Applications of user-based collaborative filtering in a business-to-business online marketing environment

Olga Vieru
SNR: 2008612

THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
DEPARTMENT OF COGNITIVE SCIENCE & ARTIFICIAL INTELLIGENCE
SCHOOL OF HUMANITIES AND DIGITAL SCIENCES
TILBURG UNIVERSITY

Thesis committee:

Dr. Drew Hendrickson
Dr. Paul Vogt

Tilburg University
School of Humanities and Digital Sciences
Department of Cognitive Science & Artificial Intelligence
Tilburg, The Netherlands
December 2019

Preface

First and foremost, I would like to thank everyone who made this project possible. A special thank you goes to my first supervisor at Tilburg University, Dr. Martijn van Otterlo, for setting a solid direction for this project, and for helping me narrow down my focus. Another big thank you goes to Dr. Drew Hendrickson for stepping in, and offering his guidance and support. And a huge thank you goes to my partner, Nick Pavlov, for his everlasting patience and invaluable help with any programming-related questions – your support means the world to me. Last but not the least, the biggest thank you of all goes to my parents – without whom none of this would have been possible. Once again, I am grateful for all your guidance as I was navigating this new and exciting field of artificial intelligence. Thank you all for your support, suggestions, words of encouragement, and for always being there for me.

Olga Vieru, December 2019

Applications of user-based collaborative filtering in a business-to-business online marketing environment

Olga Vieru

This research aims to explore whether we can predict and recommend information most relevant to individual users, based on similar users in the same demographic category. Existing user-based collaborative filtering algorithms are applied, in order to produce personalized content recommendations based on group navigation patterns (page clicks). As a result, memory-based (neighborhood) approaches and model-based (matrix factorization) techniques are tested out to compare performance and results. The final dataset used for modelling is encoded in a 25x45 user-item feedback rating matrix. Among several algorithms that are tried out, our tuned singular value decomposition (SVD) model has the best performance and accuracy with RMSE=0.24, MAE=0.16, compared to a chance-level performance of RMSE=1.14, MAE=0.80. Generated output includes ten most relevant URLs per user group, as well as five new predicted links, that users might find interesting. Several domain specific findings are discussed further on in this report. Moreover, an approach measuring user-item predicted interest is presented, in order to quantify each user group's preference for a URL, based on the deviation from their overall mean rating. In conclusion, this thesis contributes a low-resource data collection method with Google Analytics, which could be used to inform decision-making in both commercial and non-commercial settings, and translated to other domains.

1. Introduction

This research project was carried out in cooperation with eBenefits[1], a Netherlands-based business-to-business (B2B) financial software provider specialized in insurance, pension funds, and employee benefits software solutions. eBenefits' core product is the Compass Architecture SaaS (software as a service) model, which forms the basis of the software including calculation, management, and tender engines. ASR Nederland's[2] dashboard (mijn pensioenplan) was used as the primary case study, discussed in greater detail further on in this thesis (see section 3). The core goal was to customize ASR's dashboard to tailor content, personalizing and recommending options relevant to each individual user, based on behavioral group insights. Previously, users were unable to customize their content, apart from the net or gross numbers which could be displayed on their dashboard, as well as their relationship status.

Recommender systems are tools used to filter and suggest personalized items to various users or user groups, and have been researched extensively since they emerged as a domain in the early 1990s (Ricci, Rokach, and Shapira 2015). For instance, Netflix

1 https://www.ebenefits.nl/
2 https://www.asr.nl

provides individual content recommendations based on various features such as genre, similar users who watched the same movie, user sessions, behavioral patterns, even weather and events, among others. This leads to higher retention rates and overall engagement levels, lower cancellation rates, and better customer experience (Gomez-Uribe and Hunt 2016). On a higher level, this is applicable to our context as well, where users could benefit from more information and relevant content, better overall experience, and an easier interaction with the dashboard.

As for marketing applications, past academic research focuses mostly on model-based systems, and argues that complex models barely outperform simpler ones on small and medium-sized datasets (Wedel and Kannan 2016). Furthermore, demographic systems, which split users according to their demographic profiles and then generate recommendations, have been popular in the marketing literature but with insufficient research on these systems (Ricci, Rokach, and Shapira 2011). From a marketing perspective, this represents a business need since a personalization strategy would further enhance online marketing activities and efforts, ultimately driving engagement, and aiding decision-making. The employee market is continuously growing, however pension-related issues remain a low-interest aspect for most audiences, especially for younger target groups.

From a societal perspective, the Pensions Act enforced by the Dutch government requires that all employers and pension providers present employees with most important information relevant to them within five minutes (Randstad 2018). This presents a legal demand and crucial development from a legislative point of view, and thus needs to be addressed. As a result, this adds to the overall relevance of this project, whereby our society as a whole could benefit from more personalized information relevant to each individual user.

From a scientific point of view, there is an opportunity to explore personalization from the user perspective, setting the ground for further evaluation on whether that helps users find relevant information quicker. Previous research has been extensively focused on education, healthcare, and e-commerce, and thus, it would be interesting to explore personalization strategies in marketing as well. While more popular when it comes to practical applications, machine learning methods have not been thoroughly researched in marketing academia as they arguably do not 'produce generalizable theoretical insights' (Wedel and Kannan 2016). Lastly, Felfernig et al. (2013) highlight that few recommender systems to date focused on providing better information for users and not solely on revenue generation purposes – thus emphasizing potential for more strategies centered on user perspective-driven recommendations.

1.1 Research questions

Based on the information highlighted above, the following research question (RQ1) is suggested: *to what extent can we recommend the most relevant URLs for various user groups based on their behavioral patterns (clickthrough data)?* This research aims to identify ten most relevant website pages for each user group, which could then be used to customize ASR's dashboard and recommend relevant content. In order to add more depth to the research question, top five unvisited website pages will also be explored. It is unrealistic to expect that all user groups visited all website pages during a specific timeframe, and this may not signal a lack of interest or lower preference. It may also be the case that those pages are harder to find. In conclusion, both angles will be explored in order to contribute to the research question.

As an additional question, it might be interesting to compare memory-based and model-based approaches, focusing on the performance of these methods and models. This leads to the following question (RQ2): *how do memory-based approaches differ from model-based in terms of performance and results?* Cosine similarity, the most widely used memory-based/neighborhood approach in collaborative filtering, will be applied. Moreover, matrix factorization algorithms will be tested out, which are currently considered to be the most successful applications in the field since the Netflix challenge (Jorro-Aragoneses et al. 2019). Comparing these two most used methods in collaborative filtering against a basic prediction algorithm might produce interesting results, and add to the overall value of this project. Another motivation to apply both is that model-based methods are known to address sparsity issues better than memory-based systems. Next sections discuss in greater detail how the data was collected using Google Analytics, an alternative approach that enabled us to link the website navigation data to ASR's database and user groups, while ensuring user privacy at all times. The final dataset used for modeling and generating recommendations is encoded in a 25x45 user-item feedback rating matrix.

The structure of this thesis will be as follows: firstly, the theoretical framework will be presented, and this research will be placed in a broader scientific context. Secondly, the experimental setup will be discussed, how the data was pre-processed, along with the methods applied and models used. Furthermore, the results will be presented and discussed, followed by conclusions resulting from this research.

2. Related Work

This section lays out the theoretical framework and places this research in a broader context, discussing artificial intelligence (AI) applications in marketing, and narrowing it down to related concepts such as predictive modeling, recommender systems, and collaborative filtering techniques. Additionally, it explains how this work builds on related existing research in the field.

