



DIOGO MIGUEL MARTINS ANTÃO

Bachelor of Science in Biomedical Engineering

PROMOTION OF ACTIVE AGING THROUGH A RECOMMENDATION SYSTEM BASED ON MULTIMEDIA CONTENT

MASTER IN BIOMEDICAL ENGINEERING

NOVA University Lisbon

March, 2023

PROMOTION OF ACTIVE AGING THROUGH A RECOMMENDATION SYSTEM BASED ON MULTIMEDIA CONTENT

DIOGO MIGUEL MARTINS ANTÃO

Bachelor of Science in Biomedical Engineering

Adviser: Prof. Hugo Filipe Silveira Gamboa
Associate Professor, NOVA University Lisbon

Co-adviser: Dr. David José da Silva Aresta Belo
Senior Scientist, Associação Fraunhofer Portugal-AICOS

Examination Committee

Chair: Susana Isabel dos Santos Silva Sérgio Venceslau
Associate Professor, FCT-NOVA

Adviser: Prof. Hugo Filipe Silveira Gamboa
Associate Professor, NOVA University Lisbon

Co-adviser: Dr. David José da Silva Aresta Belo
Senior Scientist, Associação Fraunhofer Portugal-AICOS

Member: Prof. Cláudia Regina Pereira Quaresma
Associate Professor, NOVA University Lisbon

Promotion of active aging through a recommendation system based on multi-media content

Copyright © Diogo Miguel Martins Antão, NOVA School of Science and Technology, NOVA University Lisbon.

The NOVA School of Science and Technology and the NOVA University Lisbon have the right, perpetual and without geographical boundaries, to file and publish this dissertation through printed copies reproduced on paper or on digital form, or by any other means known or that may be invented, and to disseminate through scientific repositories and admit its copying and distribution for non-commercial, educational or research purposes, as long as credit is given to the author and editor.

ACKNOWLEDGEMENTS

These last few years in which I have been studying Biomedical Engineering been filled with sweat and tears but also laughter and personal realization. I never once regretted my choice, even in the most difficult moments.

Firstly, I would like to thank my adviser Professor Hugo Gamboa for the opportunity to develop this project, his patience for my shortcomings, his advice and his contribution towards my education even outside of this project.

I am also very thankful for my co-adviser, David Belo, whose motivation, patience and support were essential to the development of this dissertation. In the same vein, I would like to thank Philip Probst from LIBPhys, which also supplied key advice. I would like to also thank Cátia Cepeda, also from LIBPhys for accompanying me in early stages of the project.

Finally I would like to thank my family and friends for their support and motivation. To my father and mother for their patience when tolerating my mood and specially to my cat for always giving me company, even with nobody else could.

*“Deus ao mar o perigo e o abismo deu,
Mas nele é que espelhou o céu” (Fernando Pessoa)*

ABSTRACT

Due to the increase in life expectancy, promotion of active aging has become a raising concern for human society. Machine Learning applications allow for dynamic and personalized solutions to support the chronic and complex healthcare challenges for elderly people. In particular, recommendation systems in the healthcare domain have shown positive results in the promotion of well being with non-intrusive methods. Considering how aging populations are some of the biggest consumers of television, there is an opportunity for recommendation systems specialized on that type of media to be used in the promotion of active aging. But existing systems in this context lack the ability to detect elderly users, which limits their usage to predetermined groups.

This dissertation investigates the creation of an explainable recommendation system for television contents that can be used in the promotion of active aging. It also presents a method to detect older users from a dataset pertaining to television usage. The recommendation system was developed using both content-based and collaborative techniques, implemented with [K-Nearest Neighbors \(KNN\)](#) and [Singular Value Decomposition \(SVD\)](#) algorithms as well as cosine similarity. Explanations were proposed utilizing post-hoc and model-agnostic methods based on item and user similarity and evaluated with [Mean Explainability Precision \(MEP\)](#). The identification of elderly users was conducted with a clustering approach featuring [Principal Component Analysis \(PCA\)](#) and [t-Distributed Stochastic Neighbor Embedding \(t-SNE\)](#). Each of the explanation style that were used reflected a [MEP](#) value above 0.5 for both algorithms. The clustering from [t-SNE](#) allowed the identification of which division of the dataset was most likely to feature elderly users when compared to available statistics. These results reflect potential in application of the proposed system to an active aging context.

Keywords: Machine Learning, Recommendation Systems, Active Aging, Explainable Artificial Intelligence, Post-hoc Explanations

RESUMO

Devido ao aumento da esperança média de vida, a promoção de envelhecimento ativo tem-se tornado uma preocupação crescente na sociedade humana. Algoritmos de aprendizagem automática permitem o desenvolvimento de soluções dinâmicas e personalizadas para o apoio dos desafios de saúde apresentados por pessoas idosas. Em destaque, sistemas de recomendação aplicados ao domínio da Saúde têm mostrado resultados positivos na promoção de bem-estar utilizando métodos não-intrusivos. Considerando como as populações envelhecidas são dos maiores consumidores de televisão, existe uma oportunidade para sistemas de recomendação especializados nesse tipo de media serem utilizados na promoção de envelhecimento ativo. No entanto, os sistemas existentes aplicáveis a este contexto não possuem a capacidade de detetar utilizadores idosos, o que limita a sua utilização a grupos predeterminados.

Esta dissertação investiga a criação de um sistema de recomendação de conteúdos televisivos explicável que possa ser usado na promoção do envelhecimento ativo. Apresenta também um método para detetar utilizadores idosos de entre um conjunto de dados sobre visualizações de programas televisivos. O sistema de recomendação foi desenvolvido utilizando técnicas de filtragem colaborativa e baseadas no conteúdo, implementadas com algoritmos de [KNN](#) e [SVD](#), juntamente com semelhança de cosseno. Explicações foram propostas usando métodos post-hoc e de natureza agnóstica em relação aos algoritmos escolhidos, baseadas em semelhanças entre utilizadores e itens e avaliadas com [MEP](#). A identificação de utilizadores idosos foi realizada com métodos de agrupamento de dados utilizando [PCA](#) e [t-SNE](#). Cada estilo de explicação foi usado obteve um [MEP](#) superior a 0.5 para ambos os algoritmos. O agrupamento que recorreu a [t-SNE](#) permitiu distinguir em qual o grupo de utilizadores é mais provável existirem idosos através de comparações às estatísticas disponíveis. Estes resultados refletem o potencial na aplicação do sistema proposto ao contexto do envelhecimento ativo.

Palavras-chave: Aprendizagem Automática, Sistemas de Recomendação, Envelhecimento Ativo, Inteligência Artificial Explicável, Explicações Post-hoc

CONTENTS

List of Figures	x
List of Tables	xi
Abbreviations	xiii
1 Introduction	1
1.1 Context and Motivation	1
1.2 Objectives	2
1.3 Thesis Structure	2
2 Theoretical Background	3
2.1 Machine Learning	3
2.1.1 K-Nearest Neighbors	4
2.1.2 Singular Value Decomposition	5
2.2 Dimensionality Reduction	6
2.2.1 Principal Component Analysis	6
2.2.2 t-Distributed Stochastic Neighbor Embedding	7
2.3 Clustering	7
2.3.1 Silhouette Method	8
2.4 Recommendation Systems	8
2.5 Recommendation Explanations	9
2.6 Machine Learning Metrics	11
2.6.1 Precision and Recall	11
2.6.2 Root-Mean-Squared Error	12
2.6.3 Cosine Similarity	12
2.6.4 Mean Explainability Precision	12
3 Literature Review	14
3.1 Active Aging	14

3.2	Recommendation Systems in Healthcare	15
3.3	Recommendation Systems Explainability	16
4	Methods	18
4.1	Dataset	18
4.2	Computational Methods	18
4.3	Processing of the Dataset	19
4.3.1	Exploratory Data Analysis	20
4.4	Proposed Recommendation System	20
4.4.1	Content-Based Filtering	21
4.4.2	Collaborative Filtering	21
4.5	Recommendation Explanations	21
4.6	Identification of Elderly Users	22
4.6.1	Clustering	24
5	Results and Discussion	26
5.1	Exploratory Data Analysis	26
5.2	Recommendation System Prototype	28
5.2.1	Generated Explanations	28
5.3	Clustering	30
5.3.1	Principal Component Analysis	30
5.3.2	t-Distributed Stochastic Neighbor Embedding	31
5.3.3	Silhouette Coefficient	32
5.4	Identification of Elderly Users	33
6	Conclusion and Future Work	36
6.1	Conclusion	36
6.2	Future Work	37
	Bibliography	38
	Appendices	
	Annexes	

LIST OF FIGURES

2.1	SHAP explanation, adapted from [65].	10
2.2	Visual representation of set theory. Adapted from [68].	11
2.3	Visual representation of cosine similarity, adapted from [28].	12
3.1	Histogram with ratings from similar users, the best performing explanation method in [82].	16
4.1	Television viewership statistics for the elderly in Portugal, adapted from [93].	23
5.1	Examples of histograms obtained for the television genres "biography" 5.1(a) and "documentary" 5.1(b).	26
5.2	Histogram with ratings for the recommended item, from the ten most similar users.	29
5.3	Scatter-plot for the PCA dimension reduction.	31
5.4	Scatter-plots of t-SNE dimensions for different perplexity values.	32
5.5	Silhouette Coefficient for different numbers of clusters.	33
5.6	Plot of the clusters.	33
5.7	Mean values for each cluster.	34

LIST OF TABLES

2.1	Commonly used distance metrics, to calculate distance between vectors x_s and y_t belonging to matrices X and Y respectively.	5
4.1	Features of the processed dataset.	19
4.2	Metrics displayed in the Marktest data.	24
4.3	Conversion of television genres into the content typologies studied by Marktest.	24
5.1	Statistics for the viewing time of each television genre in the dataset, in minutes.	27
5.2	Performance and accuracy of each model.	28
5.3	Evaluation of explanations.	30
5.4	Statistics for the clustering features.	30
5.5	Evaluation of Principal Components.	31
5.6	Statistics for the clustering features.	35

ABBREVIATIONS

AI	Artificial Intelligence (<i>pp.</i> 3, 8, 9)
KNN	K-Nearest Neighbors (<i>pp.</i> vi, vii, 4, 9, 21, 28, 36)
MEP	Mean Explainability Precision (<i>pp.</i> vi, vii, 12, 13, 22, 29, 30, 36)
ML	Machine Learning (<i>pp.</i> 1–5, 8, 11, 12, 15, 16, 18, 21, 36, 37)
PCA	Principal Component Analysis (<i>pp.</i> vi, vii, x, 4, 6, 24, 25, 30, 31, 36, 37)
RMSE	Root-Mean-Square Error (<i>p.</i> 12)
SVD	Singular Value Decomposition (<i>pp.</i> vi, vii, 4, 5, 10, 21, 28, 30, 36)
t-SNE	t-Distributed Stochastic Neighbor Embedding (<i>pp.</i> vi, vii, 4, 7, 24, 25, 31–33, 36, 37)
WHO	World Health Organization (<i>p.</i> 1)

INTRODUCTION

1.1 Context and Motivation

According to the [World Health Organization \(WHO\)](#), life expectancy has increased substantially over the past two centuries, specially in developed countries [2]. More people today can expect to live into their 60s and beyond than ever before in human history. In Europe, it is estimated that up to 34% of all Europeans will be 60 or older by the year 2050 [3].

Unhealthy aging is associated with a high risks of metabolic and cardiovascular diseases [4], chronic pain [5] and severe cognitive decline [6]. This as motivated a number of initiatives to foster healthy habits among the population which are often supported by biomedical applications [7]. To clarify the objectives of such interventions, [WHO](#) as defined Active aging as "the process of optimizing opportunities for health, participation and security in order to enhance quality of life as people age"[8]. The word "active" doesn't merely refer to physical activity, but also the ability to maintain participation in civic, social and economic activities [8].

The complex and chronic nature of health challenges for the elderly has influenced the development of healthcare applications employing [Machine Learning \(ML\)](#) algorithms [7, 9]. These allow for personalized and adaptable solutions as well as dynamic condition-based assessments. [ML](#) as been successfully used to monitor and improve the diets and lifestyle of elderly people [10], in the detection of chronic ailments [11] and in the prediction of anxiety and depression [12].

Content-based recommendation systems are [ML](#) applications which typically allow the user to find multimedia more suitable for their interests or searching goals by presenting personalized suggestions. Although these are typically meant to combat information overload, they also have their use in health monitoring, particularly because they allow the dynamic evaluation of the user in a non-intrusive manner and can be used to enhance their well being [13]. Recommendation systems usually produce explanations for their suggestions which are particularly relevant in the healthcare domain, where they can increase the user's motivation to follow lifestyle recommendations and increase clinicians

trust in the system [14].

There is precedent in the use of information and communication technologies in the promotion of active and healthy aging [15] and older populations are some of the biggest consumers of television [16]. However, existing systems do not allow for the detection of elderly individuals from among the users, limiting their scope to pre-selected groups. This dissertation, which was developed in collaboration with NOS Inovação® aims to study the creation of a recommendation system capable of identifying elderly users and promoting active aging, based on data from television viewership.

1.2 Objectives

The main objective of this dissertation is the study of multimedia content recommendations in the context of active aging and the promotion of emotional well-being, using ML algorithms. It also comprises the suggestion of explanations that may justify the ML-based recommendations to the user.

