

A Project Report on

**Recommender system for suggesting “Frequently used together” SAP Fiori Apps for SAP S/4 HANA system**

Submitted in partial fulfilment for award of degree of

**PGDBM**

In **Business Analytics**

Submitted by

**Jyoti Singh**

<SRN>

Under the Guidance of

**Dr. J. B. Simha**

<Designation>

REVA Academy for Corporate Excellence

**REVA University**

Rukmini Knowledge Park, Kattigenahalli,

Yelahanka, Bangalore – 560064

**August, 2021**



# Candidate’s Declaration

I, Jyoti Singh hereby declare that I have completed the project work towards the first year of Master of Business Administration in Business Analytics at, REVA University on the topic entitled Recommender system for suggesting “Frequently used together” SAP Fiori Apps for SAP S/4 HANA system under the supervision of <JB Simha and designation>. This report embodies the original work done by me in partial fulfilment of the requirements for the award of degree for the academic year 2021.

Place: Bengaluru Name of the Student: Jyoti Singh

Date: Signature of Student



# Certificate

This is to Certify that the Project work entitled <Title of the Report> carried out by <Student’s Name> with <SRN>, is a bonafide student of REVA University, is submitting the first year project report in fulfilment for the award of <Program Name> in Business Analytics during the academic year <Year>. The Project report has been tested for plagiarism, and has passed the plagiarism test with the similarity score less than 15%. The project report has been approved as it satisfies the academic requirements in respect of PROJECT work prescribed for the said Degree.

Signature of the Guide Signature of the Director

Name of the Guide Name of the Director

Guide Director

External Viva

Names of the Examiners

1. <Name> <Designation> <Signature>
2. <Name> <Designation> <Signature>

Place: Bengaluru

Date:



# Acknowledgement

I would like to convey a heartfelt thanks to all my mentors at RACE; Dr. J. B. Simha, Mr. Mithun D J and Mr. Ratnakar Pandey for their continuous support throughout the learning journey. A special mention to Dr. J. B. Simha for his valuable feedback and guidance as a Guide and Mentor throughout the project lifecycle.

I would like to express a special thanks to Dr. Shinu Abhi, Director REVA Academy of Corporate Excellence for her cordial support, valuable guidance, and information at various stages, that helped in completing the project.

Also, I would like to extend my heartfelt thanks to all the program office members; Mr. Pullikeshi M, Mr. John Wilson, Yusuf, Hasitha, Prajwal and others for being available round the clock to support in technical/program related assistance.

I would also like to thank Hon’ble Chancellor, Dr. P Shayma Raju, Hon’ble Vice Chancellor, Dr. M. Dhanamjaya, and Registrar, Dr. N. Ramesh, for providing this platform to learn and implement new ideas.

Lastly, I would like to thank my Manager Mr. Joy Ramachandran for supporting me in my decision to embark this project and providing encouragement in carrying out the project activities.

Place: Bengaluru

Date: 25th Aug 2021



# Similarity Index Report

This is to certify that this project report titled Recommender system for suggesting “Frequently used together” SAP Fiori Apps for SAP S/4 HANA system was scanned for similarity detection. Process and outcome is given below.

Software Used:

Date of Report Generation:

Similarity Index in %:

Total word count:

Name of the Guide:

Place: Bengaluru Name of the Student:

Date: Signature of Student

Verified by:

Signature

Dr. Shinu Abhi,

Director, Corporate Training

# List of Abbreviations

|  |  |  |
| --- | --- | --- |
| **Sl. No** | **Abbreviation** | **Long Form** |
| 1 | SAP | Systems, Applications & Products in Data Processing |
| 2 | SAP S/4 HANA | SAP Business Suite for SAP HANA |
| 3 | SAP HANA | SAP High-Speed Analytical Appliance |
| 4 | ERP | Enterprise Resource Planning |
| 5 | UX | User Experience |
| 6 | OTT | Over-the-Top |
| 7 | GUI | Graphical User Interface |
| 8 | UI | User Interface |
| 9 | CSAT | Customer Satisfaction |
| 10 | CRISP-DM | Cross Industry Standard Process for Data Mining |
| 11 | EWA | Early Watch Alert |
| 12 | SVD | Singular Value Decomposition |
| 13 | API | Application Programming Interface |
| 14 | LOB | Line of Business |
| 15 | k-NN | K Nearest Neighbours |
| 16 | surPRISE | Simple Python Recommendation System Engine |
| 17 | NMF | Non-negative Matrix Factorization |
| 18 | PMF | Positive Matrix Factorization |
| 19 | RMSE | Root Mean Square Error |
| 20 | MAE | Mean Absolute Error |
| 21 | CV | Cross Validation |
|  |  |  |

# List of Figures

|  |  |  |
| --- | --- | --- |
| **No.** | **Name** | **Page No.** |
| Figure No. 1 | SAP Fiori Apps on various platforms | 10 |
| Figure No. 2 | Similar Apps Recommendations | 11 |
| Figure No. 3 | Frequently used Apps Recommendations | 12 |
| Figure No. 4 | CRISP-DM Lifecycle | 18 |
| Figure No. 5 | Project Approach Snapshot | 19 |
| Figure No. 6 | High Level Platform Architecture | 21 |
| Figure No. 7 | SAP HANA Tables – Entity Relation diagram | 22 |
| Figure No. 8 | Data Extract Snapshot | 22 |
| Figure No. 9 | Top 10 Used Apps | 23 |
| Figure No. 10 | Top 15 Application Component | 23 |
| Figure No. 11 | Fiori Apps Statistics | 24 |
| Figure No. 12 | Distribution of Usage Count | 24 |
| Figure No. 13 | Snapshot of Pre-processed Dataset | 26 |
| Figure No. 14 | Customer-App Interaction Matrix | 26 |
| Figure No. 15 | Cosine Similarity between vectors | 27 |
| Figure No. 16 | Cosine Similarity formula | 28 |
| Figure No. 17 | Top 10 similar Fiori Apps Recommendation | 28 |
| Figure No. 18 | Snapshot of Top 10 Fiori Apps Recommendations using SVD model | 29 |
| Figure No. 19 | Deep Neural Network | 30 |
| Figure No. 20 | RMSE Values for 20 Epochs | 30 |
| Figure No. 21 | MAE Values for 20 Epochs | 31 |
| Figure No. 22 | Snapshot of Top 10 Fiori Apps Recommendations using Residual Network | 31 |

# List of Tables

|  |  |  |
| --- | --- | --- |
| **No.** | **Name** | **Page No.** |
| Table No. 1 | Key Evaluation Metrics | 33 |
| Table No. 2 | Customers for Model Validation | 34 |

# Abstract

There are hundreds of SAP Fiori Apps in SAP S/4 HANA system, with simplified user interface supporting multiple device types. However, only few of them are getting actively used by Customers, as many of them are ignorant and unfamiliar with the full library of SAP Fiori Apps. Customers are unable to use these simplified and the award-winning user interface in the form of Fiori Apps, for most of their day-to-day activities, resulting in the difficult and slow adoption of S/4 HANA system.

The project takes the usage statistics of SAP Fiori Apps from S/4 HANA on-premises and S/4 HANA Cloud customers. In addition, a mapping is built for the relationship between Apps based on the Line of Business, Business Role, Industry, Application Area.