2.1 Applications of AI in marketing

In 2019, taking a data-driven approach in marketing when it comes to fueling decision-making processes is no longer a novelty. Yet, only five years ago IBM estimated that over 80% of traditional marketers relied solely on their gut feeling instead of scientific evidence to inform decision-making (Jacobson 2013). This also involved targeting and personalizing offers to deliver the right content to the right audiences (Sundsøy et al. 2014). However, a more recent industry study revealed several interesting findings, which might signal a shift in marketing practices. While 43% of surveyed marketers across nearly 200 business-to-consumer (B2C) companies use artificial intelligence applications for audience expansion and 39% for audience targeting, only 6% use more advanced applications – such as personalizing their content with collaborative filtering and predictive models (Blueshift 2018). Another research study powered by MIT Sloan and the Boston Consulting Group, which surveyed over 3,000 executives globally, showed that whilst almost 85% believe AI will help them achieve competitive advantage, only one in five incorporated AI in their processes and product offerings (Ransbotham et al. 2017).

Artificial intelligence can be defined as an area of computer science, 'concerned with how to give computers the sophistication to act intelligently, and to do so in increasingly wider realms' (Nilsson 2014). Since it was coined as a term in 1956, AI has seen considerable progress and reshaped many areas and industries, including healthcare, education, business, customer relationship management, and marketing. From search engines evolving from keywords to contextual search, to natural language processing developments such as IBM's Watson or Apple's Siri, which process natural language queries beyond textual input; to applications in marketing related to emotion and facial recognition, and real-time campaign optimization and dynamic pricing (Forrest and Hoanca 2015). Other applications of AI in B2B marketing specifically include predictive analysis, personalization, and lead scoring (Williams 2018). These applications aim to predict customers with the highest conversion rate, generate 'ideal customer profiles' and compare potential customers against those, as well as personalize offers and recommend relevant products often applying collaborative filtering algorithms (MarketingLand 2016).

2.2 Predictive analysis

Predictive analysis, one relevant application of AI in marketing, concerns exploring existing data to draw insights in order to predict future outcomes and trends (Verma et al. 2016), and it has become a powerful tool used in modern marketing (Stalidis, Karapistolis, and Vafeiadis 2015). SAS defines predictive analytics as 'the use of data, statistical algorithms and machine learning (ML) techniques to identify the likelihood of future outcomes based on historical data' (SAS 2018). Predictive analysis has been extensively used in consumer markets, whilst in B2B markets it has traditionally been used to drive customer acquisition (Leventhal 2018). Along with predictive analytics, adaptive personalization approaches represent state-of-the-art techniques in marketing (Wedel and Kannan 2016). Predictive models are sub-divided into: classification models, which predict belonging to a class, or discrete variables, and regression models – which predict a number, or continuous variables (Wakefield 2018). For instance, these models could help inform business decisions by predicting what content customers will choose to view, based on similar users' viewing history (Shakeel and Limcaco 2015).

2.3 Recommender systems

Recommender systems are information filtering tools that provide content suggestions (personalized or non-personalized) to users, based on various factors such as their interests, behavioral patterns, or overall product relevance. These recommendations might include, for instance, movies to watch on Netflix or videos on YouTube, music to listen to on Spotify, products to purchase on Amazon or select restaurants on Tripadvisor, for example. Recommender systems are based on either implicit or explicit ratings, which are later used to generate new suggestions in upcoming user-system interactions (Ricci, Rokach, and Shapira 2015).

Recommender systems are divided into several types, and Razmerita, Nabeth, and Kirchner (2012) discuss various interface personalization techniques, such as customization, agent-based personalization, and automatic personalization – the latter implying adaptive interfaces which rely on user characteristics to predict future behaviours. Gao, Liu, and Wu (2010) state that following user profiling and content modelling stage, interfaces can be personalized based on four filtering approaches: rule-based filtering, content-based filtering, collaborative filtering, and hybrid filtering. Masthoff (2011) discusses how group recommender approaches can be used to model individual recommendations, which is applicable to this research project as user group insights will be used to suggest content for individual users.

A categorization approach suggests filtering recommender systems as follows: demographic filtering, collaborative filtering, content-based, and hybrid filtering (Karabadji et al. 2018). Nilashi et al. (2016) discuss a hybrid recommender method based on clustering and regression techniques to improve the predictive accuracy of multi-criteria collaborative filtering. For example, a collaborative filtering system collects all information about users' activities on a web site, calculates similarity among the users, these are ultimately clustered in the same group, so that whenever a user logs in – they are placed in a group with similar users, and thus recommended relevant (and personalized) content for that particular user group.

Furthermore, most recent research focused on item-based recommender systems (such as rated items with clear scoring) or hybrid approaches, rather than user-based (Koohi and Kiani 2016). Therefore, there is potential for research within the B2B context based on user perspectives. As for online marketing theory, Wedel and Kannan (2016) discuss segment-level personalization, where groups of consumers with homogeneous preferences are identified, and the offering is personalized in the same way for all consumers in one segment. This approach was adapted and users were segmented based on their demographic characteristics.

2.4 Collaborative versus content-based filtering

Collaborative filtering (CF) is the most widely-used approach when it comes to recommender systems, and it makes content recommendations based on similar users in the same category (Sattar, Ghazanfar, and Iqbal 2017). Collaborative filtering algorithms rely on either explicit feedback (user ratings), or implicit feedback – signaling user preferences indirectly, such as for instance their browsing history, whether they watched a movie or not, or their purchase history (Koren and Bell 2015). Unlike content-based filtering, which generates recommendations based on similar items a user liked, collaborative filtering approaches are used to predict ratings based on similar users or groups of users (Yao et al. 2014). Here, opinions of other similar users, versus the content of items, are of utmost importance (Cacheda et al. 2011). Collaborative filtering is thought to be more accurate and flexible, compared to content-based filtering (Jorro-Aragoneses et al. 2019).

The relationship between users and items is subsequently encoded in a rating feedback matrix (See Table 1) (Wei et al. 2017), where each user U is matched with an item I and a rating R is specified per each item and user. Related research (Ekstrand et al. 2011) discusses two types of tasks associated with recommender systems, which will be explored in this research project: *predict* and *recommend*. These approaches are also referred to as *best item* and *top-N recommendation* problems (Desrosiers and Karypis 2011). The first task estimates, for each user-item pair, a user's preference for a new item, and the second produces a ranked list of recommendations to fit a user's need or a specific task.

2.5 Memory-based versus model-based approaches

Furthermore, depending on how the matrix is analyzed, two types of algorithms are considered: memory-based and model-based (Cacheda et al. 2011). Both of these approaches have been thoroughly researched, and depending on the purpose of the recommender system as well as the dataset, various studies argue whether one outperforms the other, or a combination of multiple approaches is the best way forward to build a

Table 1
An example user-item rating feedback matrix

	i1	**i2**	**i3**	**i4**	**i5**
u1	1	2	3	3	2
u2	1	2	1	2	1
u3	2	3	1	4	2
u4	4	1	4	1	3
u5	3	1	5	1	3

successful recommender system. Comparing these approaches will hopefully generate some domain-specific insights, building on previous related research in the field.

A memory-based collaborative filtering model calculates similarities between users (i.e. user-based approach) and/or items (i.e. item-based approach), and uses a weighted average of ratings to calculate and predict preference values (Ghazarian and Nematbakhsh 2015). Memory-based approaches, also called neighborhood-based techniques or methods, are widely applied for group recommender systems, and sometimes perform as well as more complex models. However, sparsity is an issue which occurs when data is insufficient for drawing meaningful conclusions and identifying relevant neighbors (Huang, Chen, and Zeng 2004). Common memory-based approaches used to calculate similarities between users and items are cosine similarity and Pearson's correlation coefficient.

Model-based systems, also called latent factor models, became the top choice for implementing collaborative filtering following the Netflix challenge (Ricci, Rokach, and Shapira 2015). These models are ahead of the memory-based approaches, as they factor in confidence levels and implicit feedback (Koren, Bell, and Volinsky 2009), and also address the sparsity issue better (Ghazarian and Nematbakhsh 2015; Dou, Yang, and Deng 2016). Algorithms widely applied in related collaborative filtering research are considered to be advanced and complex strategies, and include probabilistic matrix factorization (PMF), variations of SVD such as SVD++, principal component analysis (PCA), among others. However, a challenge often mentioned in related research on these models is domain specificity. That is, ratings are influenced by latent features highly specific to the domain, and these features which are directly learnt from rating data are not necessarily interpretable (Konstan 2019). This project aims to compare memory-based (neighborhood) approaches with model-based methods (matrix factorization), and dive into the predicted URLs as well as the recommendations these will generate. Since applications of these algorithms differ across domains, it will be interesting to see what domain-specific insights these will bring about.

