (or characteristics) of each item (Thorat et al., 2015). There is no general high level categories for which content based filtering algorithms are universally recognized as like we have for Collaborative based filtering (model based vs memory based). Regardless, the idea is that a content based filtering algorithm analyses the features of an item and then recommends other items with similar features. These features could be the attributes of an item, like the colour or style of a product, or they could be keywords, like the genres or artist names of a song.

We often find in practice that these two types of recommender systems techniques (collaborative and content based) are combined to form hybrid systems. These use both types of information, with the idea to leverage the strengths of each individual method to improve the accuracy and relevance of the recommendations, whilst also avoiding problems that are generated when working with just one kind of system (Thorat et al., 2015). These hybrid modelling approaches have been shown to outperform individual algorithms in many cases (Çano and Morisio, 2017). Whilst combining different models can often produce a more accurate model, there are other approaches that can be used to improve the accuracy and relevance of recommendations. One such approach is multi-modal recommender systems,

which leverage multiple types of data or information sources to make recommendations. Figure 1.1 shows at a very high level, the main types of recommender systems and the popular subcategories within these branches. Note that hybrid recommender, again, do not necessarily have recognised high level categories, however some approaches can be bundled into some high level categories such as weighted approaches.

Multi-modal refers to the integration of different types or modes of data in a system (Truong et al., 2021). In the context of recommender systems, multi-modal refers to the use of multiple types of data or information sources to make recommendations. By combining information from different modalities, such as text, images, audio, and video, multi-modal recommender systems can provide a more comprehensive understanding of user preferences and offer more personalised recommendations. Using multi-modalities also poses the opportunity to alleviate data sparsity (footnote: the situation where the available data is highly incomplete, resulting in numerous missing values in the user-item interactions. This occurs due to the vast number of items and limited user interactions, making accurate predictions for less observed items challenging) by leveraging or including auxiliary information that may encode additional clues on how users consume items. Examples of such data (referred to as modalities) are social networks, item's descriptive text, or customer product reviews (Truong et al., 2021). Studies on multi-modal recommender systems have quickly become a hot topic in the field as many understood the potential upside of incorporating multiple data types into a prediction algorithm (Truong et al., 2021). In fact, multi-modal recommendation systems have become the industry standard and are widely used in across many domains as companies try to personalise their products and services to better meet the needs and preferences of their customers (Liu et al., 2023). For example, in e-commerce, a multimodal recommender system can take advantage of product images and descriptions, user reviews, and purchase history to make recommendations (Liu et al., 2023). Beyond multi-modalities recommender systems have evolved to incorporate deep learning technologies. Traditional recommender systems have shown to be effective in many cases, however the recent advances in deep learning over the past decade have opened up new opportunities for improving the accuracy and personalisation of recommendations.

Deep learning is a a subset of machine learning that deals with models that are comprised of multiple layers (hence the "deep" in Deep Learning). In the context of recommender systems, deep learning algorithms can be used to effective in extracting complex patterns and relationships from large amounts of data (He et al., 2017). This is particularly attractive since the goal is to predict user preferences based on past behaviours or interactions. We shall explore the application of deep learning in the context of our RS by employing a neural collaborative framework which essentially generalises the matrix factorisation approach (which is extensively used in collaborative filtering) to improve the accuracy of recommendations (He et al., 2017). Amazon, amongst many other internet giants, also use neural networks in their recommendation engine to make product recommendations (Steck et al., 2021).

Although these systems can easily be mistaken as simply a tool for information retrieval and data discovery - in industry their importance is unwavering. With the tremendous amount of data available online, recommender systems have become the forefront of large internet corporations research and investment (Steck et al., 2021).

1.2   Research Problem and Objectives

The aim of this thesis is to directly enhance the predictive accuracy of traditional recommender systems by developing a hybrid model that incorporates data from multiple modalities, exploiting explicit numeric product ratings and product review text data. While existing recommender systems have shown promise in generating accurate recommendations, they often fail to account for the nuances and complexities of user preferences that may be expressed through textual data. Therefore, this thesis seeks to investigate the potential of incorporating product review text and review text sentiment as additional

sources of information to improve the accuracy and relevance of recommendations. This is done by building two individual models, a neural collaborative filtering RS to process the explicit ratings data and a content based filtering RS to handle the textual features (product reviews and review sentiment analysis), and combine them to form a hybrid recommender system. Building these two neural network architectures (for collaborative filtering and content based filtering) the system can learn a non-linear mapping between the input features (i.e., item attributes) and the target output (i.e., user preferences). Through comparative evaluation with benchmark recommender models, the efficacy of the proposed hybrid model will be assessed, with a particular focus on its ability to provide personalised recommendations that better match user preferences. To this end, a number of different models will be trained, evaluated and compared to one another in an effort to establish the performance benefits of different models. The research questions presented in Section 1.3 will guide this process.