From the combination of the usage data and the relationship data, Recommendation system is built that Recommends Similar Apps based on the Content Based Recommendation filtering method and the Best Bets for Apps that a customer is not yet using, but which are trending at other customers who use related apps using the User-Based Collaborative Filtering method.

In this way a personalized recommendation of trending apps is provided to the customers and the end-users, thereby increasing the overall Customer Satisfaction.

Keywords: Recommendation Engine, Content Based, User Based Collaborative Filtering, Surprise Models, Residual Network

# Table of Contents

[Candidate’s Declaration 2](#_Toc80916790)

[Certificate 3](#_Toc80916791)

[Acknowledgement 4](#_Toc80916792)

[Similarity Index Report 5](#_Toc80916793)

[List of Abbreviations 6](#_Toc80916794)

[List of Figures 6](#_Toc80916795)

[List of Tables 7](#_Toc80916796)

[Abstract 8](#_Toc80916797)

[Table of Contents 9](#_Toc80916798)

[Chapter 1: Introduction 10](#_Toc80916799)

[Chapter 2: Literature Review 13](#_Toc80916800)

[Chapter 3: Problem Statement 16](#_Toc80916801)

[Chapter 4: Objectives of the Study 17](#_Toc80916802)

[Chapter 5: Project Methodology 18](#_Toc80916803)

[Chapter 6: Business Understanding 20](#_Toc80916804)

[Chapter 7: Data Understanding 21](#_Toc80916805)

[Chapter 8: Data Preparation 25](#_Toc80916806)

[Chapter 9: Modeling 27](#_Toc80916807)

[Content Based Recommendation Model using Cosine Similarity 27](#_Toc80916808)

[User-Based Collaborative Filtering Model using Surprise Library 28](#_Toc80916809)

[Residual Networks with Keras Library 29](#_Toc80916810)

[Chapter 10: Model Evaluation 32](#_Toc80916811)

[Chapter 11: Analysis, Results and Deployment 34](#_Toc80916812)

[Chapter 12: Conclusions and Future Scope 35](#_Toc80916813)

[Bibliography 36](#_Toc80916814)

[Appendix 37](#_Toc80916815)

[Plagiarism Report 37](#_Toc80916816)

[Publications in a Journal/Conference Presented/White Paper 37](#_Toc80916817)

[Any Additional Details 37](#_Toc80916818)

# Chapter 1: Introduction

* 1. **Background Information**

SAP Business Suite for SAP HANA (S/4 HANA) is SAP's Enterprise Resource planning (ERP) for large enterprises meant to cover all day-to-day processes of an enterprise (for example, order-to-cash, procure-to-pay, plan-to-product, and request-to-service) and core capabilities. It is a next-generation, intelligent ERP business suite and is the successor of SAP ERP designed specifically for SAP’s in-memory computing SAP HANA database.

SAP Fiori is the new User Experience (UX) for SAP software that applies modern design principles. It comprises of the set of apps, that address the most broadly and frequently used SAP functions, such as workflow approvals, information lookups, and self-service tasks etc. They provide simple and easy-to-use user interface across desktops, tablets, and smartphones. SAP Fiori delivers an intuitive, role based UX interface that improves both employee productivity and satisfaction.



Figure 1: SAP Fiori Apps on various platforms

SAP S/4 HANA uses SAP Fiori as its front-end technology to provide personalized, responsive, and simple user experience to customers. All new functions, features and innovations of SAP S/4 HANA are accessible in SAP Fiori Launchpad. Using this launchpad customers can call the backend functionalities for which they have been granted access.

* 1. **Problem Statement**

There are hundreds of SAP Fiori Apps in SAP S/4 HANA system, with simplified user interface supporting multiple device types. However, only few of them are getting actively used by Customers, as many of them are ignorant and unfamiliar with the full library of SAP Fiori Apps.

Customers are unable to use these simplified and the award-winning user interface in the form of Fiori Apps, for most of their day-to-day activities, resulting in the difficult and slow adoption of S/4 HANA system.

* 1. **Proposed Solution**

Buyers at Amazon know this very prominent offer, after they have picked a product; a bundle of 2-3 products under the title “Frequently bought together” is showcased to buyers. Using this idea, I would like to drive the adoption of SAP Fiori apps among customers, because it is essential for the success of SAP S/4HANA, which is the most widely used product of SAP.

This project will take the usage statistics of SAP Fiori Apps from S/4 HANA on-premises and S/4 HANA Cloud customers. In addition, a mapping will be built for the relationship between Apps based on the Line of Business, Business Role, Industry, Application Area and Customer’s demographic information like Country and Region.

Based on the usage pattern of individual customer, Apps with maximum similarity will be recommended to the customer as ‘Similar Apps’.

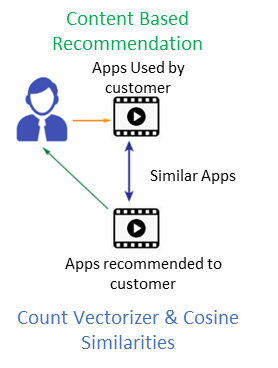


Figure 2: Similar Apps Recommendations

Also, from the combination of the usage data of different customers and the relationship data of Apps and Customers, system will recommend the Best Bets for Apps that a customer is not yet using, but which are trending at other customers who use related apps as ‘Frequently used Apps’.

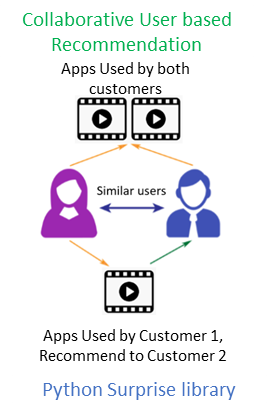


Figure 3: Frequently used Apps Recommendations

In this way a personalized recommendation of trending apps can be provided to the customers and the end-users. This will result in enhanced Customer Satisfaction and in turn better adoption of SAP S/4 HANA system.

# Chapter 2: Literature Review

In today’s world due to pandemic and with the rise of usage of Over-the-top (OTT) media services like Amazon prime videos, Netflix and many other online E-commerce websites like Amazon, Flipkart; Recommender systems have taken more and more place in our lives. From e-commerce (where Recommender system suggest the products of interest to Buyers) to online movie platform (where Recommender system suggest the movies to Users similar to their preferences), recommender systems has become the inseparable entity in our daily online journeys. (*Introduction to Recommender Systems | by Baptiste Rocca | Towards Data Science*, n.d.)

Recommender systems are algorithms that aim to predict customers’ interests and recommend product items that quite likely are interesting for them. They are among the most powerful machine learning systems that online retailers implement in order to drive sales and generate a huge amount of income when they are efficient or also be a way to stand out significantly from competitors.

Recommender systems are used for providing personalized recommendations based on the user profile, previous behaviour, and other customer’s usage pattern. Recommendation systems help the users to find and select items (e.g., books, movies, restaurants) from the wide collection available on the web or in other electronic information sources. Among a large set of items and a description of the user’s needs, they present to the user a small set of the items that are well suited to the customer’s preference. Similarly, a movie recommendation system provides user with the set of movies that is best suited for his watching pattern. (Geetha et al., 2018)

Collaborative Filtering and Content based filtering are the two major paradigms followed for building Recommender systems.

**Collaborative-filtering**: In this approach recommendation for each active user is calculated by comparing with the preferences of other users who have rated the product in similar way to the active user. Basically, Collaborative filtering methods are established on gathering and examining a large amount of information which based on user’s demeanour, activities or preferences and anticipating taste of that particular user by using their similarity with other users.