3. Experimental Setup

The following section outlines the general approach of this research project, and discusses in detail how the data was collected, pre-processed, and analyzed.

The SQL database, which stores detailed employee user data (i.e. pension information, pension options, investment options, income projections, employment details), and various demographic information, constituted the starting data source at this stage. Domain knowledge was used to inform feature selection and demographic features were opted for. Demographic information is widely used in marketing to segment consumers and map their behaviors accordingly, and yet often this data is unavailable due to privacy considerations (Dong et al. 2014). Due to personal data privacy laws, Google's analytics reports prevent viewers from inferring demographic data about individual users (Google 2018), and offer limited demographic audience insights. As a default option, group demographic data cannot be mapped down to behavioral insights, unless specifically tracked. Therefore, by implementing a solution discussed further on in this thesis, it was possible to collect behavioral insights per groups of employee users and connect the database to Google Analytics.

3.1 Feature selection

The structure of the database was analyzed, relevant tables were explored, and features were discussed with eBenefits. Due to the complexity of the SQL database, privacy considerations, as well as the sensitive nature of the data, the final feature selection was conducted in consultation with eBenefits based on domain knowledge. Less relevant features (i.e. company names, individual financial projections) were left out from further analysis, as they were deemed less informative. SQL tables which were considered at this analysis stage included 'organization', 'intermediary', and 'metadata' tables among others. The 'organization' table listed all company related information, such as: company names, descriptions, addresses; the 'intermediary' table included all information related to eBenefits' intermediaries, ASR being one of them. Features included various intermediary data, settings, names, IDs, and other similar metadata. Since none of these features were relevant to the primary objective related to grouping employee users, the 'employee' relation table was deemed as the most relevant and informative, and thus, it was used to select features and group users for the next phase. All available features from the 'employee' relation table are described and presented in the Appendix 3.

Sensitive and less relevant information for user segmentation was left out, such as: BSN (social security) numbers, home addresses, bank account numbers, phone numbers, information about one's family members, and other personal information. Moreover, regions and cities were left out, as these could be inferred from Google Analytics as needed. Demographic filtering in recommender systems is centered around the idea that users sharing personal attributes (i.e. gender) will also share preferences (Bobadilla et al. 2013). Therefore, the following demographic features were used to group users, and are each explained below:

1. **User status** – active versus sleeper. Active users have monthly pension contributions set aside, while sleeper users no longer contribute to their own pensions. Sleeper users were still included since they are an important employee user group and have pensions built up, even though no longer actively contribute to them;
2. **Income** – the income feature represents the last known income range for a particular employee user, and groups were split as follows based on the available data: 0-20 thousand, 21-50 thousand, 51-100 thousand, 101-200 thousand, 201-500 thousand, 501 thousand-1 million and higher;
3. **Gender** – the gender feature includes two options, male and female (man, vrouw - in Dutch accordingly);

4. **Marital status** – this feature includes the following statuses as listed in the database: single, married, cohabitation (aleenstand, gehuwd, samenwonend - in Dutch accordingly). Since this feature defines some of the content presented to users (i.e. partner pension related information), it was crucial to include it as well;

5. **Age** – age groups were split based on Google Analytics demographic tracking data, as follows: 18-24, 25-34, 35-44, 45-54, 55-64, 65 and higher. These six age groups are fixed in Google Analytics, and thus, were used to draw age-related insights;

6. **Pension** – this feature includes available pension scheme types: gross, net, and combined, as defined by ASR. Gross pension represents a user's basic pension, whilst the net pension – reflects an income over 100 thousand euros per year. As of 1 January 2017, this amount has been indexed to 103,317 euros (ASR 2018a). Lastly, the combined pension represents both options. To illustrate the combined pension scheme, if an employee earns 150 thousand euros per year, the gross pension is calculated up until 103,317 thousand, and the remaining 46,683 thousand could then be saved via a net pension scheme;

7. **Investment profile** – this final selected feature represents a user's profile as defined by ASR: defensive, offensive, or neutral (voorzichtig, aanvallend, gemiddeld - in Dutch accordingly). A defensive profile implies pension money is invested in low-risk funds; an offensive profile means pension money is invested in high-risk funds, and finally a neutral profile represents both risk scenarios combined. Depending on the user's profile, a portfolio is created and money is invested in shares, bonds, real estate, or cash – and the risk is either balanced out across funds or not (ASR, n.d.). In order to illustrate this, investments for a defensive profile could be spread out as follows: shares (15-35%), government bonds (20-40%), corporate bonds (20-40%), real estate (0-15%), cash funds (0-20%) (ASR 2018b).

The features mentioned above were used to create 25 user groups, and further on were implemented in Google Tag Manager (GTM) for tracking and data collection purposes. The following sub-section discusses this process in greater detail. However, one limitation which should be mentioned here is the composition of the groups. These were not fully independent due to the nature of the data and inevitably we dealt with some overlap, as one user could belong to multiple groups. Therefore, group independence was not a property of this dataset.

3.2 Google Tag Management

Data was logged for employee users specifically as the core target group, and not for administrators or managers. This was due to the fact that the ASR portal offers a feature where administrators are able to log in as employee users (i.e. for testing purposes, or bug fixing). Therefore, those sessions were not logged as they did not reflect actual user behaviour, and would have distorted the data. Each user is assigned a role within the SQL database, and thus, only users with the role 'employee' were tracked. Firstly, tags were set up within the testing environment. Variables were defined based on the selected features from the SQL database, as discussed above. Triggering was implemented based on page views as the key configuration, and ASR as the intermediary and case study (i.e. from multiple eBenefits' clients). Next, each tag configuration was thoroughly explored, which included the following options: tracking type (event), and event tracking parameters such as category, action, label, and value.

Tags represented actions that were performed (fired), for instance a page view. Page views represented loaded pages that a user group viewed. Primary testing was implemented via the Google Manager plugin, which represented a preview mode to check whether the data layer was being recorded and updated correctly. Firstly, it was crucial to assess whether fired events were recorded correctly on Google Tag Manager (GTM). Secondly, it was important to assess whether the link between GTM and Google Analytics was performed by checking if events were reflected and visible in real-time. Following on from these initial layers of checks, changes were published within the testing environment in GTM. Following a successful initial trial run within the testing environment, values were exported and implemented within the production (live) environment. Another test was performed on the production environment, which allowed for the data to be collected for a period of four weeks.

Due to personal data privacy laws and considerations, Google Analytics offers limited audience insights (i.e. general demographic data, such as age, gender, geographical location). For ASR, the largest target group in March 2018 was comprised of people aged 45-54, which represented 27% of all sessions. Out of these people 41.9% were male and 58.1% were female, their average session lasted 4 minutes 34 seconds, and they explored roughly six pages per session. This might be considered a positive signal, implying that these individuals were interested to learn about their pension options, or it could signal that relevant information was hard to find – which would explain why they had to navigate up to six pages per session. Therefore, it was hypothesized that information most relevant to users may not have been easily accessible via the dashboard. As a result, it was decided to select and recommend ten most relevant website pages, as well as predict five new pages which user groups might find equally interesting. For the purposes of this research and modeling, user groups are seen as 'individual' users.