Ultimately, this thesis aims to contribute to the development of more effective and efficient recommender systems that can potentially leverage the additional textual information provided by users for products which can potentially aid in generating recommendations tailored to the unique preferences and needs of individual users.

1.3 Research Questions and Significance

This paper seeks to explore the impact of incorporating product reviews into a recommender system model whilst also evaluating the performance of the hybrid model. Recommender systems have become pivotal to the way companies interact and sell their products to their consumers. A RS's ability to accurately predict consumer interests and desires on a highly personalised level make them a very valuable tool for content and product providers like Amazon, Google amongst other technology giants.

The key research questions addressed in this paper are discussed and explained below.

1. How does incorporating product reviews into a recommender system model impact the accuracy of the system's recommendations?

Here we are interested in determining whether integrating user reviews into a recommendation algorithm leads to more accurate and relevant product suggestions for users. Another way of looking at this is examining the accuracy of a multi-modal recommender system compared to that of a single-modal recommender system - i.e., a RS that relies on only one type of data. User reviews for products could potentially provide additional insights into user preferences or product features that are not captured by the metadata. However, reviews may also introduce noise or biases into the recommendation algorithm which can lead to less accurate suggestions. Findings from this study could provide further scope for incorporating review text into product recommendation designs or highlight potential pitfalls to avoid when integrating user reviews. The study also may prove to provide further support for multi-modal recommender systems.

2. How does incorporating product reviews as well as additional features from sentiment analysis into a recommender system impact the accuracy of the system's recommendations? Additionally, how does the accuracy of a hybrid recommender system vary based on the sentiment analysis techniques used to interpret the reviews?

Incorporating sentiment analysis into a recommender system introduces an avenue for the integration of additional features that can contribute to enhancing recommendation accuracy. Sentiment analysis, as a technique to assess the emotional tone and polarity of textual content such as product reviews (Medhat et al., 2014), offers a means of extracting valuable insights from user-generated content. By incorporating sentiment analysis outcomes alongside traditional recommendation factors, the system gains access to a richer set of attributes that can potentially capture nuanced user preferences and sentiments.

The impact of sentiment analysis on accuracy of recommender systems provides scope to further utilising review text in possibly improving recommendation accuracy. In addition, by comparing the accuracy of the system when different sentiment analysis techniques are used, the study can identify which techniques are most effective in improving recommendation accuracy. Potential findings from this study could provide insights into the most effective sentiment analysis techniques for hybrid recommender systems, and inform the design of more accurate and effective recommendation algorithms. Conversely, the study could highlight potential challenges and limitations of using sentiment analysis in a recommender system, such as accuracy issues due to sarcasm, ambiguity, or variations in language use.

3. How does the performance of a hybrid recommender system using neural collaborative filtering and content-based filtering compare to that of popular benchmark recommender systems and single-models?

The general idea is that hybrid models should combine the strengths of multiple recommendation algorithms to improve accuracy of recommendations, however this is not guaranteed. This is another key area of interest in our study - examining the potential capacity of a hybrid system to outperform other popular recommender systems. In evaluating our hybrid recommender system which comprises of a neural collaborative filtering component and a content based filtering component, we assess its prediction accuracy against other benchmark models. These models are listed and described in detail in Section 4. The results of our study can have huge ramifications. Hybrid models are by design more complex and often have greater computational expense and slower processing than non-hybrid models. By investigating the potential advantage these hybrid systems have could highlight the potential trade-off between the level of desired complexity and performance. Whilst this will be very domain specific and depend on the needs and resources available, the study will contribute to providing additional insight into this field.

4. What are the potential trade-offs of incorporating product reviews into a recommender system, such as increased complexity or potential biases in the recommendations?

Product reviews can often contain large amounts of unstructured data which is cause for concern for recommender. systems needing to quickly process and integrate this information in their algorithm. Another potential concern may be the bias inherent in user reviews. These will be discussed in depth later on in this paper. However, findings from this study could prove to be useful in highlighting the potential unearthed benefits and drawbacks of incorporating product reviews into a recommender system. Furthermore, it could open up paths for discussion on potential strategies for mitigating the trade-offs of using reviews in recommendation algorithms, such as developing more sophisticated sentiment analysis techniques.