**Content-Based filtering**: In this approach the recommendations depend on users’ past preferences and behaviour. Item description and a profile of the users’ orientation play an important role in Content-based filtering. Content-based filtering algorithms try to recommend items based on similarity count.

**Hybrid filtering**: The hybrid filtering is a combination of more than one filtering approach. The hybrid filtering approach is introduced to overcome some common problem that are associated with above filtering approaches such as cold start problem, overspecialization problem and sparsity problem. Another motive behind the implementation of hybrid filtering is to improve the accuracy and efficiency of recommendation process and provide the personalized Recommendation. (B.Thorat et al., 2015)

An Online Book Recoomendation system is presented by Sushma Rajpurkar which uses combined features of Content filtering, Collaborative filtering and Association rule mining to produce efficient and effective book recommendations that are of Buyer’s interest. The book recommendation has considered many parameters like content of the book and quality of the book by doing collaborative filtering of ratings by the other buyers. (Rajpurkar, 2015)

A personalized Product Recommendation system is proposed by *Sung-Shun Weng*. It analyses customers’ purchasing behaviours based on product features from transaction records and product feature databases. Rules of customer interest profiles are thus derived to recommend customers products that have potential attraction with customers. The approach uses Content based filtering algorithm and has its strength to be able to recommend to customers brand new products or rarely purchased products if they fit customer interest profiles. (Weng & Liu, 2004)

A hybrid online movie recommender system has been presented by Harpreet Kaur. (Virk et al., 2015). The system uses a mix of content based as well as collaborative filtering algorithm. The context of the movies is also considered while recommending. The user - user relationship as well as user - item relationship plays a role in the recommendation.

The key ingredient of this project is to use the hybrid approach of Collaborative filtering method and Content based Filtering method to fill the gap of online Apps recommendation. It targets to provide the best and personalized Fiori Aspps recommendations to SAP customers’ using SAP S/4 HANA system.

# Chapter 3: Problem Statement

SAP is a market leader in providing ERP (Enterprise Resource and Planning) solutions and services.

In Feb 2015, SAP released their new Solution SAP S/4 HANA. It is a next-generation, intelligent ERP business suite and is the successor of SAP ERP designed specifically for SAP’s in-memory computing SAP HANA database and enhanced User Interface. SAP Fiori Apps are the integral part of on-premise or Cloud centric SAP S/4 HANA systems.

Prior to release of SAP S/4 HANA, customers used SAP GUI (Graphical User Interface) screen for all ERP related activities. SAP GUI is SAPs universal and classic UI technology for working with SAP systems, however it lacks the enhanced User experience in terms of user friendliness and supporting multiple device types.

With SAP S/4 HANA, SAP launched Fiori Apps to provide a powerful yet simplified platform to maximize Customer Experience from ease of navigation to agile Dashboards.

However, not many customers are aware of the full library of SAP Fiori Apps and are using traditional ways of accessing ERP screens. Customers are unable to use these simplified and the award-winning user interface in the form of SAP Fiori Apps, for most of their day to day activities. Also, they are ignorant about the multi-device support that is provided by SAP Fiori Apps to ease their approval or reporting needs. This is resulting in the difficult and slow adoption of SAP S/4 HANA system and thereby resulting in reduced Customer Satisfaction.

# Chapter 4: Objectives of the Study

The Objective of the study is to build an accurate and accelerated Recommender system that would recommend customers’ the Best Bets for SAP Fiori Apps that a customer is not yet using, but which are trending at other customers who use related apps.

Also, the system will recommend similar Apps based on recorded information on the customers' preferences and the usage pattern.

The system will use information filtering techniques to process information and provide the customer with potentially more relevant items.

This Recommender system will help customers find interesting and relevant Apps from within a large information space. Using the recommendations provided by system, customers would get an opportunity to explore the areas/Apps which are not used by them currently.

SAP S/4HANA is the most widely used product of SAP and the project will enhance the adoption of SAP Fiori apps among customers, thereby increasing the overall Customer Satisfaction (CSAT) score.

# Chapter 5: Project Methodology

The Cross Industry Standard Process for Data Mining (CRISP-DM) process is used to achieve the project’s objective.

The process comprises of six phases: (*CRISP-DM - Data Science Process Alliance*, n.d.)

1. Business Understanding – This phase is used in the project to comprehend the Problem statement and define the Business success criteria.
2. Data Understanding – This phase is used in the project to identify the source of data, data collection and data exploration
3. Data Preparation – This phase is used in the project for cleansing the data, data imputation and data wrangling to create new dataset that would be used for Modelling
4. Modelling – This phase is used in the project to build different models and assess these models to achieve an accurate and accelerated Recommender system
5. Evaluation – This phase is used in the project to get the models evaluated by Business for accuracy and agileness
6. Deployment – This phase is not included in the current project as deployment is a long-term plan by SAP

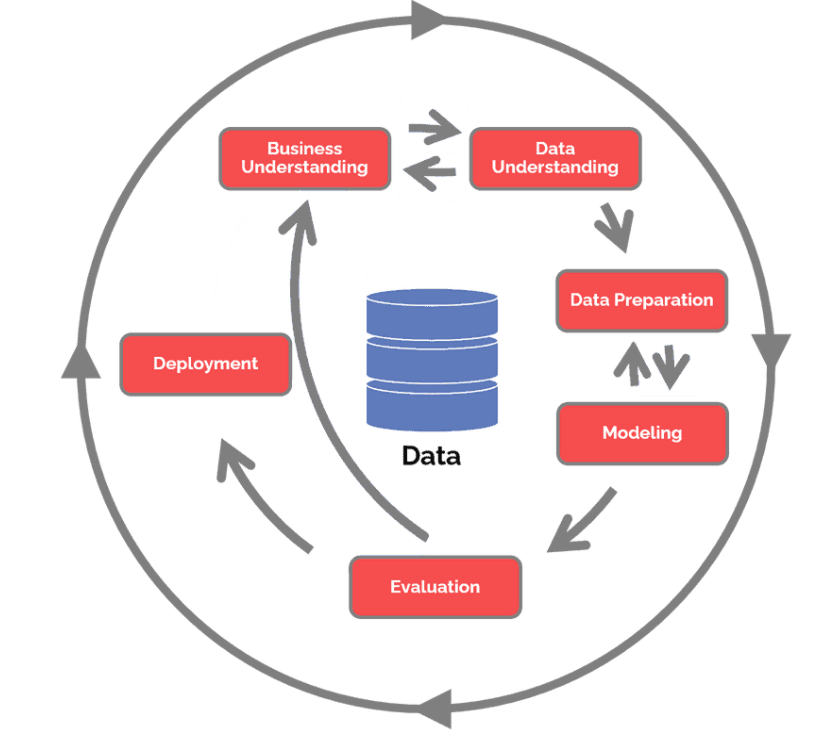


Figure 4: CRISP-DM Lifecycle

The following detailed approach is followed in this project:

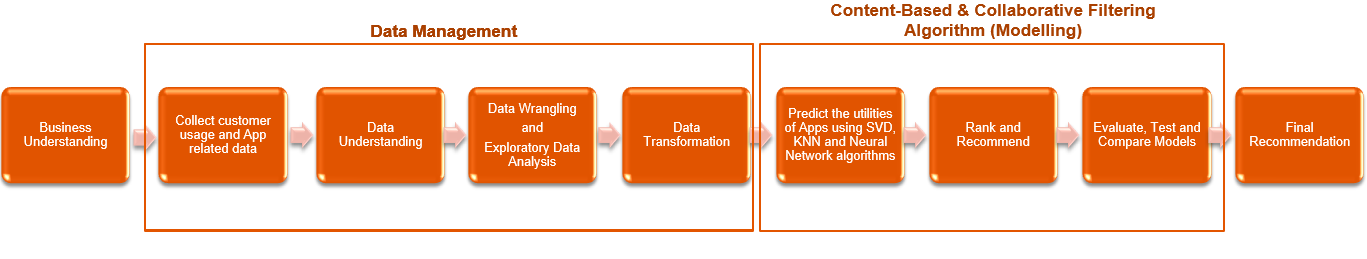


Figure 5: Project Approach Snapshot

In this project a Content-Based and User-Based collaborative filtering Recommender system is built, using the data from both SAP Fiori Apps library and SAP EarlyWatch alert usage data. A relationship mapping i.e., an interaction matrix is generated between the App usage by the customers and the Apps attributes like Line of Business, Business Role, Industry, Application Area. K-Nearest Neighbour, Singular Value Decomposition (SVD), Cosine Similarity and Deep Neural Network Algorithms are used to predict the utilities of Fiori Apps to Customers based on the interaction Matrix.

Customer’s landscape data also provide information about the SAP Fiori Apps list per system which are installed at Customer’s landscape but are not getting used. Using the information from the mapping tables, Customer’s usage data and the landscape data, a personalized recommendation of Trending Apps is generated and proposed to customers based on their Business and Industry types.

# Chapter 6: Business Understanding

Customers who have migrated to from ERP platform to new platform of SAP S/4 HANA are ignorant or unfamiliar with the SAP Fiori Apps. They are unable to utilize the dynamic and valuable capabilities of S/4 HANA platform, to get their routine activities simplified. This in turn is impacting the Customer Satisfaction Score of the tool.

The Business objective is to build Recommendation Engine that would benefit SAP S/4 HANA Customers achieve their regular activities in much simplified way. The proposed goals are as below:

1. To identify the similar SAP Fiori Apps and recommend top 10 similar Fiori Apps to a particular SAP S/4 HANA on-premise or Cloud Customer, based their usage pattern in past. These recommended Apps should be sorted based on the similarities score in descending order.
2. To identify and recommend the top 10 trending SAP Fiori Apps between other similar Customers and recommend to a particular Customer if they are not using those Apps. Similarity of the Customers can be identified based on the Industry, Application Component and Line of Business of the Apps used by Customers.

Data Source for the innovation project is identified as the Fiori App usage data which is collected from all the customer’s on-premise and Cloud landscapes on weekly basis using the SAP EarlyWatch Alert workspace and are stored in the relevant tables in SAP HANA database.

# Chapter 7: Data Understanding

SAP Early Watch Alert (EWA) Workspace is a one-stop-shop for all customer’s systems related information. It gives a comprehensive overview of the customer’s system landscape in terms of stability, configuration, hardware utilization and performance. SAP Fiori App usage data from all the customer’s on-premise and Cloud landscapes are collected on weekly basis using the SAP EWA workspace and are stored in the relevant tables in SAP HANA database.

Data from SAP HANA database is extracted using SAP HANA Python Client API for Data Science innovation.

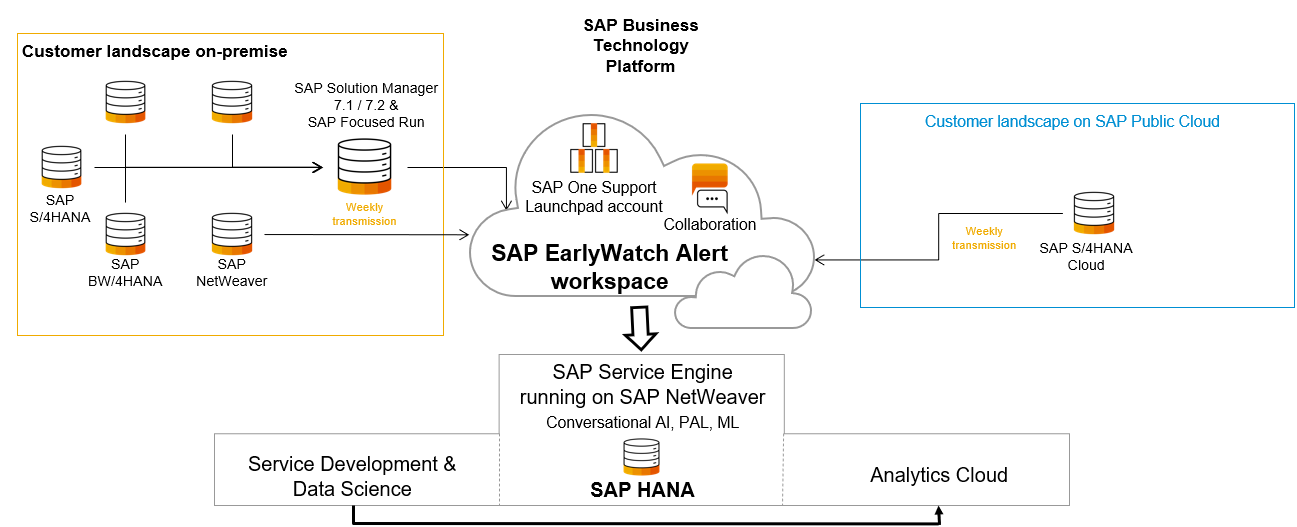


Figure 6: High Level Platform Architecture

Below are the 5 tables used to collect the information regarding Customer including their Demograhic details and FIORI App usage details. Customer data is extracted for 6 weeks and comprises of approximately 2,32,000 observations.

SAP Fiori Apps reference library i.e. FIORI\_BOM table is the comprehensive library that provides information about all SAP Fiori apps including the technical data and additional data like Line of Business, Business Role, Industry, Application Area associated with the Apps.

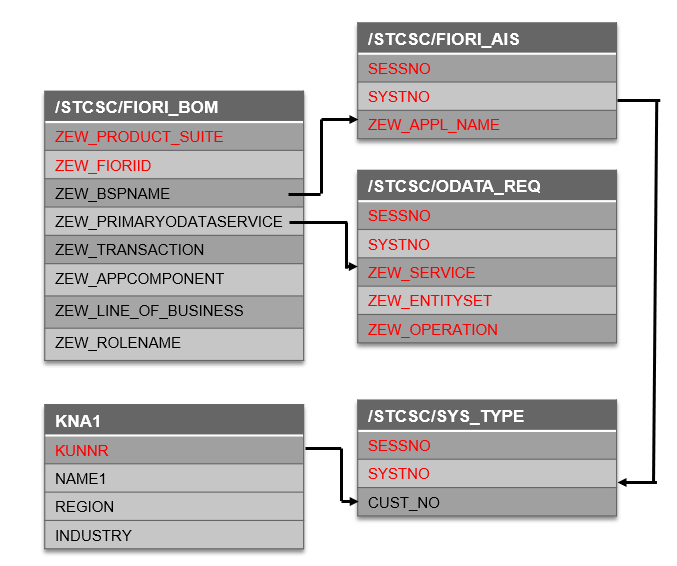


Figure 7: SAP HANA Tables – Entity Relation diagram

With the usage of table joins below information is extracted from the pool of data.