To reach these objectives, several steps were completed during the development of this dissertation:

- Processing of television usage data provided by NOS.
- Development of a traditional recommendation system for media content, using ML algorithms, which can later be adapted to promote active aging.
- Study of possible explanations for the generated recommendations.
- Identification of elderly users from among the dataset through clustering techniques.

1.3 Thesis Structure

This thesis is organized into six chapters, which are arranged as follows: The covered chapter introduced the context and motivation for this dissertation, alongside it's objectives. The second covers the theoretical concepts necessary to understand recommendations systems, such as the relevant ML techniques for their implementation and evaluation approaches, the dimensionality reduction techniques used to in the clustering process and the metrics used. The third comprises the literature review, covering recommendation systems in healthcare, explanation techniques for recommendations and approaches for active aging promotion. The forth chapter presents a detailed explanation of the methods utilized in order to fulfill the intended objectives. The fifth chapter covers the outcomes, including results and their discussion. Finally, the sixth chapter addresses the conclusion alongside some suggestions for future work.

THEORETICAL BACKGROUND

This Chapter contemplates a broad overview of the theoretical concepts relevant to this dissertation. Firstly, it will cover relevant Machine Learning (ML) methods in the context of this dissertation, including dimensionality reduction techniques and clustering. Then, it will present the concept of recommendation systems and their explanations. Finally, it features some metrics which are relevant in the evaluation of ML classifications and recommendation applications present in this project.

2.1 Machine Learning

Machine Learning (ML) is a sub-field of Artificial Intelligence (AI) comprised of algorithms which attempt to identify patterns within data samples and associate them into discrete classes [17]. ML techniques are based on the development of learning algorithms which in turn build models from an initial dataset that can be used to make predictions on new data [18].

Because ML algorithms do not require explicit programming, they are more dynamic when compared to hand-coded solutions, as the logic utilised to make decisions isn't specific to a single domain and require less intimate knowledge of the task [19].

Typically, ML approaches are classified into three categories: supervised, unsupervised and reinforced learning [20]:

- Supervised ML, in which the user provides pairs of input and output, allowing the algorithm to learn how to make predictions for unknown outputs given new inputs [19].
- Unsupervised ML, in which only the input data is known and the output must be defined with the application of the algorithm [19]. Its goal is the identification of underlying characteristics in the data structure [21].
- Reinforced learning approaches are somewhere between the other two, requiring some supervision without specifying an output [20].

Supervised ML approaches are commonly used to solve description problems in data science, in which the goal is to predict or classify an outcome of interest [21]. These techniques can be further categorized into two families, classification or regression [22]. In the former, predictions are made about categorical variables (e.g. identifying health issues in patients), while in the latter predicts numerical variables like the cost of a treatment [17].

Unsupervised ML techniques can also be divided into two families, clustering and dimension reduction. Clustering aims at dividing data into distinct classes, usually relying on distance between them and assuming that data points in the same cluster will have similar properties [17]. Dimension reduction, also known as ordination, aims to reduce the dimensionality of the data, either to compress it, or to improve its visualization, these include PCA and t-SNE techniques [19].

ML models are also classified according to their interpretability. [23] defines this concept as an algorithms ability to be "understood by a human, either by introspection or through a produced explanation". These include linear and logic regressions, decision trees, Naive Bayes and KNN [24]. Non-interpretable models, which are often called black-box models, include neural networks, deep neural Networks, support vector machines, tree ensemble and non-linear models [25]. These algorithms usually have better accuracy when compared with interpretable ones [26].

For this dissertation, both an interpretable model and a non-interpretable one were selected in order to compare their general performance and conduct a more thorough examination of the systems explainability. The chosen algorithms were the KNN and SVD respectively, which are commonly used in media recommendations [27, 28].

2.1.1 K-Nearest Neighbors

KNN algorithms are based on the assumption that different items in a database are similar to each other if they are close in proximity. Using distance metrics, these algorithms attempt to perform classifications by associating samples into clusters [17, 29].

The quality of estimation is dependant on the neighborhood size k , as the radius of the local region around each item is determined by its distance to the k -th nearest neighbor. If k is too small, the local estimation will be affected by data sparseness and ambiguity. If its too large, the classification will suffer from over-smoothing, meaning it will ignore the more subtle patterns and may be degraded by outliers [30].

Some of the metrics typically used in KNN are displayed on Table 2.1, as they are described in [31].

KNN is considered a memory-based classification model as it requires all available instances of a task to be stored before it can make extrapolations about similarities in the data [32]. These types of model are characterized by their interpretability, simplicity and easy implementation [33].

Table 2.1: Commonly used distance metrics, to calculate distance between vectors x_s and y_t belonging to matrices X and Y respectively.

Distance	Metric
Minkowski	$d_{s_t} = \sqrt[p]{\sum_{j=1}^n x_{s_j} - y_{t_j} ^p}$
Euclidean	$d_{s_t}^2 = (x_s - y_t)(x_s - y_t)'$
City Block	$d_{s_t} = \sum_{j=1}^n x_{s_j} - y_{t_j} $
Cosine	$d_{s_t} = 1 - \frac{x_s y_t'}{\sqrt{(x_s x_s')(y_t y_t')}}}$

2.1.2 Singular Value Decomposition

In contrast with the memory-based ML techniques, model-based techniques aim to use the training data to construct a model of scoring rules. Model-based solutions demand a more complex implementation, but are preferable in the case of high data sparsity. A notable approach in the context of recommendation systems relies on Singular Value Decomposition SVD, a matrix factorization technique [34].

SVD is based on linear algebra theorems which allow any rectangular matrix to be broken down into the product of three matrices: an orthogonal matrix U , a diagonal matrix S and a transposed orthogonal matrix V [35]. A popular application in this context is the matrix completion problem in which given a data matrix with S rows and n columns, it can extract uncorrelated low-dimension factors. The original matrix can then be recreated from these factors [36]. Given a $m \times n$ size matrix A , it's SVD can be presented as:

$$SVD(A) = USV^T \quad (2.1)$$

In which U , S and V are of dimensions $m \times m$, $m \times n$ and $n \times n$ respectfully. These matrices represent the breakdown of the original relationships into independent components or factors. An important property of SVD is the possibility of ignoring some of the entries in the diagonal matrix S , allowing for a lower rank approximation of the original matrix A . That is, with a k number of non-zero entries in S , the dimensions of U , S and V become $m \times k$, $k \times k$ and $n \times k$ respectfully. The resulting matrix from the multiplication of these three is a rank k approximation of A . This process allows for the smoothing of the entries in the original matrix, resulting in a reduction of dimensionality [36, 37].

Recommendation systems are a prominent application of SVD, aiming to find a model of relationships between users and items to recommend by identifying latent factors. A user's rating of a particular item is predicted based on these factors. Due to it's properties as a diagonal matrix, S simply acts as a scalar on U and V^T . We can assume the approximation as merged scalar factor into U and V , so 2.1 becomes $A = UV^T$. If we consider the predicted rating of item i by user u to be $x_u i$, this will be the result of corresponding vectors p_u and q_i which are a row of U and a column of V^T respectfully [36]. The elimination of S results in the loss of biases, which are systematic tendencies of certain users to rate items more highly and for certain items to be more highly rated. In

order to avoid this information loss, [27] proposes that these biases should be added back into the model as a linear combination, resulting in:

$$x_{ui} = p_u \times q_i^T + b_{u_u} + b_{i_i} \quad (2.2)$$

In which b_{u_u} represents the singular value of user bias for u and b_{i_i} represents the singular value of item bias for i .

2.2 Dimensionality Reduction

In a dataset with high dimensionality, the detection of understandable patterns and their classification is more challenging, as many of the features included may be redundant or irrelevant [38]. On the other hand, many scientific applications, like those involving exploratory data analysis, rely of graphical visualisation to identify qualitative data structures, such as clusters [39].

Dimensionality reduction techniques aim to reduce the number of input variables by exploiting redundancy in the data. Their implementation allows the creation of a smaller set of new variables, each being a combination of the ones featured in the input, while keeping as much of the original information as possible [40].

2.2.1 Principal Component Analysis

PCA is the most widely used dimensionality reduction technique. It is a orthogonal statistical method whose main aim is to find the optimal position for the best information variance and vector dimensional features reduction. This is done by transforming a set of observations of linearly related variables into a set of non-linearly related values. [41].

More specifically, dimensionality reduction is accomplished by identifying directions, referred to as the namesake principal components, along which the variation of data is maximal [42]. Given a data matrix with p variables, the first principal component Y_1 , the one with the greatest variance, is obtained with:

$$Y_1 = a_{1_1} X_1 + a_{1_2} X_2 + \dots a_{1_p} X_p \quad (2.3)$$

In which $X_1, X_2, \dots, X_p$ are the variables, and $a_{1_1}, a_{1_2}, \dots, a_{1_p}$ are weights that can be used to make the variance as large as possible by selecting large values for each of them, with the constrain that the sum of their squares is equal to 1. The second principal component is calculated the same way, with the condition that it is uncorrelated with the first one. This process can continue until up to p principal components are determined, with each component "explaining" or accounting for some of the total variance [43].

Samples can then be plotted using the principal components as axis in order to visually evaluate if or how they can be grouped [42].

2.2.2 t-Distributed Stochastic Neighbor Embedding

t-SNE is another technique which allows visualization of high-dimensionality data by reducing it such that each sample can be described by a location in 2D or 3D spaces. This approach minimizes the Kullback-Leibler divergence between two probabilistic distributions: p_{ij} which measures pairwise similarities in the high-dimensionality input space and q_{ij} that represents pairwise similarities in the corresponding embedded low-dimensionality space. The cost function C is given by [44, 45]:

$$C = \sum_i KL(P_i || Q_i) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \quad (2.4)$$

In which P_i represents the conditional probability distribution over all other points in the data given a high dimensionality data point and Q_i represents the conditional probability distribution over all other points in the data given a low dimensionality data point. This process allows the preservation of the local structure of the data [44]. In applications for data visualisation, this technique may feature a perplexity parameter used to balance the importance local and global structures: Higher perplexity translates to a larger number of points included in the t-distribution when attempting to find neighbors, resulting in a less localized preservation of the original structure [45].

2.3 Clustering

Clustering is one of the most commonly utilized approaches to unsupervised learning and data mining. This technique is based on the division of the data into various clusters or classes, which may represent relevant similarities or distinctions between their objects [46]. Clustering algorithms are classified according to the methods used in their implementation:

- Partition methods, in which the data is split into a predetermined number of clusters, formed according to distances, like those in 2.1.1, or pattern correlations [47]. The most famous of these is K-Means, an algorithm that uses iterative computation to update the center of a cluster until the convergence criteria are met [48].
- Hierarchical methods generate an hierarchical decomposition of the data. Each object starts by forming an independent class, which is then combined with similar classes until all are combined into one or a specific condition is fulfilled. This can be achieved with a binary merge tree [49].
- Density-based methods denote non-parametric approaches where the clusters are the high density areas in the data, separated by low density regions between them. These algorithms do not require a predetermined number of clusters as input and do not make assumptions about the variance within each cluster [50].

- Grid based methods translate the data into a grid structure composed of a defined number of cells. This approach allows for parallel processing and the application of other techniques, such as density-based or hierarchical methods in the grid structure [48].

The clustering algorithm used in this dissertation is K-means, which as partition method is characterized by it's simple implementation and low processing time [47]. It's general process starts by selecting K initial points as centroids, forming a cluster around each centroid by assigning data points to the closest one and recompute the centroid for each cluster until the result does not change [46].

2.3.1 Silhouette Method

Clusters can be validated with internal measures that quantify certain characteristics of their structure. These include separation, which determines how distant different clusters are from each other, and cohesion, which measures the distance of data objects in the same cluster [51].

Cohesion is assessed by the sum of squares within cluster (SSE) which is calculated with equation 2.5, while separation is determined by the sum of squares between cluster (SSB), described in equation 2.6 [52].

$$SSE = \sum_i \sum_{x \in C_i} (x - m_i)^2 \quad (2.5)$$

$$SSB = \sum_i |C_i| (m - m_i)^2 \quad (2.6)$$

Since this project employs a partition method for clustering, an appropriate number of clusters must be chosen as input for the algorithm. A popular approach to evaluate this variable is the silhouette method. This technique uses a coefficient based on both cohesion and separation, which is computed with [53]:

$$s = \frac{SSB - SSE}{\max(SSE, SSB)} \quad (2.7)$$

This measure ranges from -1 to 1. A positive value indicates a good separation between clusters, while a negative value suggests the presence of many outliers and poor clustering [53].

2.4 Recommendation Systems

Recommendation systems are AI applications which generate suggestions tailored to user needs or preferences. This is typically achieved by analyzing users and items as input, finding hidden relationships between them with the use of ML techniques [54–56].

These are usually classified according to the type of filtering used, that is, the technique utilized to find relations between users and items that can result in a recommendation:

- Content-Based filtering focuses on item profiling, it attempts to recommend to the user items which are similar to those they have selected or rated positively in the past [54].
- Collaborative filtering relies on a user-item rating matrix, generating recommendations based on the ratings of similar users. [54, 57].
- Knowledge-based filtering approaches rely on information not usually included in the previous techniques such as explicit knowledge about an item or user preference. [55, 58]
- Hybrid filtering is obtained by utilizing multiple of the previous approaches in the same system [55].