These key research questions will be referenced greatly throughout the whole paper. The significance of the findings from the study ultimately provides further support for possible future design and implementations in this space. More specifically, findings will directly contribute to the existing literature on recommender systems and provide insights into the effectiveness and limitations of incorporating product reviews and sentiment analysis features in hybrid models, as well as the trade-offs associated with these approaches.

1.4 Dissertation Outline

In this chapter we introduced what recommender system are and the importance of them as tools to help users find the content and services they seek or may want. It has described the aim of this thesis and the research questions it will seek to answer. It has also introduced the primary hybrid multi-modal recommender system which we will be developing and examining in this thesis.

Chapter 2 is a detailed literature review which provides context for this research by first addressing the history and application of recommender systems in E-commerce. This history will start with a

brief exploration of the fundamental concepts which have remained relevant since their inception in Section 2.1. We shall also describe the immediate benefit and advantage that recommender systems have provided for E-commerce industry. From here we shall explore the two main types of recommender systems, collaborative filtering and content based filtering in Sections 2.2 and 2.3 respectively. Here, a brief history of the algorithms are discussed as well as there more recent advances. This will lead into the Deep Learning era in Section 2.4, where recent advances are of particular importance. Additionally we explore the prevalence of text and its applications for recommender systems in 2.5, Hybrid models in section 2.6 and finally the key evaluation methods used in recommender system research in Section 2.7.

Chapter 3 addresses all matters relating to data. In order to critically examine our hybrid recommender system and assess impacts of the textual features, it is necessary to have a source of product data. Our source is the Amazon Product Review dataset publicly available and is discussed in further detail in Section 3.1. In addition to describing and explaining the data collection process (Section 3.2), we explore the dataset in Section 3.3. Finally, we summarise our findings and partitioning process, as well as how we shall use the data and select features for our recommendation system in Sections 3.4 and 3.5 respectively.

Chapter 4 looks at the methods used in our research. It provides all the technical details necessary to understand the models utilised and trained in Chapter 5. We begin the chapter by detailing out our overall modelling approach in Section 4.1 before diving into the details. The groundwork for this will be built in Section 4.2 and 4.3, where the basic architecture and training procedure for neural collaborative filtering and content based filtering will be discussed in length, respectively. Additionally, Section 4.4 will discuss how we hybridise the model, including details on fusion process and the final model specification. Section 4.5 discusses. the various benchmark models that shall be built and compared with. Finally, Section 4.6 details the evaluation criteria that shall be employed to assess our recommender models.

Chapter 5 documents the implementation procedure in Section 5.1. The observed results will be presented in Section 5.2, where a first approximation at answering the research questions will be provided. This first approximation will be expanded upon in the discussion of Chapter 6. To that end, Chapter 6 will summarise and conclude the dissertation. This shall be done by first summarising the overall results in conjunction with our research question in Section 6.1.1 and then we also address the limitations and make suggestions for future work in Section 6.1.2. Finally we conclude the paper by discussing the implications and summarising the key findings of our work with respect to the research questions in Section 6.2.

Bibliography

Bawden, D., Robinson, L., 2020. Information overload: An overview .

Çano, E., Morisio, M., 2017. Hybrid recommender systems: A systematic literature review. Intelligent data analysis 21, 1487–1524.

He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T.S., 2017. Neural collaborative filtering , 173–182.

Liu, Q., Hu, J., Xiao, Y., Gao, J., Zhao, X., 2023. Multimodal recommender systems: A survey. arXiv preprint arXiv:2302.03883 .

Medhat, W., Hassan, A., Korashy, H., 2014. Sentiment analysis algorithms and applications: A survey. Ain Shams engineering journal 5, 1093–1113.

O'Donovan, J., Smyth, B., 2005. Trust in recommender systems , 167–174.

Seth, R., Sharaff, A., 2022. A comparative overview of hybrid recommender systems: Review, challenges, and prospects. Data Mining and Machine Learning Applications , 57–98.

SINTEF, 2013. Big data, for better or worse: 90percent of world's data generated over last two years. ScienceDaily .

Steck, H., Baltrunas, L., Elahi, E., Liang, D., Raimond, Y., Basilico, J., 2021. Deep learning for recommender systems: A netflix case study. AI Magazine 42, 7–18.

Thorat, P.B., Goudar, R.M., Barve, S., 2015. Survey on collaborative filtering, content-based filtering and hybrid recommendation system. International Journal of Computer Applications 110, 31–36.

Truong, Q.T., Salah, A., Lauw, H., 2021. Multi-modal recommender systems: Hands-on exploration , 834–837.