1. CUST\_NO – Customer Number
2. NAME1 – Customer Name
3. REGION – Customer Region
4. INDUSTRY – Customer business industry/domain
5. CALWEEK – Calendar Week for which data is extracted
6. ZEW\_SERVICE – Fiori App Service
7. ZEW\_FIORIID – Fiori App Id
8. ZEW\_APPL\_NAME – Fiori App Name
9. ZEW\_APPCOMPONENT – Application Component
10. ZEW\_LINE\_OF\_BUSINESS – Line of Business
11. ZEW\_ROLENAME – Business Role using the Fiori App

Using the Group function, Fiori App usage Count is calculated for each Customer per Fiori App and App Component. Below is the snapshot of extracted data.

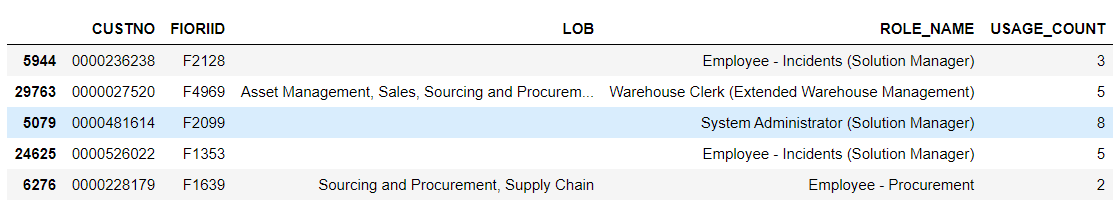


Figure 8: Data Extract Snapshot

Below Exploratory Data Analysis is performed for better understanding of the data.

1. Top 10 used Apps by all Customers with their usage count

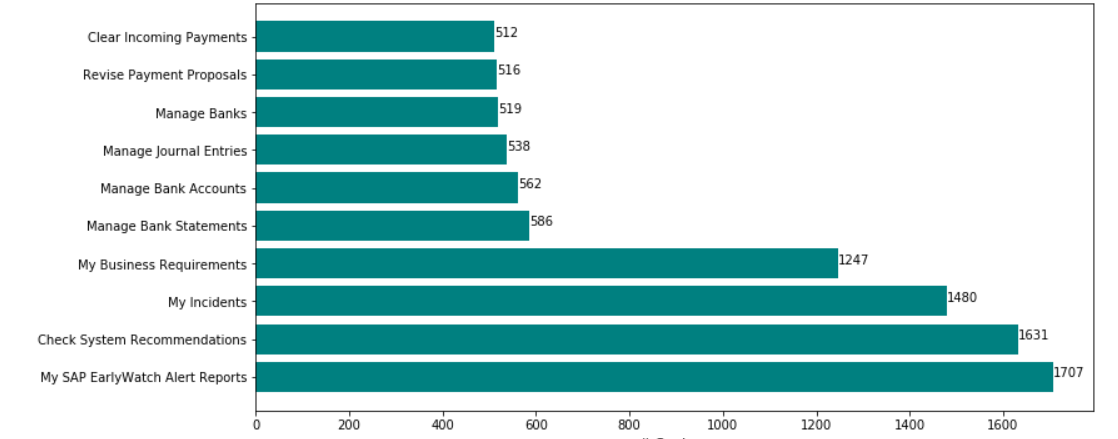


Figure 9: Top 10 Used Apps

1. Top 15 Application Component based on Usage Count

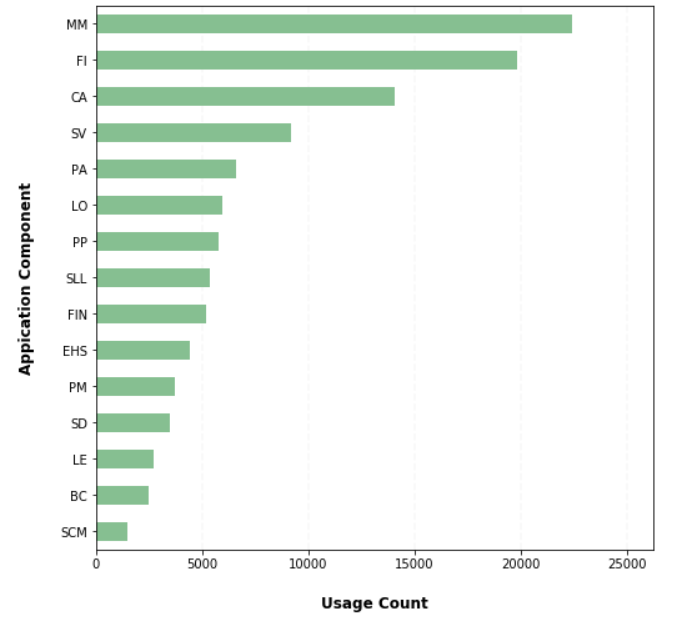


Figure 10: Top 15 Application Component

1. Fiori Apps Statistics

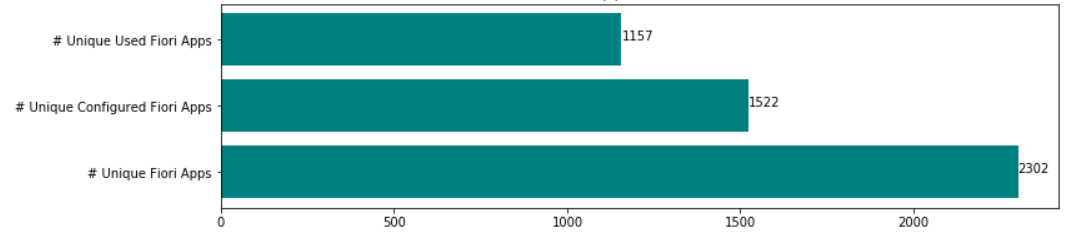


Figure 11: Fiori Apps Statistics

1. Distribution of Usage Count

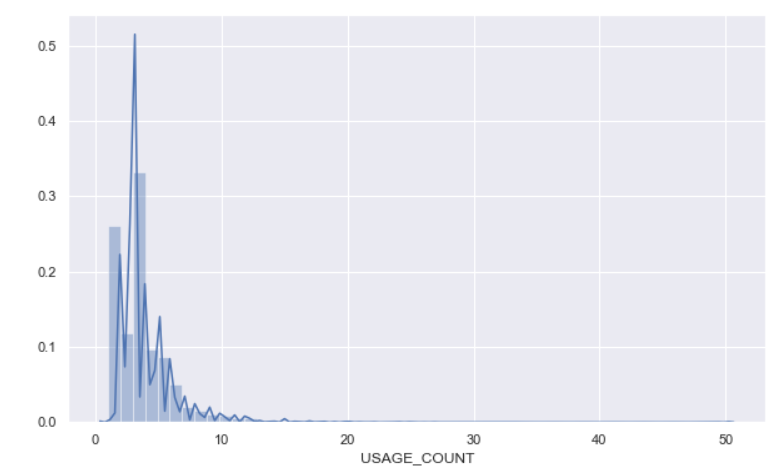


Figure 12: Distribution of Usage Count

# Chapter 8: Data Preparation

The raw data of Customer and Fiori Apps collected from EWA Workspace comprised of Customer Number, Fiori App Id, Role Name, Application Component in the textual format. In order to prepare the final dataset to be fed to Recommender system models, dataset was pre-processed to assign unique Integer id to the unique Customer Number, Fiori App Id, Role Name and Application Component respectively.