Collaborative filtering is the most widely utilized and successful recommendation technology, it's traditional aim being to recommend items to a user based on the behaviour of similar users. This approach is however requires a substantial profile of each user and item to make reliable predictions. This limitation is known as the "cold start" problem[55, 59]. It's also susceptible to issues caused by data sparsity, making the similarities between users more challenging to correctly identify [60].

Content-Based filtering on the other hand is not as affected by cold start, since it directly relies on item characteristics and may require only a small number of ratings to create a user profile. This method is however, susceptible to scalability issues, as it requires the characterization of each new item added to the database. Another limitation is the technique's dependence on previously established item attributes, which may be potentially incorrect or incomplete [55].

Hybrid filtering attempts to solve the limitations of the previously mentioned techniques by combining them. They may also feature a knowledge filtering component, in the form of contextual information, which improves the effectiveness of the resulting recommendations [54].

2.5 Recommendation Explanations

In the context of research and development for recommendation systems, there is an increasing awareness of the need for explanations that improve the interpretability of such systems and user's confidence in their results [61]. In the healthcare domain, the biggest obstacle for the implementation of AI tools in general is the lack of transparency. Since clinicians have a responsibility to give the best care possible to their patients, they require an understanding of how the model operates [14].

For recommendation systems featuring interpretable algorithms, like KNN, typically employ the explainable modelling approach, making use of their inherent explainability to develop task models which are easily understandable by the user. These are however

susceptible to large numbers of objects in the data and the comprehensibility of input features, which negatively affect explainability [14].

Recommenders based on black-box algorithms often attempt to generate explanations with the help of interpretable surrogate systems with an approximated behaviour to that of the non-interpretable algorithm. This is called a post-hoc approach, which does not describe the actual operations of the recommendation system, but instead seek to justify them with approximations. In the case of matrix factorization algorithms, like **SVD**, this can be achieved with model-specific techniques that attempt to find the latent factors used by the model [62].

Post-hoc explanation techniques that can be applied to any algorithm are classified as model-agnostic [14]. These include methods like Local Interpretable Model-Agnostic Explanations (LIME) and Shapley Additive Explanations (SHAP). LIME seeks to explain the predictions for each particular input by sampling its neighboring inputs and learning a sparse linear model based on the predictions for those neighbors [63]. SHAP determines the influence of each feature by repeatedly computing a prediction result given different subsets of features, presenting explanations like the one in Figure 2.1 [64, 65].

The following films you have rated influenced the recommendation the most:



Figure 2.1: SHAP explanation, adapted from [65].

Model-agnostic approaches also include simple explanation methods based on user and item similarities. These include the usage of neighbor ratings for the recommended item or explanations based on common features between items appreciated by the user and the suggested ones. Although such methods present weaker approximation to the algorithm, they're easier to implement and more versatile [66].

The importance of explanations for recommendation systems is usually divided into seven characteristics, which are meant to reflect their impact on the users [67]:

- **Transparency** refers to the capability of describing how the recommendation was generated.
- **Scrutability** reflects whether users can tell when the system is wrong.
- **Trust** describes the user's confidence in the system.
- **Effectiveness** is determined by the ability to help users make good decisions.

- **Persuasiveness** denotes how the explanation incentivises the user to buy products in a context of commercial recommendations.
- **Efficiency** describes whether the explanation decreases the time required by users to make decisions.
- **Satisfaction** reflects whether the explanation increases user's satisfaction with the overall system.

2.6 Machine Learning Metrics

This section introduces relevant metrics used in development this dissertation to evaluate the performance and accuracy of the employed ML algorithms and to compute similarities between objects in the data for the recommendation systems.

2.6.1 Precision and Recall

Precision and recall are the most used metrics in the assessment of performance for information retrieval or pattern recognition applications. Precision is defined as the number of correct results relative to the number of all results while recall is determined by the number of correct results relative to the number of expected results [68, 69]. They are calculated using concepts from set theory, which can be visualized in Figure 2.2.

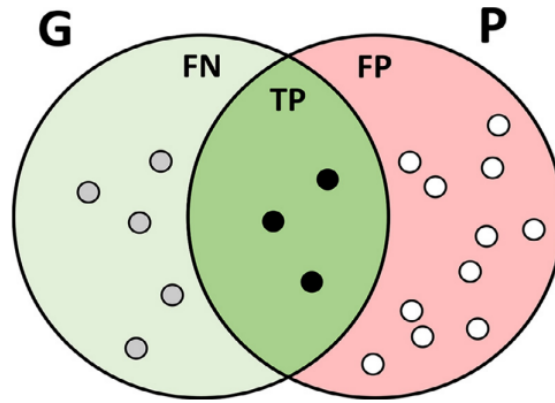


Figure 2.2: Visual representation of set theory. Adapted from [68].

Ground truth (G) denotes the truth classifications while the predicted result (P) comprise the prediction made by ML algorithms. The interception presents the correct predictions, although usually only the true positives (TP) are relevant and true negatives are discarded [68].

Considering the concepts from Figure 2.2, precision and recall can be calculated with equations 2.8 and 2.9 respectively.

$$Precision = \frac{|G \cap P|}{|P|} \quad (2.8)$$

$$Recall = \frac{|G \cap P|}{|G|} \quad (2.9)$$

2.6.2 Root-Mean-Squared Error

Root-Mean-Square Error (RMSE) is a standard statistical metric of accuracy for ML applications. Considering n observations of y and the n corresponding predictions $\hat{y}$, the value of **RMSE** is computed by [70]:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2} \quad (2.10)$$

Equation 2.10 results in a mean squared value for the residual (difference between the prediction and the true value). This difference is expressed in the same units as the observations, allowing for a direct interpretation of the mean error.

2.6.3 Cosine Similarity

Cosine similarity is metric used to compute the proximity of objects in a dataset. From a visual perspective, it can be understood as a measurement of the angle of cosine between two vectors as represented in Figure 2.3, a lesser angle denotes greater similarity [28].

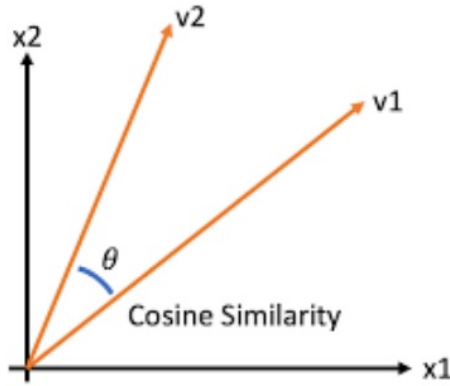


Figure 2.3: Visual representation of cosine similarity, adapted from [28].

It is obtained by dividing the scalar product of the vectors by the product of their lengths as can be seen in equation 2.11, allowing for a similarity measurement on a normalized scale.

$$CosSim(x, y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2} \sqrt{\sum_i y_i^2}} \quad (2.11)$$

2.6.4 Mean Explainability Precision

This dissertation does not cover user evaluations for the developed explanations. However, their performance can be assessed using the concept of **MEP**, introduced by [71]. This

approach requires a threshold to be defined for each style of explanation. For example, an explanation based on the rating of similar users for the recommended item may take as threshold the presence of a minimum number of neighbors who've rated it. Only recommendations that fulfill this condition are considered explainable.

To compute the [MEP](#), the number of explainable recommendations generated for each user is divided by the number of all recommendations for that user and then the users mean is calculated [\[72\]](#), as in :

$$MEP = \frac{1}{|U|} * \sum_{u \in U} \frac{|R^{rec} \cap E|}{|R|} \quad (2.12)$$

In which U is the set of users, R^{rec} is the set of recommendations for each user, R represents all the recommendations and E is the set of all the explainable recommendations.

LITERATURE REVIEW

This Chapter contains the literature review, initially pertaining to the promotion of active aging with a focus on information and communication technology approaches. It will also cover recommendation systems have been used in the context of healthcare as well as current approaches to explanations for their recommendations.

3.1 Active Aging

The topic of active aging raises a few challenges such as how to convince people to follow healthy behavior advice and how to promote social inclusion for elderly populations [73]. The following studies present examples of media-based applications for the promotion of healthy aging.

In 2015, Chessa et al. developed a home care service based on information and communication technologies targeting elderly patients. This system proposes active aging lifestyle protocols, which receive constant feedback from a set of mobile applications. These include social games, cognitive games and diet applications which also motivate the users to continue following the protocol. Although in this case the supervision was complemented with wearable and environmental sensors, it exemplifies the adaptation of active aging solutions into the target's environment through the use of media [74].

Also in 2015, Caprara et al. conducted a study in order to evaluate the effectiveness of a psycho-educational multimedia program designed to promote healthy aging. This program, "Vital Aging-Multimedia" was implemented over 3 months through 35 hours of video lessons and addressed four components of healthy aging: cognitive functioning, aging self-efficacy, social participation and general health habits. A total of 115 senior citizens aged 52-82 participated in the study, of which 73 were exposed to the program, while 42 were used as control. The participants were assessed before and after being exposed to the materials with a battery of psychological instruments and were found to score higher than the control group and reported better health, attesting to the positive effects of the program [75].

In an article from 2019, Vázquez et al. propose the development of a cognitive-behavioral intervention for active aging based on an interactive multimedia online game. This intervention, which consists of 8 modules of approximately 45 minutes each, would be administered through a mobile application at the rate of one per week. The game is a graphic adventure, in which the main character is introduced to a variety of techniques that promote well-being and increase quality of life. The overall intervention focuses on prevention of depression, cognitive stimulation and promotion of healthy lifestyle habits. This approach counts anonymity, the ability for the participant to choose the time of application and accessibility as advantages when compared to traditional therapy [76].

3.2 Recommendation Systems in Healthcare

Recommendation systems are utilized in a number of fields, such as e-learning and e-commerce, usually attempting to suggest the items which will better gauge the users attention. However, they have also become a popular application of artificial intelligence in the healthcare context. Utilizing a variety of ML techniques, it can solve difficult issues featuring unstructured information [77].

A study from 2017 comparing the quality of recommendations in the context of online health forums found hybrid methods to be the most successful [78]. An article from the same year comparing the recommendation system in a healthcare focused social network which aimed at suggesting specific medical professionals and facilities for each patient, concluded that pre-filtering techniques for similar patients greatly increased recommendation accuracy [79].

In 2018, researchers applied recommendation systems in to the EU-funded AGILE project. Their purpose was to build a "modular hardware and software gateway" for a networked infrastructure of interconnected devices, which would allow mobile tracking of biomedical data. They proposed two recommendation systems in this context: one which recommends activity plans and devices(such as step-count watches and wristbands), based on demographic data and user similarities, and another system which motivates chronic patients to reach the goals of a recommended treatment plan, using a content-based filtering approach [80].

A systematic review was conducted in 2021 focusing on recommendation systems targeting non-medical staff. This study divided the featured recommendation systems into 4 categories. Of the 73 unique papers analysed, 24 dealt with lifestyle recommendations, 26 dealt with nutrition, 23 were about information and 9 focused on specify health conditions. The authors noted a majority presence of knowledge-based approaches. While only 13 of the papers featured exclusively knowledge-based systems, 32 included them in hybrid systems, resulting in a total of 49 out of 73. According to this review, the most popular offline evaluation metrics include precision, accuracy, performance and recall [58].

3.3 Recommendation Systems Explainability

Recommendation systems often rely on black-box ML models, which provide no explanatory information. When that is the case, the recommendation should be accompanied of an explanation that describes why specific items were recommended [81].

In 2000, Herlocker et al. studied explanations for automated collaborative filtering systems. They evaluated 21 different explanation styles in a movie recommendation context, with two experiments. In the first one, users preferences for each explanation method was measured, resulting in histograms of neighbors rating, as depicted in Figure 3.1, past performance and similarity to items in the user’s profile being the highest rated. The second experiment measured the response to adding an explanation interface to their recommendation system, which wielded positive results [82].

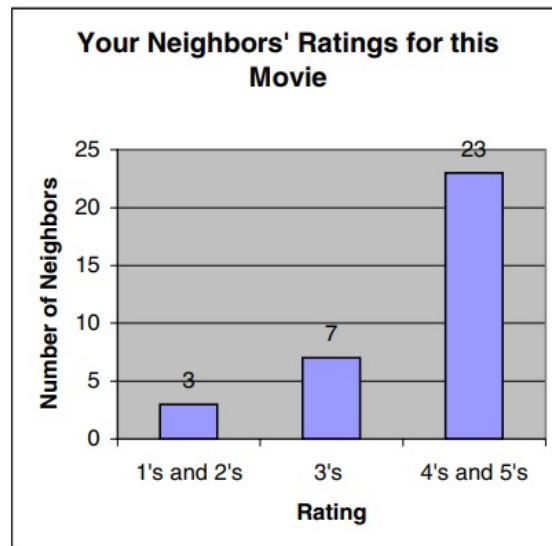


Figure 3.1: Histogram with ratings from similar users, the best performing explanation method in [82].

A paper from 2012 empirically studied the effectiveness of feature-based explanations in recommendation system for films and cameras. The movie features covered in the study included cast, director and genre. These were used to generate textual explanations that were then evaluated by users. By comparing effectiveness and satisfaction, the investigators found that even when feature-based explanations do not help the user make better decisions, they still increase their satisfaction with the system [83].