3.3 Data

Data was tracked and collected between 25 April 2018 and 30 June 2018. In total, the unprocessed exported dataset (see Appendix 5 for a partial snapshot) included 6358 logged group events (actions). Following pre-processing, the final dataset included 1124 group actions. This dataset included features and relevant selected metrics, as follows: event category, event action, event label, users, and page views. These features are explained in more detail below:

- **Event category –** categories assigned to triggered events;
- **Event action –** actions assigned to triggered events;
- **Event label –** labels used to describe triggered events (URLs);
- **Users –** users who initiated at least one session during the date range;
- **Page views –** total number of pages viewed (repeated views of a single page counted as well);

The final exported data was anonymized (any company names and user IDs were excluded for privacy considerations) and grouped. Therefore, our data could not have been tracked down to individual users, or linked to individuals within the SQL database. In total, 12416 unique users logged onto the dashboard during the specified date range: 8995 male users and 3851 female users, as can be seen in Figure 1. Idle users with a single session of over thirty minutes were considered to be outliers, and thus, excluded from further analysis. These users represented less than 1% of the entire dataset, but considerably skewed their own groups (in instances where they were the single user within their group).

3.4 Pre-processing

Data was pre-processed in Python and R, and subsequently analyzed using PyCharm (version 2019.2.4) and RStudio (version 1.1.383). The final dataset was exported as a comma-separated values (CSV) file for analysis and pre-processing. As previously mentioned, administrators and managers were excluded from the dataset. Upon initial analysis, it was noted that several event labels had two different referrals to the same URL, one with *https* and one with *www*, some included capitals and some – lowercase letters, some included double slashes versus a single slash, and some pointed to sub-domains versus main domain name (asr.nl) – all these issues with the data had to be addressed.

For example, *https://www.mijnwerknemerspensioen.asr.nl/Employee/dashboard.aspx* and *https://mijnwerknemerspensioen.asr.nl/Employee/dashboard.aspx* both point to the same address, the only difference is that one has the *www* prefix, and the other one does not. Python packages used for preprocessing the data were pandas (McKinney et al. 2010), numpy (Oliphant 2006), and sklearn (Pedregosa et al. 2011). The primary objective was leaving in the pathname, and removing any attributes resulting in multiple URLs. Firstly, any URLs which included lowercase and uppercase characters had to be formatted and merged. For example, the following two URLs pointed to the same end location, but were treated as two separate entities by Google (i.e. mutation overview versus Mutation Overview):

https://mijnwerknemerspensioen.asr.nl/Mutations/Employee/mutationoverview.aspx
https://mijnwerknemerspensioen.asr.nl/Mutations/Employee/MutationOverview.aspx

Secondly, as previously mentioned, *https* and *www* had to be removed as well, to ensure more accurate data and less duplication. Moreover, any trailing spaces were removed as well, such as referrals from various sources (i.e. email). In addition, whenever the same URL had double slashes instead of a single slash – both referring to the same address – had to be streamlined as well. Furthermore, comma-decimal separators had to be replaced with semicolons. As some links contained commas, it was important to ensure that multiple imports and exports of the CSV file had no impact on the data. Finally, the domain and sub-domain issue had to be addressed as well, again to ensure no duplication within the data.

All employee users have individual documents (such as financial statements), each with their own unique identifiers (IDs). For the purposes of this research, it was less relevant to look at these documents since they are individual user-based, and more appropriate to view them as a general category that users access (group-based). Therefore, options such as mutation detail, mutation get/view document, and sub-mutation detail had to be merged, since they all had individual IDs assigned to them. These identifiers reference specific mutations or documents, which are irrelevant to be explored as individual entities.

/mutations/employee/MutationDetail.aspx?ID=45e1ea71-2947-42a3-951c-569426908f99
(?ID = removed)

In the context of this data, mutations represent change requests submitted to the system. Since employee users are unable to perform these changes on their dashboards, they submit requests to have those updates implemented. These requests include the following types: mutations such as an address change, relationship status change, pension type change, and sub-mutations which represent the mutation type, or how the action was carried out (i.e. means of delivery).

The dashboard link (employee/dashboard.aspx) was excluded from the analysis, and used as a reference point for the number of users who logged in during the specified time period, as all traffic considered was directed via the dashboard (i.e. referrals were excluded). Following the pre-processing discussed above, the final dataset included 45 unique links. These URLs were assigned IDs to facilitate further analysis. In order to do this, URLs were re-coded as follows: 1 = main page or dashboard, 2 = any links from the dashboard, 3 = any remaining links (i.e. URLs clicked via top navigation). Moreover, their relationships to the main page were analyzed, as well as the navigation and page hierarchy (i.e. main page or subpage). The full overview is presented in Appendix 1.

As a final step in pre-processing, page clicks were normalized per group based on the percentage of visits from the dashboard as the entry point, and converted to ratings on a 0-5 scale accordingly – a commonly used rating scale in recommender systems. Usually, recommender systems work with such scales, binary (like versus dislike) scales, or unary data (Ekstrand et al. 2011).

3.5 Method

Memory-based methods

In order to answer our research questions, a memory-based method was explored. The following steps were undertaken as discussed extensively in related research: firstly, similarities between users were computed, and secondly – weighted rating averages were calculated to estimate user preference values (Ghazarian and Nematbakhsh 2015). Adjusted cosine similarity (see formula 1) was used versus regular cosine similarity, in order to account for the considerable difference in the ratings between various user groups. As a result, instead of a user group's 'raw' rating, similarity was measured based on the difference between a user group's rating for an item *i*, and their average rating across all items (Kane 2019). In the formula below, $\bar{x}$ denotes the average of all user x's ratings, and $\bar{y}$ stands for all of user y's ratings. Therefore, the main difference here from conventional cosine similarity is the variance from the mean of each user group's ratings, and not the just the ratings themselves.

$$\text{CosSim}(x,y) = \frac{\sum_i \left(\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)\right)}{\sqrt{\sum_i \left(x_i - \bar{x}\right)^2}\sqrt{\sum_i \left(y_i - \bar{y}\right)^2}} \tag{1}$$

Firstly, both URL and user rating mean scores were calculated. Furthermore, the difference between user group ratings and averages was computed. Moreover, NaNs were replaced by user group average ratings, and ultimately user group similarity was calculated. Replacing NaNs is a method widely used in related research to address the sparsity issue and improve recommender performance, since not all user groups visited all URLs. Related research (Dou, Yang, and Deng 2016) discusses various strategies to deal with that: setting NaNs to a default value, but this would arguably lead to a low reliability; and another is to use rating averages between URLs and users, in our scenario. Matrix factorization algorithms are known to address the sparsity issue better, and will therefore be explored further on in this research project. As a result, two similarity matrices were created: between URLs and between users. Finally, in order to predict how much a user group would like a particular item, the below formula was applied, where s is the predicted score of a user u and item i pair, r is user rating. Cosine similarity was applied to calculate the weight in the formula 2 below, and the score was equal to all the ratings each user gave to that specific item minus their average rating, multiplied by the weight denoting how much this user is similar to another user (Konstan 2019).

$$s(u,i) = \bar{r}_u + \frac{\sum_{v \in V} (r_{vi} - \bar{r}_v) * w_{uv}}{\sum_{v \in V} w_{uv}} \tag{2}$$

Moreover, we could identify nearest neighbors using the similarity matrix discussed above, and recommend content based on similar users within a neighborhood. An example will be provided further on in the results section. Finally, our baseline consisted of mean and median ratings per user group (Appendix 4), so that we could estimate the difference between average ratings and predicted preference levels.

Model-based methods

In order to address our research questions, various model-based strategies were tested out, to explore which one would have the best performance with our data. Given the current popularity within the field as discussed in the theoretical framework, several SVD implementations were explored as well as PMF – and these were compared to our baseline which was a chance-level performance (random recommendations). As for evaluation, two commonly used accuracy measures (Desrosiers and Karypis 2011) were applied, in order to evaluate the performance of our recommender: RMSE (root mean squared error) and MAE (mean absolute error). For evaluation purposes, our ratings dataset containing 1124 group actions was split into 80% training and 20% test set. Furthermore, hyperparameter tuning was performed. With SVD, for instance, we were interested to see how many latent factors, or dimensions (n factors in the surprise library), should ideally be extracted. Moreover, the learning rate and how many epochs (or steps) the algorithm should take, were both adjusted. To sum up, the GridSearchCV package from the surprise library[3] was used to tune and find the best parameters for the learning rate, latent factors or dimensions, and number of epochs, and a three-fold cross-validation was applied on the training dataset each time. These are illustrated in table 3.