Raw dataset comprised of comma separated values for Line of Business(LOB) and approximately 20% of missing values. Using get\_dummies() function with separator attribute as comma (,) , Line of Business was converted to columns with values as 0 and 1. pandas.get\_dummies() function is used convert categorical variable into dummy/indicator variables.

An additional column was added as LOB\_NAN to identify if the LOB comprises of missing value. Value 1 was assigned to LOB\_NAN if it had missing value else it was assigned value 0.

Missing data imputation of LOB is done using KNeighborsClassifier of Scikit-learn library which is K-Nearest Neighbours (KNN) Classification Algorithm. The K in the name of this classifier represents the number of nearest neighbors (k), where k is an integer value specified by the user.

For imputing missing LOBs, K value is chosen as 15 and multi-class classification is performed using KNN classifier. Fiori Apps in general are associated with one or more LOBs, hence Multi-Class classification is used to classify the instances into one or more classes.

K-NN model is trained using the Application Component and Role Name fields of the Fiori App data with known LOBs. The prediction target is set as the list of unknown LOBs of Fiori Apps.

Below is the snapshot of final dataset post pre-processing of raw dataset.

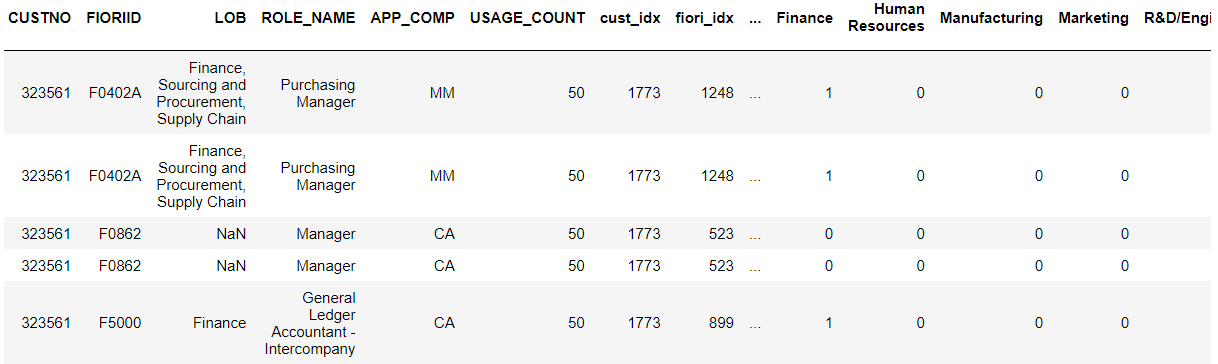


Figure 13: Snapshot of Pre-processed Dataset

For Recommendation Engine Algorithm, dataset needs to be converted into matrix comprising of the App usage count by a set of Customers to Fiori Apps from a set of Apps, called as Interaction Matrix. Each row contains the usage count by a Customer, and each column contains the Usage Count of an App.

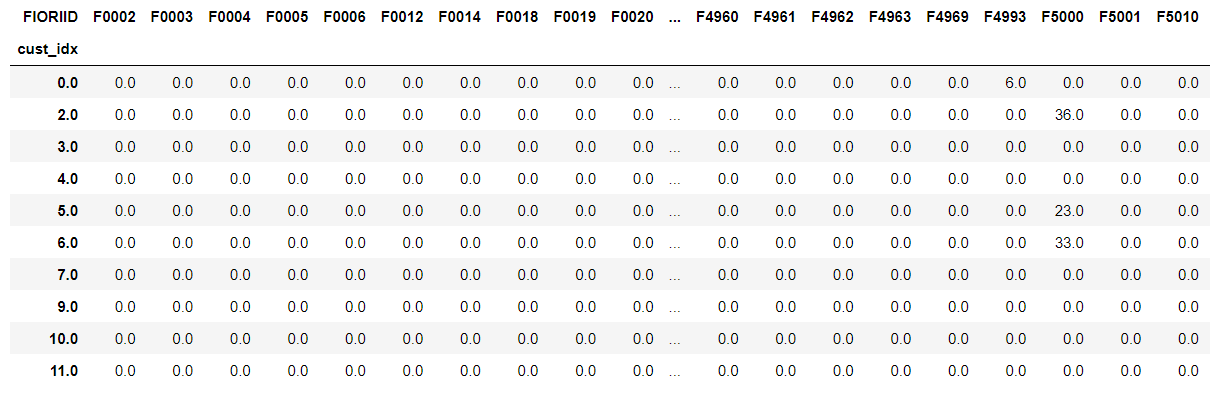


Figure 14: Customer-App Interaction Matrix

In majority of cases, the cells in the matrix are empty, as each customer use only a few Apps. It is highly unlikely for every Customer to use every other available Apps, thereby resulting in mostly empty cells. A matrix with mostly empty cells is called as sparse, and with mostly filled matrix is called dense. Sparse Interaction Matrix sometimes leads to weak recommendations. The sparsity level of extracted dataset that will be used for modelling is 96.5%.

# Chapter 9: Modeling

The project uses below 3 approaches to achieve the Recommendation engine’s objective to recommend similar and most trending SAP Fiori Apps to SAP S/4 HANA customers.

## Content Based Recommendation Model using Cosine Similarity

Content Based Recommendation Engine recommends products based on the attributes /metadata of the product already used by the Customer rather than using the interactions between different Customers. This method is used in the project to recommend similar Fiori Apps to customer based on their previous usage history. This algorithm does not involve usage pattern of other Customers.

The dataset comprises of metadata like Application Component, LOB and Role name for each Fiori Apps. These metadata are merged to create new column named *Features*. Using the CountVectorizer() method of Scikit-learn library, bag of words is generated for Apps *Features*. These Bag of Words are used to train the Machine Learning Model.

The project uses Cosine Similarity method to compute the similarities between various Apps. Cosine similarity between two objects measures the angle of cosine between the two objects. It compares two vectors on a normalized scale. It can be done by finding the dot product between the two identities.