In 2020, Shmaryahu et al. proposed a post-hoc approach for explaining black-box recommendation systems. Instead of trying to extract latent factors from matrix factorization applications, this approach is based on attempting to find justifications for the recommendation that still provide value for the user despite not shedding light on the model’s inner mechanism. After the recommendation is realized by the black-box model, the recommended item is used as input by the explanation system, which in turn produces a score for each possible explanation and picks the most highly rated. These may include

simply item popularity, similarity between recommended item and items chosen by the user in the past, similarity between the user and users who've picked said item, etc. This explanation method was tested in a movie recommendation context with 207 participants, in which most of the explanation types were positively received. [66]

METHODS

This Chapter presents an overview of the methods used to achieve the proposed objectives. It starts with an introduction of the dataset. Then, it includes a description of the computational methods used, followed by a presentation of the proposed recommendation system and explanation suggestions. Lastly, it explains the process utilized to identify elderly users.

4.1 Dataset

The data provided by NOS Inovação[®] was collected over the course of April, May and June in 2019. It was comprised of two *.tsv* files. One contained information on the available VOD (Video on Demand) and EPG (Electronic Program Guide) television content, featuring 331,865 items, each identified with a content ID, consisting of 58.7 Megabytes. The other file contained the usage data, consisting of 12,018,379 content visualisations from 100,000 devices, corresponding to users, each identified with an anonymised user ID, containing 1.16 Gigabytes of data. The information featured on the content file included the television genres of each entry, as well as directors, year of release, duration, age rating and some of the actors who participated in the content in question. The usage file contained the time and date when the user started watching each content (identified with its respective ID) and when they finished the visualisation.

4.2 Computational Methods

The project's coding was realized with the Python language, utilizing the Spyder code editor. To organize the dataset and extract features, different libraries such as Pandas [84] and NumPy [85] were used, while the Scikit-Learn [86] and Surprise [87] libraries were employed in the ML applications. Visualisations of the data and graphical explanations for the recommendations were developed with the usage of the Seaborn and Matplotlib libraries.

4.3 Processing of the Dataset

The first step in processing the data was to load each *.tsv* file into a Pandas Dataframe. This is a two-dimensional tabular data format, in which the rows represent entries while the columns are used to represent their features.

In this project, we focused on VOD content, which is usually associated with recommendation systems used by popular media services such as Netflix [88] and large television providers [89]. As such, the EPG information was filtered out of both Dataframes, alongside columns expressing characteristics exclusive to this type of content like the television channel they are featured in. VOD trailers were also removed, resulting in 364,230 visualisations and 24,493 items remaining, still including different episodes of the same series.

As previously mentioned, the usage data included the beginning and end of each visualisation, but not the actual duration. The next stage was therefore to calculate it. This was done by converting it's ISO 8601 format into minutes and subtracting the time of day in which each visualisation started from when it ended, adding this information into a new column in the usage Dataframe. It was followed by the removal of entries from this Dataframe that were less than 5 minutes long, in order to remove overly short interactions, which may negatively affect exploration results. Entries lacking in user ID or content ID were also removed.

The two Dataframes were then combined by matching content IDs, resulting in a new usage Dataframe which includes each item's characteristics. Because many series were divided into episodes, some more processing was required to organize a user's visualisation of an item into a single row. With this, we have finally organized the relevant data into a single Dataframe containing the features displayed in table 4.1, describing 84,322 content visualisations in total.

Table 4.1: Features of the processed dataset.

Column	Description
UserID	A string used to identify the user
Genres	A string containing each genre of the content item, separated by coma.
Season	The season of the series, it features a <i>nan</i> instead if it's a movie.
Year	The year when the item was released.
Rating	The age rating for the item.
Directors	The item's directors.
Actors	Notable actors featured in the item.
ContentName	Name of the item.
ContentID	String number used to identify the item.
ViewTime	Time duration of the user's visualisation for the respective item.
Duration	Total time duration of the item. Considers all episodes of a season.

4.3.1 Exploratory Data Analysis

When it comes to content-based filtering for recommendation systems, television genres, actors and directors are amongst the most useful meta-data features recognised in the literature [90, 91]. Out of these, the content's genre in particular seemed the most likely to allow us to identify elderly users through their viewership habits. This was because, cultural surveys tend to include statistics about said genres, as well as viewing time [16]. As such, an exploratory data analysis of the dataset was performed with a focus on the genre classification of contents.

First, an histogram was plotted for each television genre, allowing the visual representation of the distribution of viewing times for items included in it. Then, a mean of these times, as well as their standard deviation and maximum and minimum values, were calculated.

4.4 Proposed Recommendation System

In the literature, healthcare recommendation systems often gauge the interests of each user in order to generate more personalised suggestions, which increases the chances that users actually following said recommendations [58, 80]. Collaborative techniques can also group users in a way that is useful when addressing similar needs [79].

With these considerations, the next step in the development of the recommendation system was to obtain a quantifiable appreciation for the items visualised by each user. This was done by dividing the *ViewTime* by the *Duration* for each entry. The result is a ratio between the viewing time of the user and the total duration of the item, which was then normalized into a range from 1 to 10 added as a new column in the Dataframe named *ViewerRating*.

Next, a user-item matrix was created, with a column for each item and a row for each user, featuring the *ViewerRating* for each pair. When a user didn't watch a certain piece of content, the rating was considered 0. From this matrix, it could be seen that 18,124 users and 6537 items remain from the previous processing.

The metadata from the previously obtained Dataframe and ratings matrix allowed the application of collaborative and content-based techniques into the creation of a conceptual recommendation system. As was previously mentioned in Chapter 2, hybrid recommendation systems aim to reduce the weaknesses of other filtering techniques by combining them. By employing these two approaches at once, we hoped to avoid the "cold start" problem and to decrease the effects of data sparsity, as each user only watched a few of the items featured in the dataset.

4.4.1 Content-Based Filtering

As was mentioned in the Data Exploration section of this chapter, the *Genres*, *Directors* and *Actors* columns of the Dataframe contain prime features to be used in the content-based component of a recommendation system. Following the method proposed in [92], a new column was added into the usage Dataframe containing these features merged into a "metadata soup". That is, a string containing each genre, actor and director respective to the item. This format allows the application of Scikit-Learn's CountVectorizer, a feature extraction technique which converts text into a dimensional vector of token (word) counts. These vectors can then be compared to find the content similarity of different items.

A 6537 by 6537 cosine similarity matrix was then built, in where the cell in the i^{th} row and j^{th} column represents the similarity between items i and j . This matrix was the foundation of the proposed recommendation system's content-based component. The concept was tested in the form of a top 10 recommendation which took a contentID as an argument and generated a recommendation for the ten most similar items.

4.4.2 Collaborative Filtering

For the collaborative component of the proposed recommendation system, ML algorithms were used to predict user's rating for items they have not yet seen. The chosen approach was user-based, meaning the predictions are based on the behaviour of similar users.

In order to properly utilise the Surprise library's functionalities, a *ratings* Dataframe was built, containing only the IDs for the user and series, as well as the rating for each particular pair. The algorithms chosen were the KNN and SVD, as they are often cited in the literature [27, 28, 59].

To complete the desired hybrid system, predictions generated by these algorithms were incorporated into the previously mentioned content-based concept: taking both a user and item as arguments, it finds the 25 contents that are the most similar to the item and orders them according to the predicted rating of the user, presenting only the top 10 as recommendations.

To evaluate the accuracy of the ML components of the system, the RMSE was calculated for each algorithm. The precision and recall were also computed in order to assess their performance. This was done through a K-fold method, in which the data was divided into 5 subsets, with the last one being used as the test set and the other as training sets.

4.5 Recommendation Explanations

To explain and justify the recommendations generated by the proposed system, a model agnostic approach was applied. Although these explanations do not uncover the latent factors in black box algorithms, like SVD, they still provide a personalised justifications for the suggestions that tend to increase user's trust [66].

The first type of explanation developed was a of a textual nature. It attempts to explain both the content-based and the collaborative components of the recommendation through common characteristics between items. This approach compares the metadata between the suggested items and the one used as argument in order to explain their similarity. It also compares the recommended content to each item that was highly rated by the user. The common features found in these comparisons are then displayed to justify the recommendation.

Another proposed explanation style was the nearest neighbors recommendation, which was found to be well rated by users in the literature [82]. For each item suggested, an histogram is presented, displaying the ratings of said content by the 10 users most similar to the target one. Cosine similarity was the metric used in this process.

The evaluation of explanations in this project focused on explainability of the chosen approaches, calculated with MEP as was proposed in [71]. For the textual explanations, having at least two common features with content previously enjoyed by the user was the minimum to be considered explainable. While for the nearest-neighbors explanations the existence of 10 nearest users who have seen the recommended content was the threshold utilized to classify an item recommendation as explainable.

4.6 Identification of Elderly Users

One of the objectives of this project was the identification of elderly users based on the data provided by NOS. In order to fulfil this purpose, statistical data from Marktest, a major market study group in Portugal was requested. In response, the statistics described in Figure 4.1 were supplied. This data concerns the types of content watched by elderly users, compared with the total population.

Typologies per year - 2019 to 2021

	Total Day							
	Older than 65				Universe			
	rat%	rat(000)	shr%	ats	rat%	rat(000)	shr%	ats
2019								
FUN	7,5	153,5	22,5	00:36:00	3,5	333,0	15,2	00:31:49
INFORMATION	7,2	146,2	23,1	00:27:41	3,5	335,5	16,4	00:25:37
SPORT	5,7	116,0	13,7	00:27:59	3,3	316,2	10,9	00:25:56
OTHER	3,9	76,5	10,3	00:01:44	1,7	166,3	6,3	00:01:45
FICTION	3,7	75,6	13,8	00:22:38	2,5	237,7	11,7	00:25:14
KNOWLEDGE	1,5	30,9	5,6	00:13:28	0,7	68,2	3,6	00:14:08
ART	1,3	25,9	3,4	00:13:06	0,5	50,2	1,7	00:12:28
DIVERSE	1,3	25,6	11,4	00:18:25	0,9	90,4	8,1	00:22:34
KIDS	0,2	4,3	0,7	00:11:24	0,3	31,4	1,7	00:19:05
PUBLICITY	0,1	2,9	5,6	00:20:54	0,2	16,5	4,6	00:35:23
2020								
FUN	8,1	174,8	22,1	00:36:46	4,0	380,4	14,5	00:32:41
INFORMATION	7,8	167,2	23,4	00:30:14	4,0	382,9	16,5	00:27:48
SPORT	6,8	148,2	14,6	00:28:20	4,2	399,0	11,8	00:26:37
OTHER	5,1	109,6	12,4	00:02:02	2,0	188,6	6,1	00:02:07
FICTION	4,5	97,4	14,8	00:23:27	3,1	291,6	12,1	00:24:56
DIVERSE	2,4	50,5	16,7	00:14:51	1,6	154,0	10,8	00:15:51
KNOWLEDGE	2,0	43,0	6,4	00:13:47	0,9	86,9	3,8	00:13:55
ART	1,1	24,6	2,8	00:12:44	0,5	44,3	1,3	00:11:39
PUBLICIDADE	0,2	4,2	7,8	00:28:13	0,2	17,4	4,6	00:36:20
JUVENTUDE	0,2	4,1	0,6	00:11:22	0,3	28,6	1,3	00:20:26
2021								
FUN	7,6	168,3	21,2	00:36:46	3,8	360,7	14,5	00:33:35
INFORMATION	7,6	167,9	22,8	00:28:55	4,0	374,9	16,3	00:27:22
DIVERSE	6,4	141,1	17,9	00:02:05	3,6	341,2	13,1	00:02:04
SPORT	4,5	100,2	12,9	00:29:25	3,0	283,7	11,3	00:29:46
FICTION	4,4	97,7	14,9	00:22:19	2,9	277,5	11,9	00:23:09
OTHER	4,3	94,6	9,6	00:02:14	2,5	240,9	7,5	00:02:12
KNOWLEDGE	1,7	38,6	5,3	00:13:52	0,8	79,4	3,4	00:14:24
ART	0,9	18,9	1,9	00:10:56	0,4	34,0	0,9	00:09:51
PUBLICITY	0,2	4,9	8,3	00:29:34	0,2	17,2	4,6	00:38:24
KIDS	0,2	4,7	0,7	00:12:12	0,2	22,2	1,1	00:18:25

Source: CAEM/MediaMonitor
data obtained from YUMI Analytics Desktop

audience with live viewership and recordings viewed in the same day
Universe: 9.471.100 individuals
Older than 65: 2.228.800 individuals

Figure 4.1: Television viewership statistics for the elderly in Portugal, adapted from [93].

According to [93], the metrics used in Figure 4.1 are described as in Table 4.2. Among these, Average Time Spent was considered to be the most useful for this project, allowing for some comparative understanding of elderly television viewing times, in relation to content genres. Priority was given to the statistics from 2019, since that was the year in which the NOS dataset was also collected.

The data in Figure 4.1 is not without limitations however: Considering its size, the study's universe likely includes individuals who, despite not being older than 65, may be very similar in age and possessing similar viewership habits. For future work, a study stratified by age groups would likely to allow for more consistent results.

Table 4.2: Metrics displayed in the Marktest data.

Metric	Description
Rating	Mean audience per minute.
Share	Audience share.
Average Time Spent(ATS)	Average time spent by each user watching a piece of content.

4.6.1 Clustering

In order to obtain features similar to those presented in [93], as they're described in Marktest's website [94] a new Dataframe was created. With each row representing a user, each column contains the mean amount of viewing time among items seen containing certain genres as presented in Table 4.3. It also contained a column for the mean viewing time from among all items seen by respective user.

Table 4.3: Conversion of television genres into the content typologies studied by Marktest.