3 https://surprise.readthedocs.io

4. Results

In this section, several interesting research findings will be presented to illustrate recommendations that were generated for two user groups. Moreover, model performance will be discussed, as well as evaluation metrics and baselines that were used to compare model accuracy and performance. Following on from the steps described in the previous section, the relationship between user groups and unique URLs was encoded in a 25x45 rating feedback matrix, where each unique URL represented an instance on the X axis, and each user group was represented as rows with normalized ratings on the Y axis. For illustration purposes, findings from two contrasting age groups are presented below: group 11 – users aged 55-64, and group 21 – users aged 18-24. Below are shown output examples to ensure a better understanding of the results.

Examples output

```
Output (k=5)
UserID1: 2,4,3,6,5
UserID2: 1,4,3,6,5
UserID3: 6,2,1,9,8
UserID4: 5,1,2,10,3
UserID5: 4,2,1,10,3
```

The output above shows five nearest neighbors for user groups with ID1 through 5. The output below shows the top ten recommended URLs for user 11, as well as top five predicted new URLs which the user group might find interesting.

```
URLs recommended for User 11
/employee/documents.aspx
/employee/notification/notifications.aspx
/employee/cms.aspx?type=myscheme
/employee/customerjourney/presentvalue2.aspx
/employee/pension/pensionchoices.aspx
/mutations/employee/mutationoverview.aspx
/employee/customerjourney/yourdata.aspx
/employee/customerjourney/personaldata.aspx
/profile.aspx
/employee/cms.aspx?type=dashboard-mobile
```

```
URLs predicted for User 11
/mutations/employee/nettoparticipationscheme.aspx
/employee/klantreis.aspx?type=overlijden
/employee/klantreis.aspx?type=nettowerknemerspensioen
/employee/klantreis.aspx?type=scheiden
/mutations/employee/endnettoparticipationscheme.aspx
```

As can be seen from the output presented above, the top ten URL recommendations computed based on adjusted cosine similarity, for user group with ID11 (users aged 55-64), are as follows: documents folder, messages, pension scheme overview, and general pension data, among other links. Interestingly, some of these pages are sub-pages hierarchically, and thus are not present in the top navigation bar. Mobile dashboard page received the lowest ranking among the top ten s(u,i)=0.58, with the average user rating being s(u,i)=0.65. The top five predicted URLs which this user group might find interesting include pages about divorce, death, and on ending their pension net settlement.

```
URL recommendations for User 21
/employee/documents.aspx
/employee/notification/notifications.aspx
/employee/cms.aspx?type=myscheme
/employee/cms.aspx?type=dashboard-mobile
/employee/customerjourney/presentvalue2.aspx
/employee/pension/pensionchoices.aspx
/profile.aspx
/employee/customerjourney/personaldata.aspx
/mutations/employee/mutationoverview.aspx
/employee/customerjourney/yourdata.aspx

URL suggestions for User 21
/employee/klantreis.aspx?type=variabel-pensioen
/mutations/employee/nettoparticipationscheme.aspx
/employee/klantreis.aspx?type=nettowerknemerspensioen
/employee/klantreis.aspx?type=scheiden
/mutations/employee/endnettoparticipationscheme.aspx
```

As can be seen above, the mobile version of the dashboard was ranked higher up for the user group aged 18-24 (ID21), with the score s(u,i)=1.47 and average rating 0.85, which signals the importance of the mobile site for younger user groups. Suggested links also included the ‘scheiden/divorce’ URL. Our tuned SVD model predicted pages including information on ending cohabitation s(u,i)=0.46, starting s(u,i)=0.48 and ending s(u,i)=0.58 a net participation scheme. A table providing mean and median URL ratings per user group can be found in the Appendix 4, along with reference user IDs (Appendix 2). This data, along with the top ten most popular links overall for all users, was used as the baseline measure to compare ‘individual’ recommendations per group for the memory-based neighborhood method. This naïve approach is referenced in previous research as ‘surprisingly powerful’, as most users focus on the few of the many available items – and this is therefore considered a suitable baseline value (Hu, Koren, and Volinsky 2008). As for model-based approaches, results below (see Table 2) showcase the algorithms that were tried out with our dataset. The best performing model was the tuned SVD (RMSE=0.24, MAE=0.16) with SVD++ also performing rather well (RMSE=0.28, MAE=0.18), compared to a chance-level performance.

Table 2
An overview of the algorithms including chance-level performance

Algorithm	RMSE	MAE
SVD	0.40	0.27
SVD [Tuned]	**0.24**	**0.16**
SVD++	0.28	0.18
PMF	0.95	0.58
PMF [Tuned]	0.56	0.31
Random	1.14	0.80

A basic algorithm[4] (formula 3 below) was used as a baseline, predicting random ratings based on the distribution of the training dataset. The formula below was applied: where R_{train} is the training set, r_{ui} is the true rating of user *u* for item *i*, and where $\hat{\mu}$ and $\hat{\sigma}$ are estimated from the training data using Maximum Likelihood Estimation:

$$\hat{\mu} = \frac{1}{|R_{train}|} \sum_{r_{ui} \in R_{train}} r_{ui}$$
$$\hat{\sigma} = \sqrt{\sum_{r_{ui} \in R_{train}} \frac{(r_{ui} - \hat{\mu})^2}{|R_{train}|}} \tag{3}$$

As can be seen in table 3 below, the models were tuned and best hyperparameters were found with GridSearchCV, and a three-fold cross-validation was applied on the training dataset each time.

5. Discussion

The goal of this research was to predict and recommend information most relevant to individual employee users, based on group navigation patterns (clicks). More specifically, this study aimed to apply existing user-based collaborative filtering algorithms, in order to personalize dashboard content based on similar users in the same category. An alternative data collection approach with Google Analytics was suggested, which could be applied to similar contexts and across various domains, as it is a rather low-resource implementation compared to other strategies (i.e. such as implementing an explicit feedback solution). Moreover, two collaborative filtering approaches were explored:

4 https://surprise.readthedocs.io/en/stable/prediction_algorithms_package.html

Table 3
An overview of the tuned hyperparameters and best values (GridSearchCV)

SVD model parameters, CV=3	Tuning values	Best values
Learning rate	0.005, 0.010	0.01
N_epochs	20, 30	30
N_factors	5, 50	5

PMF model parameters, CV=3	Tuning values	Best values
Learning rate	0.005, 0.010	0.01
N_epochs	5, 10	10
N_factors	10, 20	20

memory-based (cosine similarity) and model-based techniques (matrix factorization), generating several interesting findings – both from an online marketing point of view, as well as from business and societal standpoints. Besides the low-resource data collection method, this study also contributes a strategy to use group insights to model individual recommendations, using demographic features which have been popular but insufficiently researched in the marketing literature (Ricci, Rokach, and Shapira 2015).

On a higher level, there was little variance among the top ten most popular pages among all user groups. For instance, the mobile dashboard was understandably ranked higher among younger user groups, elderly people relied on notifications more, and several sub-pages received considerable attention despite being scattered throughout the website. The memory-based method included a custom function which assigned scores for each user-item pair, so these could be compared to mean ratings. Thus, we managed to quantify user preferences. As for model-based techniques, several algorithms were tested out and the tuned SVD had the best performance and accuracy with RMSE=0.24, MAE=0.16, against a chance-level performance of RMSE=1.14, MAE=0.80. However, predicting five new pages for each user group revealed surprising findings, as married people were recommended the 'divorce' page, and elderly people – the 'information about death' page. Other users from similar groups visited those pages, so that might be a reason why they were recommended to those users who have not seen them yet.