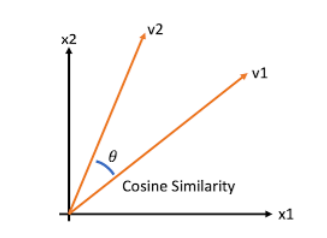


Figure 15: Cosine Similarity between vectors

In the above figure, angle q represents the similarity between vector v1 and v2. Lesser the angle more is the similarity between the vectors and vice-versa. (Singh et al., 2020)

Cosine Similarity is calculated as:

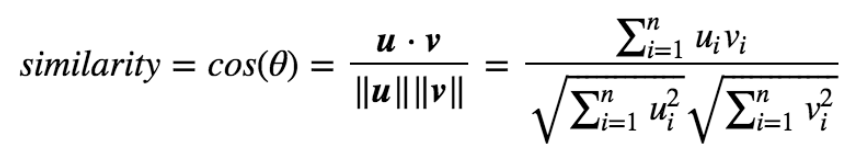
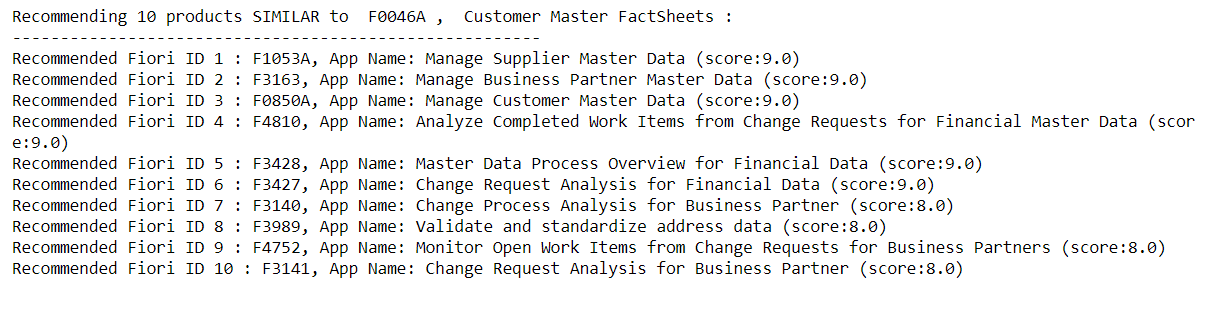


Figure 16: Cosine Similarity formula

Using the Cosine Similarity method, recommendation of top 10 similar Fiori Apps are generated based on the matching score with the already used Apps by the customer. Cosine similarity score is used since it is independent of magnitude and is simpler and faster to calculate.

 Figure 17: Top 10 similar Fiori Apps Recommendation

## User-Based Collaborative Filtering Model using Surprise Library

User Based Collaborative Filtering Model recommends products to a user that is not used/liked by user but is used/liked by the similar users. Interaction Matrix between multiple users acts as the input for this type of Model.

The project uses Surprise library for User-based recommendation model. Surprise is a Python library for creating and analyzing Recommender systems that deals with only explicit rating/usage data. The name *SurPRISE* stands for *Simple Python Recommendation System Engine.* The library provides various ready-to-use prediction algorithms like Singular Value Decomposition (SVD), K-Nearest Neighbour (k-NN), Non-Negative Matrix factorization (NMF), Positive Matrix Factorization (PMF) and many more. It comprises of Cross-Validation procedures that can be run for various iterators. (*Surprise · A Python Scikit for Recommender Systems.*, n.d.)

In the project final dataset is split into Train and Test Data in 75% and 25% respectively. The various Surprise library models used for prediction are SVD, NMF, k-NN Basic, k-NN Baseline, k-NN with Means and Normal Predictor. The metrics Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Fit time and Test time with 5 folds Cross Validation (CV) are used to evaluate the Model’s effectiveness.

Model Evaluation details are listed in Chapter 10. Based on the Model Evaluation, SVD stands out to be the best model among all the models with average RMSE value 1.15. Final Recommendation for Top 10 Apps that are not used by Customer but are trending at other Customer location is provided using SVD model. Below is the snapshot of final Recommendation.



Figure 18: Snapshot of Top 10 Fiori Apps Recommendations using SVD model

## Residual Networks with Keras Library

As an alternate approach to User-based Collaborative Filtering approach, project uses Residual Networks approach using Keras Library.

A deep Neural Network model uses multiple layers to reduce the error rate. This holds good for less number of layers but when the number of layers increases it leads to a common problem called Vanishing/Exploding Gradient as a result it leads to over-fitting. In order to reduce this error, Residual Networks are used where there are two models learning side by side. The error from the first model is fed into the Deep Neural Network model, thereby reducing the number of layers and yielding better results.

The project uses Residual Network of Matrix Factorization model as the first model. Output of this model is fed to Deep Neural Network model with 4 network layers and 20 epochs.

Matrix factorization is a type of collaborative filtering algorithms used in recommender systems that work by decomposing the user-item interaction matrix into the product of two lower dimensionality rectangular matrices.

Deep Neural network comprises of Input Layer, multiple Hidden Layers, and an output layer. The multiple hidden layers result in better predictions, however more the number of layers, the more complex the network becomes, and it takes more resources and time to train the model. (*Deep Neural Network - an Overview | ScienceDirect Topics*, n.d.)

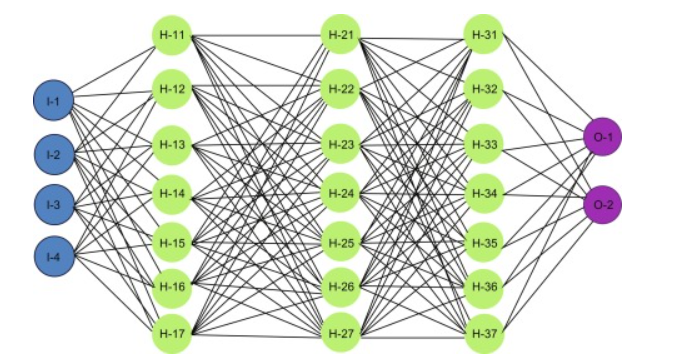


Figure 19: Deep Neural Network

Final Dataset is split into 80% Training set and 20% Test set. The dataset comprises of metadata like Application Component, LOB and Role name for each Fiori Apps and Customer’s region and Industry as Demographic information. These metadata are used to train the models along with SAP Fiori App Id and Customer Number.

Below is the graphical representation of Train/Test RMSE and MAE values for each Epochs.

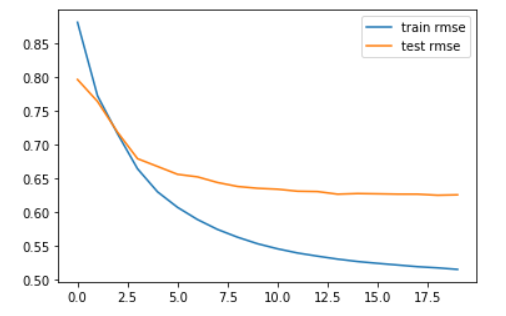


Figure 20: RMSE Values for 20 Epochs

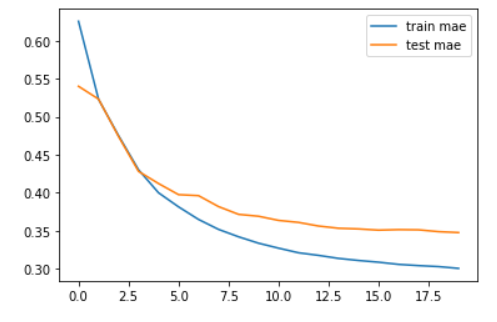


Figure 21: MAE Values for 20 Epochs

Below is the snapshot of final Recommendation.