Feature	Genres
Entertainment	<i>entertainment, magazine, music, musical</i>
Information	<i>information</i>
Sport	<i>sport</i>
Knowledge	<i>biography, documentary, history</i>
Kids	<i>kids, animation</i>
Other	<i>adult, erotic</i>
Fiction	(all genres not included in other features)

Clustering techniques were then applied to the features obtained in Table 4.3 as well as the mean viewing time for all genres, with the intention of identifying a cluster from among the dataset that is likely to contain elderly users.

In order to reduce the dimensionality of this data while losing as little information as possible, PCA and t-SNE were employed as dimension reduction techniques.

To test the performance of PCA, the explained variation for principal component was computed in order to verify if only two or three dimensions, used to plot the clusters, were appropriate.

According to the literature, the perplexity values used in [t-SNE](#) applications should range from 5 to 50 [[45](#), [95](#)]. To assess the appropriate perplexity for the [t-SNE](#) usage in cluster visualisation, values of 5, 50 and 30 were tested. This was done by producing scatter-plots using the first and second dimensions obtained from the application of [t-SNE](#) and comparing the effect of this variable between them.

In order to find the optimal number of clusters to classify the users into, the Silhouette Method was utilized [[86](#)]. A silhouette coefficient was computed for different numbers of clusters, ranging from 2 to 6 and the one with the highest coefficient was chosen.

The plots obtained by [PCA](#) and [t-SNE](#) were then visually compared, to distinguish which technique would be more appropriate for the desired study. For each cluster obtained, it's features were visually represented using bar charts and compared with the data in [Figure 4.1](#).

RESULTS AND DISCUSSION

This chapter presents and discusses the results obtained from the methods that were previously described. It starts with the presentation of the developed recommendation system and the employed explanation techniques alongside their evaluations. Then it will cover the clustering analysis and the identification of elderly users from among the dataset.

5.1 Exploratory Data Analysis

In total, 31 genres were detected in the dataset. Some of the histograms obtained from the viewing times for each genre are presented in Figure 5.1. From these representations, it was perceived a significant variety of visualisations durations between genres, which motivated a more thorough data exploration, described in Table 5.1, where the mean viewership time for content items included in each genre and its standard deviation, as well as minimum and maximum time spent viewing said items are presented.

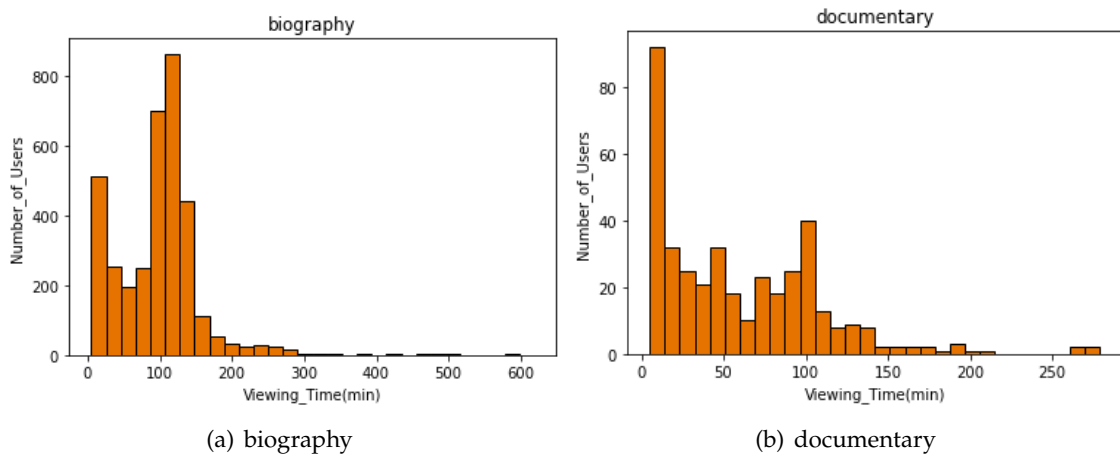


Figure 5.1: Examples of histograms obtained for the television genres "biography" 5.1(a) and "documentary" 5.1(b).

Table 5.1: Statistics for the viewing time of each television genre in the dataset, in minutes.

Genre	Mean	Standard Deviation	Minimum	Maximum
action	169	199	5	4634
adult	33	47	5	1176
adventure	140	171	5	4744
animation	107	173	5	5308
biography	97	61	5	618
classic	72	80	5	411
comedy	137	390	5	11722
crime	153	217	5	2266
documentary	60	49	5	279
drama	181	328	5	11722
entertainment	56	86	5	828
erotic	40	46	5	195
family	107	166	5	5308
fantasy	168	236	5	5043
history	135	147	5	1332
horror	116	155	5	1376
information	67	110	6	538
kids	56	133	5	1190
magazine	284	901	5	9299
martialarts	156	200	17	968
music	404	1163	5	11722
musical	332	896	5	10937
mystery	111	156	5	1427
romance	129	179	5	5526
scifi	157	176	5	2724
serie	401	407	5	2019
sport	80	86	5	989
thriller	107	112	5	1828
tutorial	29	17	5	53
war	102	110	5	977
western	97	61	5	422

From the statistics displayed in Table 5.1, the effect of the 5 minutes threshold is perceived in the minimum viewing times, indicating most genres featured some very short content visualisations that were filtered out. The high values of standard deviation when compared to the respective mean, suggest the importance of collaborative features that do not consider absolute viewing times, but instead their relation with the full duration of the respective content item, like the user rating concept that was employed in this project, described in section 4.4. The diversity and high values of maximum visualisation times are reflective of the grouping of episodes for each series and the large variety of total durations among items.

5.2 Recommendation System Prototype

As was described in the previous chapter, the proposed recommendation system takes both a userID and itemID as arguments and attempts to generate ten recommendations, ordered by their predicted rating. This simulates the generation of recommendations after a user just watched a piece of content. Before the implementation of any explanation, these consist of a simple list of suggested content.

The K-Fold assessment of the KNN and SVD classifiers is described in table 5.2. These results suggest that the SVD is the more accurate algorithm for the dataset. Although it's mean RMSE represents a deviation of more than 3 units in the context of predicted 1 to 10 rating, which is enough to affect the recommended items.

The values obtained for the Precision and Recall also point towards SVD being the better algorithm for this type recommendation system. A higher Recall suggests that the system is efficient when it comes to including the items which are predicted to be highly rated, into the recommendation. The Precision values also indicate that most recommended items are predicted to be well rated by the target user.

Table 5.2: Performance and accuracy of each model.

Algorithm	Metric	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean	Std
KNN	RMSE	3.43	3.46	3.44	3.53	3.47	3.47	0.03
	Precision	0.82	0.81	0.82	0.82	0.81	0.82	0.003
	Recall	0.84	0.83	0.84	0.83	0.83	0.84	0.004
SVD	RMSE	3.09	3.14	3.16	3.11	3.15	3.13	0.03
	Precision	0.83	0.83	0.82	0.83	0.83	0.83	0.003
	Recall	0.86	0.86	0.86	0.86	0.86	0.86	0.003

5.2.1 Generated Explanations

In the case of the textual explanation style, the genres, actors and directors shared by the recommended item, the recently watched item and contents previously watched by the

user, are presented after the title of the suggested movie or series, as can be perceived in the following examples:

Anatomia de Grey T2

You have enjoyed content featuring: Shonda Rhimes, Patrick Dempsey, Justin Chambers, Sandra Oh, Ellen Pompeo, drama

Pequena Sereia II (VP)

You have enjoyed content featuring: animation, drama, family

The neighbors style of explanation, on the other hand, plots an histogram of ratings obtained from the ten most similar users as it's presented in Figure 5.2, for each recommendation.

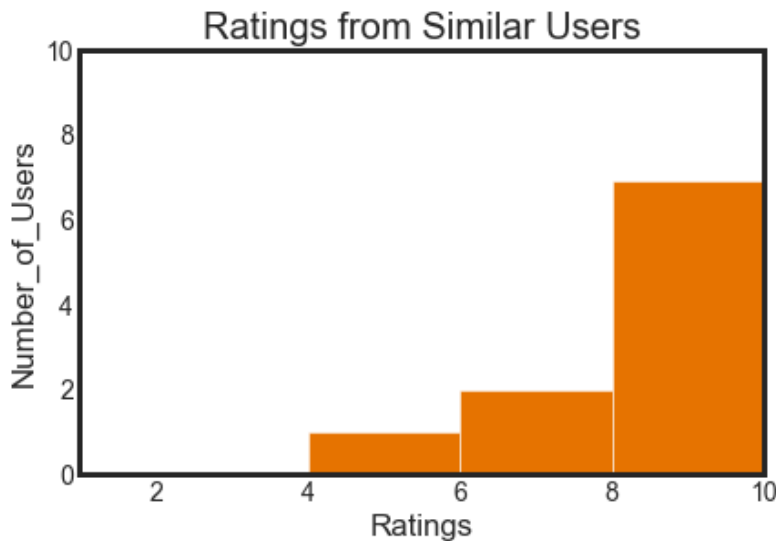


Figure 5.2: Histogram with ratings for the recommended item, from the ten most similar users.

For both explanation styles, the recommendations are only presented if they're considered above or equal to the thresholds described in section 4.5, similarly to the method introduced in [66]. Therefore, the MEP assessments, featured in Table 5.3, also suggest how many recommendations are to be expected from the implementation of each explanation style, out of the ten possible ones for each user and item combination.

These results propose that the neighbors explanation can only be applied to around half the recommendations, while the textual explanations can cover most. This indicates that finding similar users who have watched the recommended content is far more demanding than detecting similarities between items in the recommendation and those in a user's

Table 5.3: Evaluation of explanations.

Algorithm	Explanation Style	MEP
KNN	Textual Explanation	0.9
	Neighbors Histogram	0.52
SVD	Textual Explanation	0.9
	Neighbors Histogram	0.52

profile. However, it's also important to consider how the literature suggests a higher approval for graphical explanations among users [82]. If such possibility were to be verified in future work, the results obtained in this project could indicate a trade-off between the amount of explainable recommendations, and their appreciation by the users.

Table 5.3 also suggests a confirmation of the model-agnostic nature of the chosen explanation style, considering the lack of different results for the distinct algorithms. This proposes that the SVD could be chosen without a loss of explainability using while following this approach, despite being classified as black-box.

5.3 Clustering

The mean and standard deviation values were computed for each of the features described in section 4.6.1 used in the clustering application. The results are presented in Table 5.4.

Table 5.4: Statistics for the clustering features.

Feature	Mean	Standard Deviation
All	72.70	36.39
Entertainment	10.24	34.46
Information	0.05	1.71
Sport	3.20	17.29
Knowledge	11.58	31.26
Kids	20.43	35.52
Other	1.27	7.19
Fiction	72.33	38.10

These values are quite different from those of *ATS* in Figure 4.1, likely due to being obtained from videos on demand instead of passive television viewership. The standard deviation of the Marktest data is also not included so establishing direct comparisons is challenging. However, it was considered that a relative comparison focused on which typologies are more relevant for elderly users would still allow for some results.

5.3.1 Principal Component Analysis

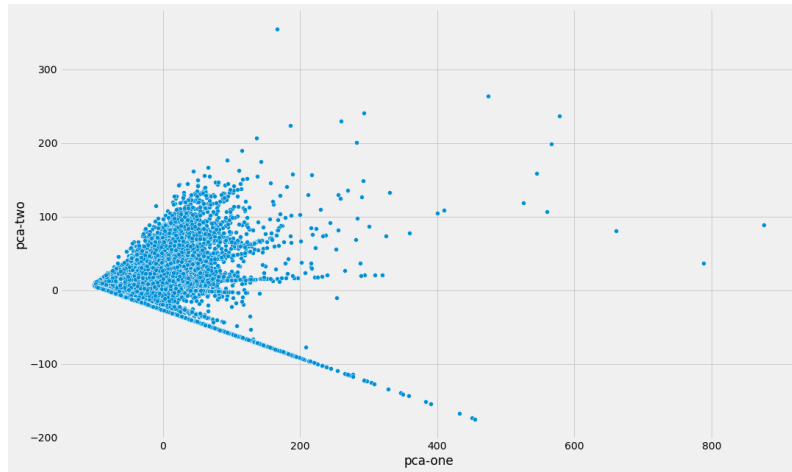
In the initial stage of PCA application, the explained variation calculated for each of the first three principal components is displayed in Table 5.5.

Table 5.5: Evaluation of Principal Components.

Principal Component	Explained Variation
First	0.65
Second	0.27
Third	0.05
Cumulative	0.97

These results suggest that most of the variation for the features can be explained using only the first two principal components. A reduction of those initial seven features into said components was therefore considered appropriate for the [PCA](#) application in the clustering classification.

The dispersion plot resulting from taking the two first principal components of the data as axis is presented in [Figure 5.3](#).

Figure 5.3: Scatter-plot for the [PCA](#) dimension reduction.

It's not possible to visually perceive any separable clusters in [Figure 5.3](#). This indicates that [PCA](#) is not an appropriate technique with which to separate the dataset. Considering that [PCA](#) assumes linearity in the data [\[41\]](#), this result suggests that the linear relationship between the features utilized is very low.

5.3.2 t-Distributed Stochastic Neighbor Embedding

The obtained scatter-plots for the assessment of the [t-SNE](#), using perplexity values of 5, 30 and 50 are presented in [Figure 5.4](#). The horizontal and vertical axis correspond to the first and second [t-SNE](#) dimensions respectively.