As for limitations of the data, an overlap between user groups should be noted. Due to privacy laws and considerations, as well as the sensitive nature of the data discussed earlier in this report, grouped data was used for the purposes of this study. As a result, these user groups were not completely independent of each other, i.e. a user could belong to both 'gender – male' and 'marital status – married' groups, but not to multiple age or income groups simultaneously. Future research could address this issue to further enhance dashboard optimization and individual user experience. Moreover, future research opportunities could include considering more behavioral metrics, aggregating various types of feedback, and adding to the complexity of the recommender system. Metrics such as pages per session, or time spent on page, could also be considered to dig deeper into user navigation patterns. Recommender systems improve as more data is collected, so perhaps building a larger dataset with a better representation of all user groups, could be explored.

Furthermore, implicit data is considered to be noisy in recommenders (Hu, Koren, and Volinsky 2008), so we could only estimate user preferences based on their behavioral patterns. Several pages were visited often whilst other received hardly any hits, but it is important to note that not all users had the chance to interact with all items – a typical real-life scenario. Therefore, the first research question was two-fold: recommending top ten most relevant URLs for each user group, as well as predicting five new pages which might be equally interesting. Recommenders do not always get these predictions right, and even though the 'divorce' page might be relevant to married users, or the 'information about death' page to elderly users, it is unlikely that recommending them would work well in a real-life setting.

Finally, further evaluation and A/B testing could potentially be explored to measure how these changes would affect user engagement and satisfaction levels. It is challenging to estimate how well a recommender system would work without an in-depth user study or a survey (Hu, Koren, and Volinsky 2008). With no standardized way to evaluate all the different approaches in the field, this 'user-centered holistic evaluation' might work well but it is only applied by few organizations (Ekstrand et al. 2011).

6. Conclusion

This thesis contributes a rather low-resource data collection method, which could be used to inform decision-making in both commercial and non-commercial settings. Moreover, this project shows how group insights could be used to model recommendations relevant to individual users, converting implicit feedback into 'ratings' to indicate user preferences. This project compared widely used memory-based (neighborhood) approaches with more advanced and complex model-based techniques (matrix factorization), in order to see if any domain specific insights would arise in the process. As a result, it was possible to answer our research questions by recommending the most relevant URLs for various user groups based on their navigation patterns (clickthrough data). Moreover, top five new pages were predicted per user group, which could potentially be interesting as well. Finally, we compared the results and performance of several matrix factorization algorithms. Our tuned SVD model had the best performance and accuracy with RMSE=0.24, MAE=0.16, against a chance-level performance of RMSE=1.14, MAE=0.80.

As there is no 'one size fits all' approach in recommender systems, their performance is highly dependent on multiple factors – such as input data and specific domains. It is likely that some methods and strategies used in this study could be translated to other domains, business or research settings. Future research could address the group independence issue, consider additional behavioral metrics or aggregate several types of feedback - both implicit and explicit - as user preferences and performance may diverge (Herlocker et al. 2004). Furthermore, user-centered evaluation strategies could be explored such as A/B testing, in-depth user studies, or surveys. Other researchers state that there is still work to be done on the recommender experience as a whole, from the data collection to user experience (Ekstrand et al. 2011), and so future research could take many directions on the basis of the contribution of this study.

References

ASR. 2018a. Netto werknemers pensioen. Retrieved from https://www.asr.nl/.

ASR. 2018b. Onze beleggingsprofielen. Retrieved from https://www.asr.nl/.

Blueshift. 2018. Activating customer data for ai powered marketing.

Bobadilla, Jesús, Fernando Ortega, Antonio Hernando, and Abraham Gutiérrez. 2013. Recommender systems survey. *Knowledge-based systems*, 46:109–132.

Cacheda, Fidel, Víctor Carneiro, Diego Fernández, and Vreixo Formoso. 2011. Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems. *ACM Transactions on the Web (TWEB)*, 5(1):2.

Desrosiers, Christian and George Karypis. 2011. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook*. Springer, pages 107–144.

Dong, Yuxiao, Yang Yang, Jie Tang, Yang Yang, and Nitesh V Chawla. 2014. Inferring user demographics and social strategies in mobile social networks. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 15–24, ACM.

Dou, Yingtong, Hao Yang, and Xiaolong Deng. 2016. A survey of collaborative filtering algorithms for social recommender systems. In *2016 12th International Conference on Semantics, Knowledge and Grids (SKG)*, pages 40–46, IEEE.

Ekstrand, Michael D, John T Riedl, Joseph A Konstan, et al. 2011. Collaborative filtering recommender systems. *Foundations and Trends® in Human–Computer Interaction*, 4(2):81–173.

Felfernig, Alexander, Michael Jeran, Gerald Ninaus, Florian Reinfrank, and Stefan Reiterer. 2013. Toward the next generation of recommender systems: applications and research challenges. In *Multimedia services in intelligent environments*. Springer, pages 81–98.

Forrest, Edward and Bogdan Hoanca. 2015. Artificial intelligence: Marketing's game changer. In *Trends and innovations in marketing information systems*. IGI Global, pages 45–64.

Gao, Min, Kecheng Liu, and Zhongfu Wu. 2010. Personalisation in web computing and informatics: Theories, techniques, applications, and future research. *Information Systems Frontiers*, 12(5):607–629.

Ghazarian, Sarik and Mohammad Ali Nematbakhsh. 2015. Enhancing memory-based collaborative filtering for group recommender systems. *Expert systems with applications*, 42(7):3801–3812.

Gomez-Uribe, Carlos A and Neil Hunt. 2016. The netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems (TMIS)*, 6(4):13.

Google, Analytics. 2018. About demographics and interests. Retrieved from https://support.google.com/analytics/answer/2799357?hl=en.

Herlocker, Jonathan L, Joseph A Konstan, Loren G Terveen, and John T Riedl. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53.

Hu, Yifan, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *2008 Eighth IEEE International Conference on Data Mining*, pages 263–272, Ieee.

Huang, Zan, Hsinchun Chen, and Daniel Zeng. 2004. Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering. *ACM Transactions on Information Systems (TOIS)*, 22(1):116–142.

Jacobson, Ralph. 2013. Ibm industry insights: 80% of cpg & retail marketers still make decisions based on gut feel. Retrieved from: https://www.ibm.com/blogs/insights-on-business/consumer-products/80-of-cpg-retail-marketers-still-make-decisions-based-on-gut-feel/.

Jorro-Aragoneses, Jose, Marta Caro-Martinez, Juan Antonio Recio-Garcia, Belen Diaz-Agudo, and Guillermo Jimenez-Diaz. 2019. Personalized case-based explanation of matrix factorization recommendations. In *International Conference on Case-Based Reasoning*, pages 140–154, Springer.

Kane, Frank. 2019. Building recommender systems with machine learning and ai. Retrieved from: https://www.skillshare.com/classes/Building-Recommender-Systems-with-Machine-Learning-and-AI/.

Karabadji, Nour El Islem, Samia Beldjoudi, Hassina Seridi, Sabeur Aridhi, and Wajdi Dhifli. 2018. Improving memory-based user collaborative filtering with evolutionary multi-objective optimization. *Expert Systems with Applications*, 98:153–165.