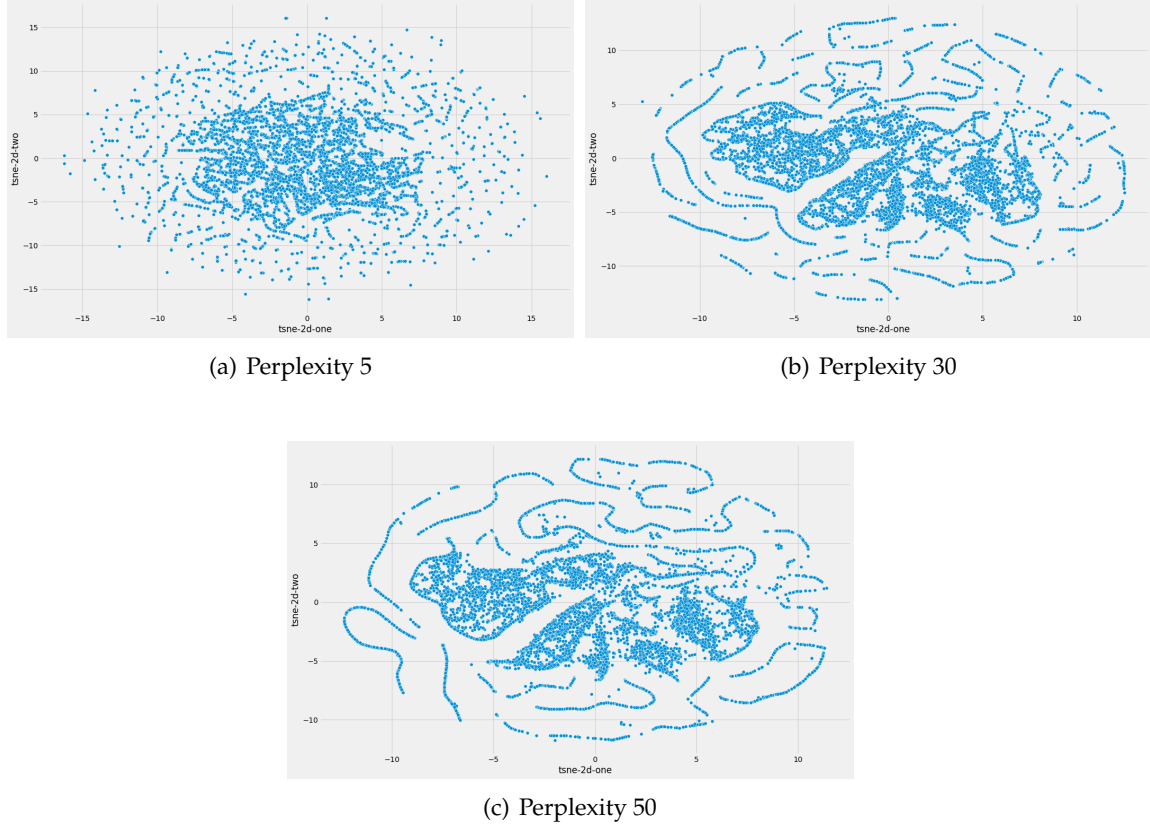


Figure 5.4: Scatter-plots of t-SNE dimensions for different perplexity values.

The lack of any perceivable clusters in 5.4(a) suggests that low perplexity is not appropriate for the intended clustering. On the other hand, the similarities between 5.4(b) and 5.4(c) indicate that high perplexity values do not cause too much variation of results. This suggests that separable structures in the dataset are only discernible at the higher ends of recommended perplexity values. As such, the intermediate value of 40 was chosen in the application of t-SNE for the following steps in the project.

According to the literature, higher perplexity translates to a better preservation of global structure in detriment of the local one [45, 95]. As such, the results obtained in Figure 5.4 indicate that this data is only separable when considering its global structure.

5.3.3 Silhouette Coefficient

The silhouette coefficients obtained for different k numbers of clusters, applied to t-SNE dimensions is presented in Figure 5.5. As was mentioned in 2.3.1, a coefficient of 1 would have suggested a perfect separation between clusters.



Figure 5.5: Silhouette Coefficient for different numbers of clusters.

Figure 5.5 indicates that a division of the data into 5 clusters is the most appropriate, as it presents the highest coefficient and a visually discernible maximum. The silhouette coefficient computed for this number of k was approximately 0.40. This value suggests that some separation is possible, although far from perfect, a result which reflects the shapes observed in the Figures 5.4(b) and 5.4(c).

5.4 Identification of Elderly Users

In order to visually identify the clusters, each data point, representing each user, was colored according to the result of the K-means clustering algorithm applied to the two first t-SNE dimensions of the dataset. This process produced the plot represented in Figure 5.6.

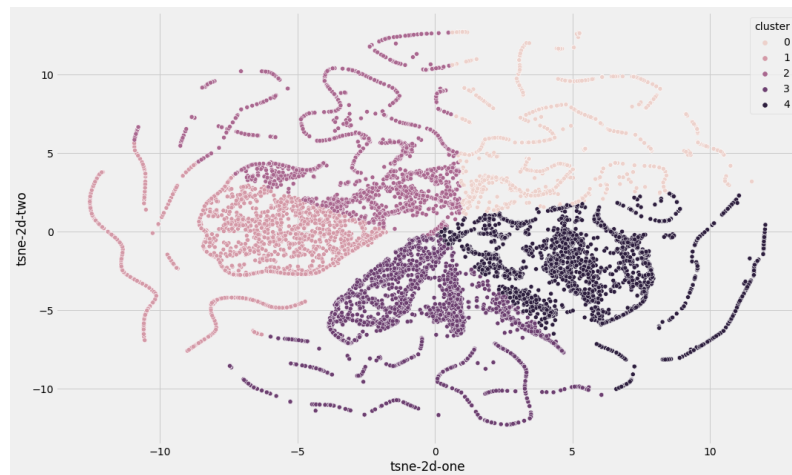


Figure 5.6: Plot of the clusters.

From Figure 5.6 it can be perceived how each cluster includes the shapes identified in previous plot, yet also contain many outliers. These are possibly a reflection of the

high values of standard deviation obtained for the features in Table 5.4, which in turn are congruent with the high variation of viewing times in the original dataset.

To visualise the mean values for each feature in each cluster, the histograms presented in Figure 5.7 were produced.

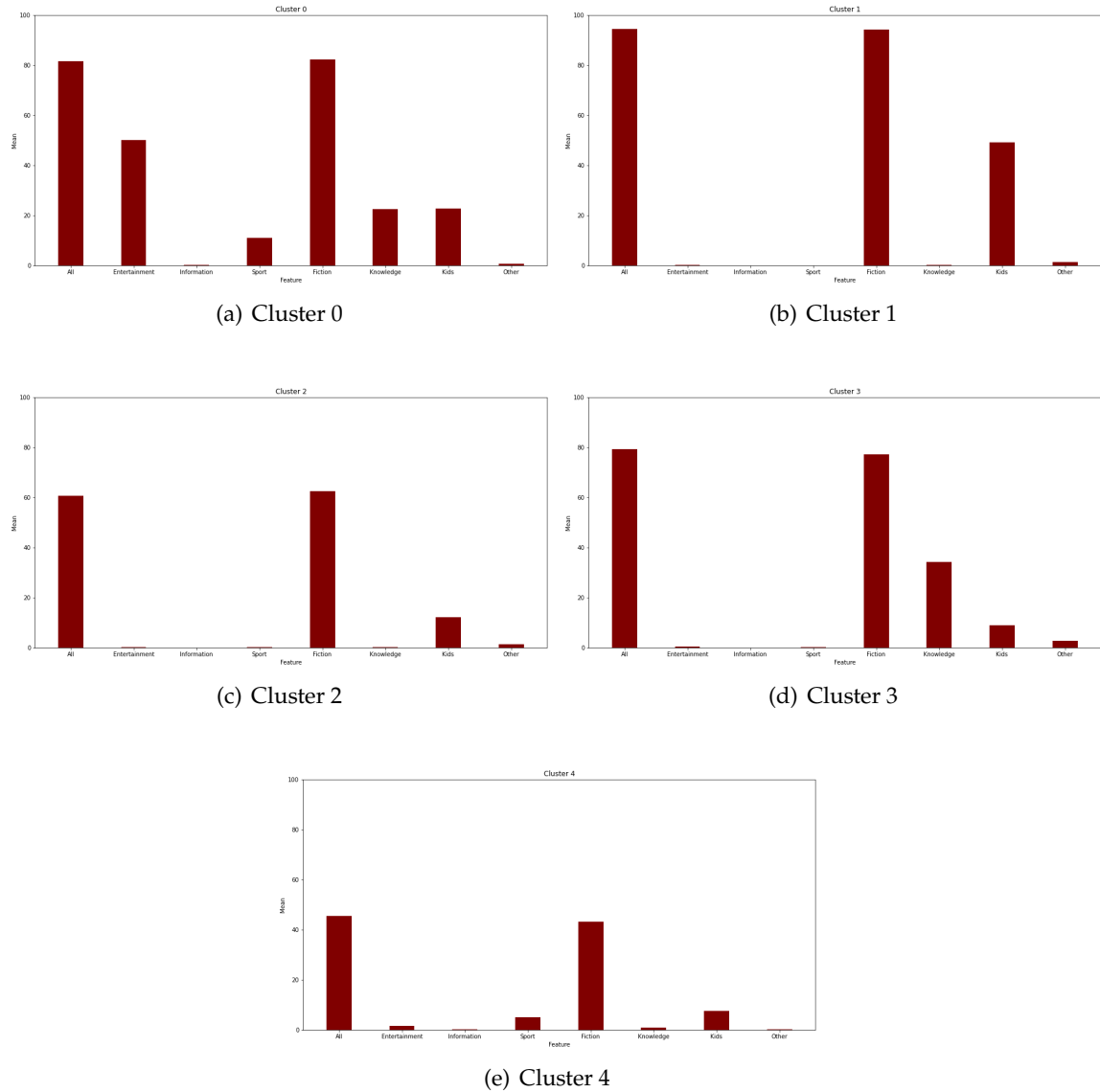


Figure 5.7: Mean values for each cluster.

As was mentioned in section 5.3, to identify which cluster is more likely to contain elderly users, comparisons were made between which features in Figure 5.7 are above or below their respective mean in Table 5.4. This was then compared to which typologies in Figure 4.1 are above or below in *ATS* for elderly individuals when compared to the universe. Among the years presented in the Marktest statistics, 2019 was chosen because it's the year in which the NOS data was collected.

Upon an initial visual inspection, 5.7(a) appeared to be the most plausible to include elderly users as it features the highest value for *Entertainment* and an above average

value for *Sport*. It was assumed that the former might be particularly important, as it includes content usually associated with elderly viewers, such as telenovelas and rural cultural shows.

In order to obtain a more detailed perspective, the mean values for each cluster are presented in Table 5.6.

Table 5.6: Statistics for the clustering features.

Cluster	All	Entertainment	Information	Sport	Knowledge	Kids	Other	Fiction
0	81.64	50.12	0.09	11.06	22.58	22.64	0.73	82.38
1	94.42	0.19	0.01	0.04	0.14	49.26	1.44	94.32
2	60.65	0.19	0.01	0.09	0.18	12.05	1.25	62.58
3	79.38	0.34	0.00	0.21	34.26	8.96	2.71	77.26
4	45.51	1.61	0.12	5.00	0.77	7.52	0.14	43.28

When comparing these results with Table 5.4, it can be perceived that Cluster 0 has above average means for *Entertainment*, *Information* and *Sport*, as well as below average mean values for *Other*, which is consistent with elderly viewing habits in Figure 4.1. The fact that the overall mean watch time for contents is above average can also be an indication of older users [16]. However, the other features do not reflect the statistics for individuals who are older than 65, in particular the mean value for *Kids* and *Knowledge* are above average. These may be a result of outliers, which were very perceptible in Figure 5.6. They can also be explained by limitations in the genre classification of items. For example, contents with the *animation* genre will always count towards the *Kids* feature even if they're meant for adults.

Cluster 4 presents above average *Sport* and *Information*, as well as below average *Knowledge*, *Kids* and *Fiction*. However, its *Entertainment* value is very low.

The remaining clusters do not display many similarities with the viewership patterns of elderly individuals proposed in the Marktest data. Although the importance of the *Entertainment* many require some future work to be better understood, these results suggest that Cluster 0 is the more likely one to contain elderly users.

CONCLUSION AND FUTURE WORK

This chapter summarizes the work produced during the development of this dissertation alongside its conclusions, followed by suggestions for future work.

6.1 Conclusion

The objective of this dissertation was the development of an explainable recommendation system for television contents which could be employed to promote active aging and well being, while also identifying elderly individuals through their television viewership habits.

Using a dataset supplied by NOS Inovação®, this project resulted in the creation of a hybrid recommendation system for television content, developed by employing both collaborative filtering, implemented through [KNN](#) and [SVD](#) algorithms, and content-based filtering, obtained with methods based on cosine similarity between items. Two explanation styles were proposed for the system: a textual explanation, featuring the similarities between content favoured by the user and the recommendations, and a visual explanation, presented in the form of an histogram of neighbors ratings for the suggested item. For the recognition of elderly people within the dataset, clustering techniques were applied to divide users into groups with similar habits, employing [PCA](#) and [t-SNE](#). The group with whose features were most consistent with statistics supplied by Marktest, a major market study group in Portugal, concerning the television consumption characteristics of older populations was selected.

The assessment of the [ML](#) algorithms used in the system indicated better results for [SVD](#) across all metrics, including accuracy, precision and recall. This attests to the importance of the development of explanation models, since unlike [KNN](#), this algorithm isn't inherently interpretable.

When evaluating the explanation styles, their model-agnostic nature was evidenced by obtaining the same [MEP](#) with both algorithms. Although the textual explanation was shown to be less demanding, neither style could explain all the recommendations, reflecting the trade-off between explainability and performance.

The clustering application did not present separable results with [PCA](#), suggesting that method isn't reliable for the intended study. [t-SNE](#) allowed for the separation of clusters, and it was possible to select one as the most likely to contain elderly users. However, the comparison of the features obtained from the dataset to the available statistics was challenging and required some assumption about the importance of certain genres. Considering how elderly people are some of the biggest consumers of this type of media, there is potential in the creation of better studies about their viewing habits.

Overall, this dissertation produced an explainable recommendation system that could be implemented into other health applications and proposed a method to detect older users from a dataset pertaining to television usage. These types of systems show a lot of potential for non-intrusive solutions in the healthcare domain that can be personalised to individual taste. This dissertation intends to provide a basis and motivation for further research into real-world applications of [ML](#) and media usage to promote active aging.

6.2 Future Work

The developed work still contains some limitations that could be improved in further studies which are presented in this section alongside some ideas for additional research.

The recommendation system was tested in a static context, without the constant additions of new users and item that would be present in a real world application. Additionally, not all types of data were utilized. Future studies could find a way to use other formats such as EPG in conjunction with the VODs. Further research could also point towards a knowledge filtering component for the recommendations, which was mentioned in the literature to improve performance.

Testing the explanation styles in interactions with real users would allow for a more thorough evaluation, using metrics such as scrutability and trust. Further research in that direction could be used to better specify thresholds with which to consider the explanations acceptable. Other approaches, such as LIME and SHAP could also be implemented in future work.

Further studies could also focus on obtaining higher quality data about the habits of elderly individuals that would allow for a more precise identification of this population within the dataset. These could evaluate the assumptions made about typologies like *Entertainment*.

A similar framework to the one proposed in this dissertation could be employed to other vulnerable members of the population, such as individuals with depression and anxiety.

An important suggestion for future work would be the development of active aging and well being focused recommendations that would allow the system to directly promote these concepts.

BIBLIOGRAPHY

- [1] J. M. Lourenço. *The NOVAthesis L^AT_EX Template User's Manual*. NOVA University Lisbon. 2021. URL: <https://github.com/joaomlourenco/novathesis/raw/master/template.pdf> (cit. on p. ii).
- [2] W. H. Organization et al. "Decade of healthy ageing: baseline report". In: (2020) (cit. on p. 1).
- [3] W. H. Organization. "World report on ageing and health". In: World Health Organization, 2015, 246 p. (Cit. on p. 1).
- [4] L. J. Dominguez and M. Barbagallo. "The biology of the metabolic syndrome and aging". In: *Current Opinion in Clinical Nutrition & Metabolic Care* 19.1 (2016), pp. 5–11 (cit. on p. 1).
- [5] C. Chen et al. "Body weight, frailty, and chronic pain in older adults: a cross-sectional study". In: *BMC geriatrics* 19.1 (2019), pp. 1–10 (cit. on p. 1).
- [6] C.-P. Chung et al. "Physio-cognitive decline syndrome as the phenotype and treatment target of unhealthy aging". In: *The journal of nutrition, health & aging* (2021), pp. 1–11 (cit. on p. 1).
- [7] D. Griol Barres, J. M. Molina López, and M. A. Sanchis de Miguel. "A multimodal conversational coach for active ageing based on sentient computing and m-health". In: (2020) (cit. on p. 1).
- [8] W. H. Organization. "Active ageing : a policy framework". In: Active ageing series (2002), Japanese publication also includes the title: Global age friendly cities a guide 2007 (cit. on p. 1).
- [9] T. D. Robbins, S. N. Lim Choi Keung, and T. N. Arvanitis. "E-health for active ageing; A systematic review". In: *Maturitas* 114 (2018), pp. 34–40. ISSN: 0378-5122. DOI: <https://doi.org/10.1016/j.maturitas.2018.05.008>. URL: <https://www.sciencedirect.com/science/article/pii/S0378512218301725> (cit. on p. 1).

-
- [10] D. Bacciu et al. "Smart Environments and Context-Awareness for Lifestyle Management in a Healthy Active Ageing Framework". In: *Progress in Artificial Intelligence*. Ed. by F. Pereira et al. Cham: Springer International Publishing, 2015, pp. 54–66. ISBN: 978-3-319-23485-4 (cit. on p. 1).
 - [11] R. Priya and S. V. Jinny. "Elderly healthcare system for chronic ailments using machine learning techniques—a review". In: *Iraqi Journal of Science* 62.9 (2021), pp. 3138–3151 (cit. on p. 1).
 - [12] A. Sau and I. Bhakta. "Predicting anxiety and depression in elderly patients using machine learning technology". In: *Healthcare Technology Letters* 4.6 (2017), pp. 238–243. DOI: <https://doi.org/10.1049/htl.2016.0096>. eprint: <https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/htl.2016.0096>. URL: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/htl.2016.0096> (cit. on p. 1).
 - [13] M. F. Alhamid et al. "Leveraging Biosignal and Collaborative Filtering for Context-Aware Recommendation". In: *Proceedings of the 1st ACM International Workshop on Multimedia Indexing and Information Retrieval for Healthcare*. MIIRH '13. Barcelona, Spain: Association for Computing Machinery, 2013, pp. 41–48. ISBN: 9781450323987. DOI: [10.1145/2505323.2505332](https://doi.org/10.1145/2505323.2505332). URL: <https://doi.org/10.1145/2505323.2505332> (cit. on p. 1).
 - [14] A. F. Markus, J. A. Kors, and P. R. Rijnbeek. "The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies". In: *Journal of Biomedical Informatics* 113 (2021), p. 103655 (cit. on pp. 2, 9, 10).
 - [15] C. Christophorou et al. "ICT services for active ageing and independent living: identification and assessment". In: *Healthcare Technology Letters* 3.3 (2016), pp. 159–164. DOI: <https://doi.org/10.1049/htl.2016.0031>. eprint: <https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/htl.2016.0031>. URL: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/htl.2016.0031> (cit. on p. 2).
 - [16] J. Pais, P. Magalhães, and M. Antunes. "Inquérito às Práticas Culturais dos Portugueses 2020 Síntese dos Resultados." In: (2020) (cit. on pp. 2, 20, 35).
 - [17] J. Jovel and R. Greiner. "An introduction to machine learning approaches for biomedical research". In: *Frontiers in Medicine* 8 (2021), p. 2534 (cit. on pp. 3, 4).
 - [18] Z.-H. Zhou. "Machine learning". In: Springer Nature, 2021 (cit. on p. 3).
 - [19] A. C. Müller and S. Guido. "Introduction to machine learning with Python: a guide for data scientists". In: "O'Reilly Media, Inc.", 2016 (cit. on pp. 3, 4).

- [20] M. Kaur, A. K. Shukla, and S. Kaur. "An Introduction to Machine Learning in a Nutshell". In: *2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART)*. IEEE. 2021, pp. 17–22 (cit. on p. 3).
- [21] T. Jiang, J. L. Gradus, and A. J. Rosellini. "Supervised machine learning: a brief primer". In: *Behavior Therapy* 51.5 (2020), pp. 675–687 (cit. on pp. 3, 4).
- [22] M. W. Berry, A. Mohamed, and B. W. Yap. "Supervised and unsupervised learning for data science". In: Springer, 2019 (cit. on p. 4).
- [23] O. Biran and C. Cotton. "Explanation and justification in machine learning: A survey". In: *IJCAI-17 workshop on explainable AI (XAI)*. Vol. 8. 1. 2017, pp. 8–13 (cit. on p. 4).
- [24] C. Molnar. "Interpretable Machine Learning. A Guide for Making Black Box Models Explainable". In: 2nd ed. 2022. URL: <https://christophm.github.io/interpretable-ml-book> (cit. on p. 4).
- [25] R. Guidotti et al. "A Survey of Methods for Explaining Black Box Models". In: *ACM Comput. Surv.* 51.5 (2018-08). ISSN: 0360-0300. DOI: [10.1145/3236009](https://doi.org/10.1145/3236009). URL: <https://doi.org/10.1145/3236009> (cit. on p. 4).
- [26] B. Kovalerchuk and N. Neuhaus. "Toward Efficient Automation of Interpretable Machine Learning". In: *2018 IEEE International Conference on Big Data (Big Data)*. 2018, pp. 4940–4947. DOI: [10.1109/BigData.2018.8622433](https://doi.org/10.1109/BigData.2018.8622433) (cit. on p. 4).
- [27] Y. Koren. "The bellkor solution to the netflix grand prize". In: *Netflix prize documentation* 81.2009 (2009), pp. 1–10 (cit. on pp. 4, 6, 21).
- [28] R. H. Singh et al. "Movie recommendation system using cosine similarity and KNN". In: *International Journal of Engineering and Advanced Technology* 9.5 (2020), pp. 556–559 (cit. on pp. 4, 12, 21).
- [29] A. R. Lubis, M. Lubis, et al. "Optimization of distance formula in K-Nearest Neighbor method". In: *Bulletin of Electrical Engineering and Informatics* 9.1 (2020), pp. 326–338 (cit. on p. 4).
- [30] J. Gou et al. "A new distance-weighted k-nearest neighbor classifier". In: *J. Inf. Comput. Sci* 9.6 (2012), pp. 1429–1436 (cit. on p. 4).
- [31] K. Chomboon et al. "An empirical study of distance metrics for k-nearest neighbor algorithm". In: *Proceedings of the 3rd international conference on industrial application engineering*. Vol. 2. 2015 (cit. on p. 4).
- [32] W. Daelemans and A. Van Den Bosch. "Memory-based learning". In: *The Handbook of Computational Linguistics and Natural Language Processing* (2010), pp. 154–179 (cit. on p. 4).
- [33] A. Sagdic et al. "A Scalable K-Nearest Neighbor Algorithm for Recommendation System Problems". In: *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*. IEEE. 2020, pp. 186–191 (cit. on p. 4).

- [34] Z.-Y. Chai et al. "Recommendation System Based on Singular Value Decomposition and Multi-Objective Immune Optimization". In: *IEEE Access* 7 (2019), pp. 6060–6071. DOI: [10.1109/ACCESS.2018.2842257](https://doi.org/10.1109/ACCESS.2018.2842257) (cit. on p. 5).
- [35] K. Baker. "Singular value decomposition tutorial". In: *The Ohio State University* 24 (2005) (cit. on p. 5).
- [36] P. Bhavana, V. Kumar, and V. Padmanabhan. "Block based singular value decomposition approach to matrix factorization for recommender systems". In: *arXiv preprint arXiv:1907.07410* (2019) (cit. on p. 5).
- [37] B. Sarwar et al. "Incremental singular value decomposition algorithms for highly scalable recommender systems". In: *Fifth international conference on computer and information science*. Vol. 1. 012002. Citeseer. 2002, pp. 27–8 (cit. on p. 5).
- [38] R. Zebari et al. "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction". In: *J. Appl. Sci. Technol. Trends* 1.2 (2020), pp. 56–70 (cit. on p. 6).
- [39] S. Arora, W. Hu, and P. K. Kothari. "An Analysis of the t-SNE Algorithm for Data Visualization". In: *Proceedings of the 31st Conference On Learning Theory*. Ed. by S. Bubeck, V. Perchet, and P. Rigollet. Vol. 75. Proceedings of Machine Learning Research. PMLR, 2018-06–09 Jul, pp. 1455–1462. URL: <https://proceedings.mlr.press/v75/arora18a.html> (cit. on p. 6).
- [40] C. O. S. Sorzano, J. Vargas, and A. P. Montano. "A survey of dimensionality reduction techniques". In: *arXiv preprint arXiv:1403.2877* (2014) (cit. on p. 6).
- [41] B. M. S. Hasan and A. M. Abdulazeez. "A review of principal component analysis algorithm for dimensionality reduction". In: *Journal of Soft Computing and Data Mining* 2.1 (2021), pp. 20–30 (cit. on pp. 6, 31).
- [42] M. Ringnér. "What is principal component analysis?" In: *Nature biotechnology* 26.3 (2008), pp. 303–304 (cit. on p. 6).
- [43] F. Kherif and A. Latypova. "Chapter 12 - Principal component analysis". In: *Machine Learning*. Ed. by A. Mechelli and S. Vieira. Academic Press, 2020, pp. 209–225. ISBN: 978-0-12-815739-8. DOI: <https://doi.org/10.1016/B978-0-12-815739-8.00012-2>. URL: <https://www.sciencedirect.com/science/article/pii/B9780128157398000122> (cit. on p. 6).
- [44] P. Hajibabaei, F. Pourkamali-Anaraki, and M. A. Hariri-Ardebili. "An Empirical Evaluation of the t-SNE Algorithm for Data Visualization in Structural Engineering". In: *2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2021, pp. 1674–1680. DOI: [10.1109/ICMLA52953.2021.00267](https://doi.org/10.1109/ICMLA52953.2021.00267) (cit. on p. 7).
- [45] L. Van der Maaten and G. Hinton. "Visualizing data using t-SNE." In: *Journal of machine learning research* 9.11 (2008) (cit. on pp. 7, 25, 32).