Konstan, Michael D. Ekstrand & Joseph A. 2019. Recommender systems specialization. University of Minnesota. Retrieved from https://www.coursera.org/specializations/recommender-systems.
Koohi, Hamidreza and Kourosh Kiani. 2016. User based collaborative filtering using fuzzy c-means. *Measurement*, 91:134–139.
Koren, Yehuda and Robert Bell. 2015. Advances in collaborative filtering. In *Recommender systems handbook*. Springer, pages 77–118.
Koren, Yehuda, Robert Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37.
Leventhal, Barry. 2018. *Predictive Analytics for Marketers: Using Data Mining for Business Advantage*. Kogan Page Publishers.
MarketingLand. 2016. Martech intelligence report: B2b predictive marketing analytics platforms: A marketer's guide.
Masthoff, Judith. 2011. Group recommender systems: Combining individual models. In *Recommender systems handbook*. Springer, pages 677–702.
McKinney, Wes et al. 2010. Data structures for statistical computing in python. In *Proceedings of the 9th Python in Science Conference*, volume 445, pages 51–56, Austin, TX.
Nilashi, Mehrbakhsh, Mohammad Dalvi Esfahani, Morteza Zamani Roudbaraki, T Ramayah, and Othman Ibrahim. 2016. A multi-criteria collaborative filtering recommender system using clustering and regression techniques. *Journal of Soft Computing and Decision Support Systems*, 3(5):24–30.
Nilsson, Nils J. 2014. *Principles of artificial intelligence*. Morgan Kaufmann.
Oliphant, Travis E. 2006. *A guide to NumPy*, volume 1. Trelgol Publishing USA.
Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in python. *Journal of machine learning research*, 12(Oct):2825–2830.
Randstad. 2018. Pensioenvoorlichting. Retrieved from https://werkpocket.randstad.nl/wet-en-regelgeving/pensioen/pensioenvoorlichting.
Ransbotham, Sam, David Kiron, Philipp Gerbert, and Martin Reeves. 2017. Reshaping business with artificial intelligence: Closing the gap between ambition and action. *MIT Sloan Management Review*, 59(1).
Razmerita, Liana, Thierry Nabeth, and Kathrin Kirchner. 2012. User modeling and attention support: Towards a framework of personalization techniques. In *The fifth international conference on advances in human-oriented and personalized mechanisms, technologies, and services*, pages 27–33.
Ricci, Francesco, Lior Rokach, and Bracha Shapira. 2011. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer, pages 1–35.
Ricci, Francesco, Lior Rokach, and Bracha Shapira. 2015. Recommender systems: introduction and challenges. In *Recommender systems handbook*. Springer, pages 1–34.
SAS. 2018. Predictive analytics. sas insights. Retrieved from https://www.sas.com/en-us/insights/analytics/predictive-analytics.html.
Sattar, Asma, Mustansar Ali Ghazanfar, and Misbah Iqbal. 2017. Building accurate and practical recommender system algorithms using machine learning classifier and collaborative filtering. *Arabian Journal for Science and Engineering*, 42(8):3229–3247.
Shakeel, Usman and Mike Limcaco. 2015. Leveraging cloud-based predictive analytics to strengthen audience engagement.
Stalidis, George, Dimitrios Karapistolis, and Athanasios Vafeiadis. 2015. Marketing decision support using artificial intelligence and knowledge modeling: application to tourist destination management. *Procedia-Social and Behavioral Sciences*, 175:106–113.
Sundsøy, Pål, Johannes Bjelland, Asif M Iqbal, Yves-Alexandre de Montjoye, et al. 2014. Big data-driven marketing: how machine learning outperforms marketers' gut-feeling. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, pages 367–374, Springer.
Verma, Jai Prakash, Smita Agrawal, Bankim Patel, and Atul Patel. 2016. Big data analytics: Challenges and applications for text, audio, video, and social media data. *International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI)*, 5(1).
Wakefield, Katrina. 2018. Sas insights: Predictive analytics and machine learning.

Wedel, Michel and PK Kannan. 2016. Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6):97–121.

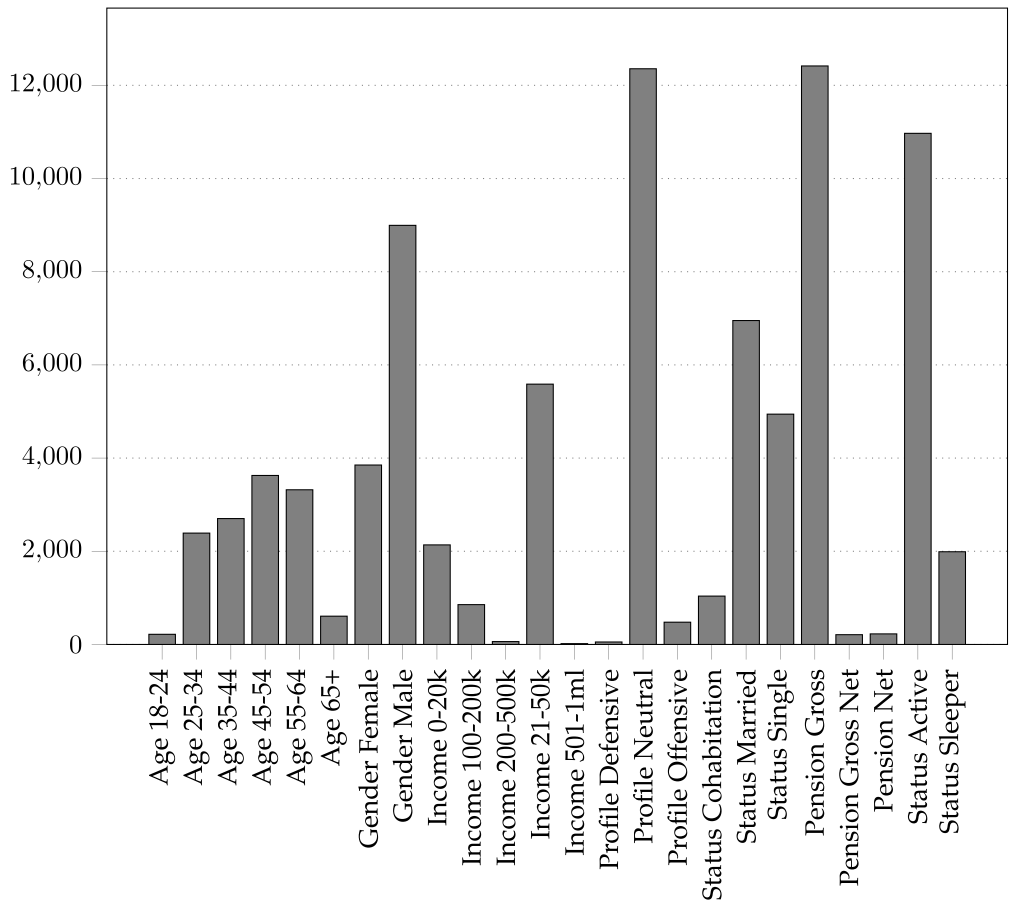
Wei, Jian, Jianhua He, Kai Chen, Yi Zhou, and Zuoyin Tang. 2017. Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69:29–39.

Williams, Amahl. 2018. Applications of artificial intelligence in b2b marketing. Retrieved from https://www.forbes.com/sites/forbescommunicationscouncil/2018/01/05/applications-of-artificial-intelligence-in-b2b-marketing/2f40cce07150.

Yao, Lina, Quan Z Sheng, Anne HH Ngu, Jian Yu, and Aviv Segev. 2014. Unified collaborative and content-based web service recommendation. *IEEE Transactions on Services Computing*, 8(3):453–466.