- [46] A. Dubey and A. Choubey. "A systematic review on k-means clustering techniques". In: *Int J Sci Res Eng Technol (IJSRET, ISSN 2278-0882)* 6.6 (2017) (cit. on pp. 7, 8).
- [47] S. J. Nanda and G. Panda. "A survey on nature inspired metaheuristic algorithms for partitional clustering". In: *Swarm and Evolutionary Computation* 16 (2014), pp. 1–18. ISSN: 2210-6502. DOI: <https://doi.org/10.1016/j.swevo.2013.11.003>. URL: <https://www.sciencedirect.com/science/article/pii/S221065021300076X> (cit. on pp. 7, 8).
- [48] D. Xu and Y. Tian. "A comprehensive survey of clustering algorithms". In: *Annals of Data Science* 2 (2015), pp. 165–193 (cit. on pp. 7, 8).
- [49] F. Nielsen. "Hierarchical Clustering". In: *Introduction to HPC with MPI for Data Science*. Cham: Springer International Publishing, 2016, pp. 195–211. ISBN: 978-3-319-21903-5. DOI: [10.1007/978-3-319-21903-5_8](https://doi.org/10.1007/978-3-319-21903-5_8). URL: https://doi.org/10.1007/978-3-319-21903-5_8 (cit. on p. 7).
- [50] R. J. Campello et al. "Density-based clustering". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10.2 (2020), e1343 (cit. on p. 7).
- [51] D. M. SAPUTRA, D. SAPUTRA, and L. D. OSWARI. "Effect of distance metrics in determining k-value in k-means clustering using elbow and silhouette method". In: *Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*. Atlantis Press. 2020, pp. 341–346 (cit. on p. 8).
- [52] P.-N. Tan, M. Steinbach, and V. Kumar. "Data mining cluster analysis: basic concepts and algorithms". In: *Introduction to data mining* 487 (2013), p. 533 (cit. on p. 8).
- [53] D.-T. Dinh, T. Fujinami, and V.-N. Huynh. "Estimating the optimal number of clusters in categorical data clustering by silhouette coefficient". In: *Knowledge and Systems Sciences: 20th International Symposium, KSS 2019, Da Nang, Vietnam, November 29–December 1, 2019, Proceedings 20*. Springer. 2019, pp. 1–17 (cit. on p. 8).
- [54] M. F. Alhamid et al. "Towards context-sensitive collaborative media recommender system". In: *Multimedia Tools and Applications* 74 (2015), pp. 11399–11428 (cit. on pp. 8, 9).
- [55] A. Felfernig and R. Burke. "Constraint-Based Recommender Systems: Technologies and Research Issues". In: *Proceedings of the 10th International Conference on Electronic Commerce. ICEC '08*. Innsbruck, Austria: Association for Computing Machinery, 2008. ISBN: 9781605580753. DOI: [10.1145/1409540.1409544](https://doi.org/10.1145/1409540.1409544). URL: <https://doi.org/10.1145/1409540.1409544> (cit. on pp. 8, 9).
- [56] D. Nawara and R. Kashef. "Iot-based recommendation systems—an overview". In: *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. IEEE. 2020, pp. 1–7 (cit. on p. 8).

-
- [57] J. Huang et al. "A Novel Position-based VR Online Shopping Recommendation System based on Optimized Collaborative Filtering Algorithm". In: *arXiv preprint arXiv:2206.15021* (2022) (cit. on p. 9).
 - [58] R. De Croon et al. "Health recommender systems: systematic review". In: *Journal of Medical Internet Research* 23.6 (2021), e18035 (cit. on pp. 9, 15, 20).
 - [59] Y. Chen. "A music recommendation system based on collaborative filtering and SVD". In: *2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS)*. IEEE. 2022, pp. 1510–1513 (cit. on pp. 9, 21).
 - [60] S. Kant and T. Mahara. "Merging user and item based collaborative filtering to alleviate data sparsity". In: *International Journal of System Assurance Engineering and Management* 9 (2018), pp. 173–179 (cit. on p. 9).
 - [61] D. McSherry. "Explanation in recommender systems". In: *Artificial Intelligence Review* 24 (2005), pp. 179–197 (cit. on p. 9).
 - [62] G. Peake and J. Wang. "Explanation mining: Post hoc interpretability of latent factor models for recommendation systems". In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2018, pp. 2060–2069 (cit. on p. 10).
 - [63] Y. Zhang et al. "" Why Should You Trust My Explanation?" Understanding Uncertainty in LIME Explanations". In: *arXiv preprint arXiv:1904.12991* (2019) (cit. on p. 10).
 - [64] G. Van den Broeck et al. "On the tractability of SHAP explanations". In: *Journal of Artificial Intelligence Research* 74 (2022), pp. 851–886 (cit. on p. 10).
 - [65] V. Kuosmanen. "Improving the explainability of user-facing recommender systems". In: (2020) (cit. on p. 10).
 - [66] D. Shmaryahu, G. Shani, and B. Shapira. "Post-hoc Explanations for Complex Model Recommendations using Simple Methods." In: *IntRS@ RecSys*. 2020, pp. 26–36 (cit. on pp. 10, 17, 21, 29).
 - [67] M. Z. Al-Taie and S. Kadry. "Visualization of explanations in recommender systems". In: *Journal of Advanced Management Science Vol 2.2* (2014), pp. 140–144 (cit. on p. 10).
 - [68] P. Fränti and R. Mariescu-Istodor. "Soft precision and recall". In: *Pattern Recognition Letters* 167 (2023), pp. 115–121 (cit. on p. 11).
 - [69] B. Juba and H. S. Le. "Precision-recall versus accuracy and the role of large data sets". In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 33. 01. 2019, pp. 4039–4048 (cit. on p. 11).
 - [70] T. O. Hodson. "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not". In: *Geoscientific Model Development* 15.14 (2022), pp. 5481–5487 (cit. on p. 12).

- [71] B. Abdollahi and O. Nasraoui. "Explainable Matrix Factorization for Collaborative Filtering". In: *Proceedings of the 25th International Conference Companion on World Wide Web*. WWW '16 Companion. Montréal, Québec, Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 5–6. ISBN: 9781450341448. DOI: [10.1145/2872518.2889405](https://doi.org/10.1145/2872518.2889405). URL: <https://doi.org/10.1145/2872518.2889405> (cit. on pp. 12, 22).
- [72] M. S. Alshammari. "A Restaurant Recommendation Engine Using Feature-based Explainable Matrix Factorization". In: *2022 Fifth National Conference of Saudi Computers Colleges (NCCC)*. IEEE. 2022, pp. 26–30 (cit. on p. 13).
- [73] A. Bilbao, A. Almeida, and D. López-de-Ipiña. "Promotion of active ageing combining sensor and social network data". In: *Journal of Biomedical Informatics* 64 (2016), pp. 108–115. ISSN: 1532-0464. DOI: <https://doi.org/10.1016/j.jbi.2016.09.017>. URL: <https://www.sciencedirect.com/science/article/pii/S1532046416301307> (cit. on p. 14).
- [74] D. Bacciu et al. "Smart Environments and Context-Awareness for Lifestyle Management in a Healthy Active Ageing Framework". In: 2015-09. ISBN: 978-3-319-23484-7. DOI: [10.1007/978-3-319-23485-4_6](https://doi.org/10.1007/978-3-319-23485-4_6) (cit. on p. 14).
- [75] M. Caprara, R. Fernández-Ballesteros, and G. Alessandri. "Promoting aging well: evaluation of Vital-Aging-Multimedia Program in Madrid, Spain". In: *Health Promotion International* 31.3 (2015-03), pp. 515–522. ISSN: 0957-4824. DOI: [10.1093/heapro/dav014](https://doi.org/10.1093/heapro/dav014). eprint: <https://academic.oup.com/heapro/article-pdf/31/3/515/6751667/dav014.pdf>. URL: <https://doi.org/10.1093/heapro/dav014> (cit. on p. 14).
- [76] F. L. Vázquez et al. "Cognitive-behavioral intervention via interactive multimedia online video game for active aging: study protocol for a randomized controlled trial". In: *Trials* 20 (2019), pp. 1–15 (cit. on p. 15).
- [77] P. Kumar and A. Kumar. "A Review of Healthcare Recommendation Systems Using Several Categories of Filtering and Machine Learning-Based Methods". In: *2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*. IEEE. 2022, pp. 762–768 (cit. on p. 15).
- [78] K. Halder, M.-Y. Kan, and K. Sugiyama. "Health forum thread recommendation using an interest aware topic model". In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. 2017, pp. 1589–1598 (cit. on p. 15).
- [79] F. Narducci, P. Lops, and G. Semeraro. "Power to the patients: The HealthNetsocial network". In: *Information Systems* 71 (2017), pp. 111–122 (cit. on pp. 15, 20).
- [80] S. P. Erdeniz et al. "Recommender systems for IoT enabled m-health applications". In: *Artificial Intelligence Applications and Innovations: AIAI 2018 IFIP WG 12.5 International Workshops, SEDSEAL, 5G-PINE, MHDW, and HEALTHIOT, Rhodes, Greece, May 25-27, 2018, Proceedings 14*. Springer. 2018, pp. 227–237 (cit. on pp. 15, 20).

- [81] J. B. Daher, A. Brun, and A. Boyer. "A Review on Explanations in Recommender Systems". In: (2017-07). URL: <https://hal.science/hal-01836639> (cit. on p. 16).
- [82] J. L. Herlocker, J. A. Konstan, and J. Riedl. "Explaining Collaborative Filtering Recommendations". In: *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*. CSCW '00. Philadelphia, Pennsylvania, USA: Association for Computing Machinery, 2000, pp. 241–250. ISBN: 1581132220. DOI: [10.1145/358916.358995](https://doi.org/10.1145/358916.358995). URL: <https://doi.org/10.1145/358916.358995> (cit. on pp. 16, 22, 30).
- [83] N. Tintarev and J. Masthoff. "Evaluating the effectiveness of explanations for recommender systems: Methodological issues and empirical studies on the impact of personalization". In: *User Modeling and User-Adapted Interaction* 22 (2012), pp. 399–439 (cit. on p. 16).
- [84] "Pandas documentation". In: *pandas documentation - pandas 1.5.3 documentation* (). URL: <https://pandas.pydata.org/docs/> (cit. on p. 18).
- [85] "NumPy documentation". In: *NumPy documentation* (). URL: <https://numpy.org/doc/> (cit. on p. 18).
- [86] "Selecting the number of clusters with silhouette analysis on kmeans clustering". In: *scikit* (). URL: https://scikit-learn.org/stable/auto_examples/cluster/plot_kmeans_silhouette_analysis.html#sphx-glr-auto-examples-cluster-plot-kmeans-silhouette-analysis-py (cit. on pp. 18, 25).
- [87] "Welcome to surprise documentation!" In: *Welcome to Surprise' documentation! - Surprise 1 documentation* (). URL: <https://surprise.readthedocs.io/en/stable/> (cit. on p. 18).
- [88] C. A. Gomez-Urbe and N. Hunt. "The Netflix Recommender System: Algorithms, Business Value, and Innovation". In: *ACM Trans. Manage. Inf. Syst.* 6.4 (2016-12). ISSN: 2158-656X. DOI: [10.1145/2843948](https://doi.org/10.1145/2843948). URL: <https://doi.org/10.1145/2843948> (cit. on p. 19).
- [89] R. Bambini, P. Cremonesi, and R. Turrin. "A Recommender System for an IPTV Service Provider: a Real Large-Scale Production Environment". In: *Recommender Systems Handbook*. Ed. by F. Ricci et al. Boston, MA: Springer US, 2011, pp. 299–331. ISBN: 978-0-387-85820-3. DOI: [10.1007/978-0-387-85820-3_9](https://doi.org/10.1007/978-0-387-85820-3_9). URL: https://doi.org/10.1007/978-0-387-85820-3_9 (cit. on p. 19).
- [90] Y. Deldjoo et al. "Recommender Systems Leveraging Multimedia Content". In: *ACM Comput. Surv.* 53.5 (2020-09). ISSN: 0360-0300. DOI: [10.1145/3407190](https://doi.org/10.1145/3407190). URL: <https://doi.org/10.1145/3407190> (cit. on p. 20).
- [91] T. Tsunoda and M. Hoshino. "Automatic metadata expansion and indirect collaborative filtering for TV program recommendation system". In: *Multimedia Tools and applications* 36 (2008), pp. 37–54 (cit. on p. 20).

- [92] R. Banik. “Hands-on recommendation systems with Python: start building powerful and personalized, recommendation engines with Python”. In: Packt Publishing Ltd, 2018 (cit. on p. 21).
- [93] “CAEM/MediaMonitor - dados retirados da YUMI Analytics Desktop”. In: (2023) (cit. on pp. 23, 24).
- [94] “Tipos de Programas na TV”. In: *Estudos de Mercado, Audiências, Marketing Research, Media - Grupo Marktest - Estudos de Mercado, Audiências, Marketing Research, Media* (). URL: <https://www.marktest.com/wap/a/n/id~5c1.aspx> (cit. on p. 24).
- [95] M. Wattenberg, F. Viégas, and I. Johnson. “How to use T-Sne effectively”. In: *Distill* (2016-10). URL: <https://distill.pub/2016/misread-tsne/> (cit. on pp. 25, 32).