Appendix A: Graphs

Figure 1
Histogram illustrating the number of users per group (dashboard)



Appendix B: Tables

Table 1
Overview of links and descriptions

Short URL	Category	Type	Navigation (top)	Page type	Parent
dashboard	1	Dashboard	Yes	Main	N/A
documents	3	Documents folder	Yes	Main	N/A
notification/notifications	3	Messages	Yes	Main	N/A

Short URL	Category	Type	Navigation (top)	Page type	Parent
cms?type=myscheme	3	Scheme overview	Yes	Main	N/A
presentvalue2	3	General pension data 2	No	Sub	Myscheme
cms?type=dashboard_mobile	1	Mobile dashboard	Yes	Main	N/A
profile	3	Profile settings	Yes	Main	N/A
personaldata	3	Personal information	No	Sub	Myscheme
yourdata	3	Employment related pension data	No	Sub	Myscheme
mutation/mutationoverview	3	Overview changes submitted	No	Sub	Myscheme
survivors	3	Data about relatives	No	Sub	Myscheme
pension/pensionchoices	2	Submit changes options	No	Sub	Myscheme
costs	3	Premiums and costs	No	Sub	Myscheme
contact	3	Contact us	Yes	Main	N/A
klantreis?type=nieuwbijasr	2	Welcome information	No	Main	N/A
klantreis?type=pensioenregeling2018	2	Information page 2018	No	Main	N/A
mutation/contactinformation	3	Contact details change	No	Sub	Submit changes options
cms?type=my_scheme_mobile	3	Mobile scheme	Yes	Main	N/A
klantreis?type=variabel_pensioen	2	Variable and fixed pension	No	Main	N/A
presentvalue	3	General pension data 1	No	Sub	Myscheme

Short URL	Category	Type	Navigation (top)	Page type	Parent
faq	3	Frequently asked questions	Yes	Main	N/A
mutation/cohabitation/start	3	Submit changes options - cohabitation	No	Sub	Submit changes op-tions
mutation/sanw	3	Voluntary ANW pension	No	Sub	Submit changes op-tions
contactform	3	Contact form	No	Sub	Contact us
klantreis?type=gegarandeerdpensioen	2	Investing or guaranteed pension	No	Main	N/A
klantreis?type=pensioenopbouwen	2	Build up extra pension	No	Main	N/A
mutation/scheme/pensiontransfer	3	Request for value transfer	No	Sub	Submit changes op-tions
mutation/valuetransfer	3	Submit value transfer	No	Sub	Submit changes op-tions
mutation/sendofrelation	3	End relationship	No	Sub	Submit changes op-tions
klantreis?type=uitdienst	2	End employment	No	Main	N/A
klantreis?type=samenwonen	2	Start cohabitation information	No	Main	N/A
klantreis?type=samenvoegen	2	Combine pensions information	No	Main	N/A
klantreis?type=bijnametpensioen	2	Retiring soon information	No	Main	N/A

Short URL	Category	Type	Navigation (top)	Page type	Parent
klantreis?type=parttimewerken	2	Working part-time information	No	Main	N/A
klantreis?type=beleggen	2	Investment choices	No	Main	N/A
getviewdocument	3	Download document	No	Sub	Documents folder
mutation/cohabitation/end	3	End cohabitation	No	Sub	Submit changes options
mutation/nettoparticipationscheme	3	Net participation scheme	No	Sub	Submit changes options
klantreis?type=overlijden	2	Death	No	Main	N/A
mutation/mutationdetail	3	Details about mutations	No	Sub	Overview changes submitted
klantreis?type=nettopensioen	2	ASR net pension information	No	Main	N/A
mutation/submutationdetail	3	Details about submutations	No	Sub	Overview changes submitted
pension/survivors	3	Relatives pension information	No	Sub	Myscheme
mutation/endnettoparticipation	3	End participation net settlement	No	Sub	Submit changes options
klantreis?type=scheiden	2	Divorce information	No	Main	N/A
klantreis?type=pensioenverkenner	2	Pension explorer information	No	Main	N/A

Table 2
User group ID references

ID	Group name
1	PensionGross
2	InvestmentProfileNeutral
3	StatusActive
4	GenderMale
5	MaritalStatusMarried
6	Income21-50k
7	MaritalStatusSingle
8	Income51-100k
9	GenderFemale
10	Age45-54
11	Age55-64
12	Age35-44
13	Age25-34
14	Income0-20k
15	StatusSleeper
16	MaritalStatusCohabitation
17	Income100-200k
18	Age65+
19	InvestmentProfileOffensive
20	PensionNett
21	Age18-24
22	PensionGrossNett
23	Income200-500k
24	InvestmentProfileDefensive
25	Income501-1ml

Table 3
Features available in the employee relation SQL table

Features name(s)	**Description**
GEZ_GEMGOED, GEZ_ROKEND_PART GEZ_STATUS_BEGIN, GEZ_STATUS_EIND GEZ_WERKGEVER_PART GEZ_VOORL_PART, GEZ_VOORV_PART GEZ_ANAAM_PART, GEZ_TITEL_PART GEZ_SALARIS_PART GEZ_WAOPERC_PART, GEZ_TNMSTEL, GEZ_NATIONALITEIT, GEZ_BSN_PARTNER	Information about an employee's family, such as: family member names, genders, salaries, social security numbers, nationalities, and other personal information. Gez – stands for *gezin*, or *family* in Dutch
DEF_ADRESTYPE_NR, DEF_TELEFOONTYPE_NR DEF_EMAILTYPE_NR	Def – stands for *definitief*, and includes personal details such as: postal and email addresses, and phone numbers
PERSONEELSNUMMER, DATUM_DIENSTVERBAND DGA, DATUM_IN_DIENST, DATUM_UIT_DIENST EXTERN_RELATIENUMMER	Employee details such as: personnel number, date in/out of service, employer remarks (missing fields), external relationship number (if available)
ORGANISATIE_NR, TITULATUUR_NR AFDELING_NR, DIENSTVERBAND_NR PENSIOENREG_NR, ZIEKTEKOSTENREG_NR FUNCTIE_NR, WAO_NR, DIVISIE_NR VERZEKREG_NR	Nr – stands for *nummer,* and lists identification numbers such as: organization's number, departmental IDs, function numbers, and other identification numbers
VOORL, VOORV, ANAAM ZOEKNAAM, VOORNAAM, TNMSTEL	Features related to an employee first names, surnames, initials, search names, titles
REDEN_UIT_DIENST, TIJDSTEMPEL WAO_VANAF, EMPLOYMENTTYPE EINDARBEIDSCONTRACT, DATUM_VERLOF, ONBETAALDVERLOF IK_PERCOBV, AOW_KORTINGS_JAREN	Features related to employment such as: reason for leaving, contract end date, employment type, occupational disability insurance (WAO) starting date, leave of absence, registration dates, discounts, unpaid leave
ADRES_SOORT, SOFINUMMER REKENING_NR, NATIONALITEIT, IDTYPE, IDNUMMER, BANKNUMMER ID_EXPIRATIE_DATUM, VERSIE_EXTERN	Personal employee information such as: ID number, type, expiration date; nationality, bank account number, social security number
GESLACHT, GEBDAT, GEZ_HUW_STATUS, STATUS_NR, PENSIOENREGEX_NR and PENSIOENREGFN_NR, INK_PERIODE_SALARIS	**Demographic features selected for user segmentation: gender, age, marital status, user status, income, pension, investment profile**

Table 4
Overview of the mean and median ratings per user group

User ID	Mean	Median
1	0,64	0,34
2	0,63	0,33
3	0,64	0,33
4	0,64	0,34
5	0,65	0,38
6	0,66	0,39
7	0,65	0,35
8	0,64	0,37
9	0,62	0,34
10	0,63	0,33
11	0,65	0,39
12	0,66	0,39
13	0,73	0,44
14	0,76	0,40
15	0,76	0,40
16	0,65	0,29
17	0,48	0,15
18	0,55	0,25
19	0,69	0,36
20	0,60	0,24
21	0,85	0,58
22	0,52	0,21
23	0,56	0,20
24	0,74	0,43
25	0,90	0,83

Table 5
A snapshot of the Google Analytics data

Event Label	Users	Page Views
mijnwerknemerspensioen.asr.nl/Employee/Dashboard.aspx	3.383	27.885
mijnwerknemerspensioen.asr.nl/Employee/Dashboard.aspx	3.346	27.499
mijnwerknemerspensioen.asr.nl/Employee/Dashboard.aspx	3.030	25.471
mijnwerknemerspensioen.asr.nl/Employee/Dashboard.aspx	2.426	20.129
mijnwerknemerspensioen.asr.nl/Employee/Documents.aspx	2.392	21.850
mijnwerknemerspensioen.asr.nl/Employee/Documents.aspx	2.360	20.825
mijnwerknemerspensioen.asr.nl/Employee/Documents.aspx	2.132	19.119
mijnwerknemerspensioen.asr.nl/Employee/Notifications.aspx	1.955	14.163
mijnwerknemerspensioen.asr.nl/Employee/Notifications.aspx	1.923	14.274