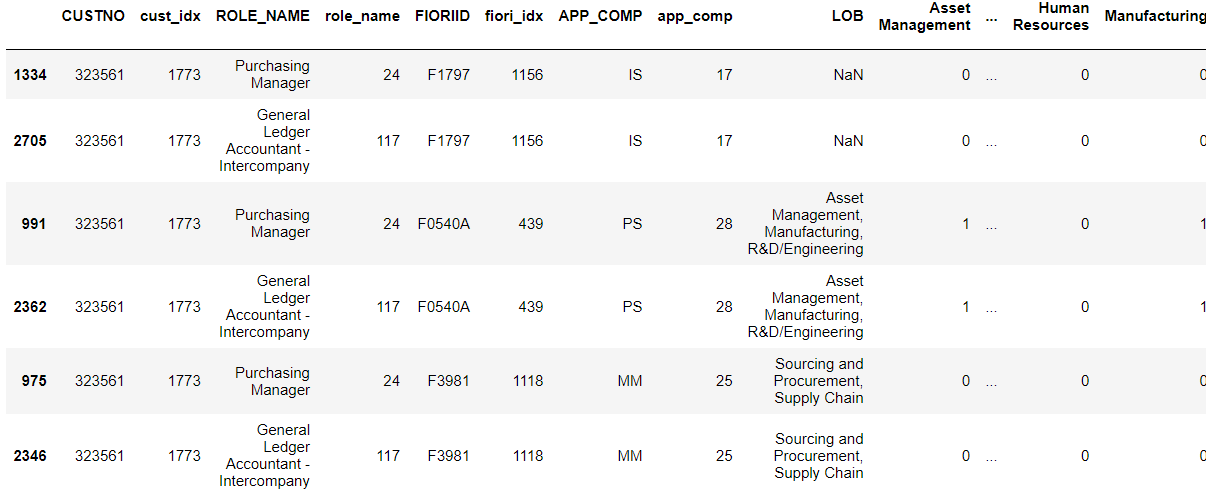
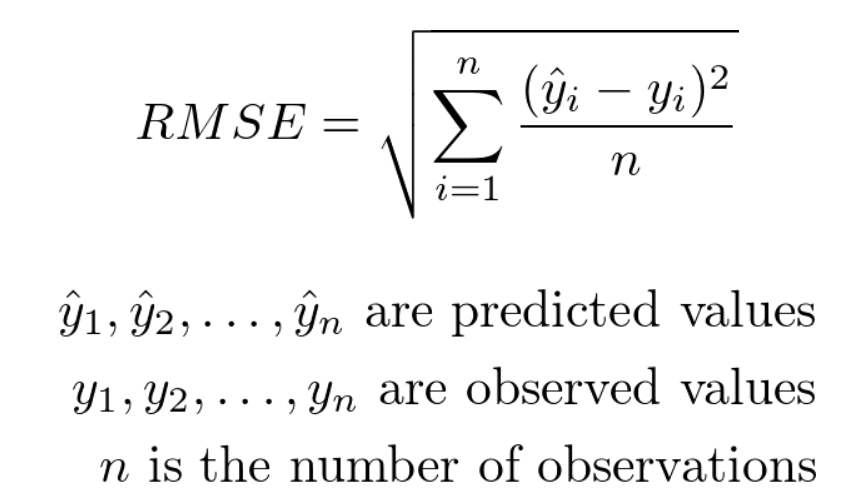


Figure 22: Snapshot of Top 10 Fiori Apps Recommendations using Residual Network

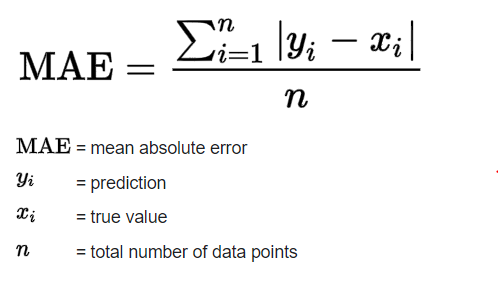
# Chapter 10: Model Evaluation

The key metric for Recommender Model Evaluation is:

1. Root Mean Square Error (RMSE) is the way in which the standard deviation of the residuals i.e., prediction errors are calculated. Residuals are a measure of how far from the regression line data points are located and RMSE is a measure of how spread out these residuals are. In other words, it tells you how concentrated the data is around the line of best fit. (*What Does RMSE Really Mean?. Root Mean Square Error (RMSE) Is A… | by James Moody | Towards Data Science*, n.d.)



1. Mean Absolute Error (MAE) is the average of all the Absolute errors i.e., the difference between the predicted value and the True value.



1. Fit Time – Time taken by a model to train itself on the Training dataset and generate the key metrics value based on the Test Dataset.
2. Prediction Time – Time taken by a model to predict values for the unknown dataset based on the Model learning/fitting.

Below table represent the Key metrics values for the predictive models used for building App Recommendation engine.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **RMSE** | **MAE** | **Fit time (secs)** |
| SVD | 1.15 | 0.58 | 187.5 |
| NMF | 1.21 | 0.76 | 192.3 |
| k-NN Basic | 1.25 | 0.83 | 226.15 |
| k-NN Baseline | 1.30 | 0.79 | 209 |
| k-NN with Means | 1.43 | 0.81 | 234.1 |
| Normal Predictor | 3.45 | 0.84 | 193.5 |
| Residual Networks (Keras) | 0.67 | 0.45 | 520 |

Table 1: Key Evaluation Metrics

# Chapter 11: Analysis, Results and Deployment

Based on the different types of Models’ Key performance metrics, Singular Value Decomposition (SVD) with Surprise library is in the process of getting selected for the final deployment.

The model provided Top 20 Fiori Apps recommendations for the below sample Customers which was validated with the Customers based on their App usage.

|  |  |
| --- | --- |
| **Customer Number** | **Customer Name** |
| 11201 | Volkswagen |
| 10718 | BSH Hausgeräte |
| 13932 | Hella |
| 1280299 | Jumbo Supermarket |
| 284268 | Axa Konzern |
| 638908 | Delhaize |

Table 2: Customers for Model Validation

The Fiori Apps recommendations suggested by SVD model are appreciable by the Customers as the next best App to be used in their routine activities.

The time taken by SVD model for training, testing and final prediction is very minimal as compared to the Deep Neural network model.

Hence, SVD stands out in terms of Apps Recommendations, usage of resources, time consumption and reduced complexity.

Production Deployment of the model will include building up the SAP Fiori Frontend that will use SVD model with Surprise library as the backend code for Apps Recommendation to SAP Customers. Production Deployment is not part of this project and will be handled outside this project activities.

# Chapter 12: Conclusions and Future Scope

The project uses various Recommender engine models and the dataset comprising of Customer’s demographic details, SAP Fiori Apps features and metadata and the usage information to recommend the similar SAP Fiori Apps and the top trending Fiori Apps at other Customers which are not getting used for a particular Customer.

The recommendations provided by the top models are personalized Recommendations for Customers based on their Region, Industry, Business Role and the Application Component used by them.

User-based Collaborative Filtering Recommendation approach ensures that the recommendations should have coverage from all the Apps in the Apps Library based on the usage index by other customers and should not only depend on the top trending Apps.

# Bibliography

B.Thorat, P., M. Goudar, R., & Barve, S. (2015). Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System. *International Journal of Computer Applications*, *110*(4), 31–36. https://doi.org/10.5120/19308-0760

*CRISP-DM - Data Science Process Alliance*. (n.d.). Retrieved August 25, 2021, from https://www.datascience-pm.com/crisp-dm-2/

*Deep Neural Network - an overview | ScienceDirect Topics*. (n.d.). Retrieved August 26, 2021, from https://www.sciencedirect.com/topics/computer-science/deep-neural-network

Geetha, G., Safa, M., Fancy, C., & Saranya, D. (2018). A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System. *Journal of Physics: Conference Series*, *1000*(1). https://doi.org/10.1088/1742-6596/1000/1/012101

*Introduction to recommender systems | by Baptiste Rocca | Towards Data Science*. (n.d.). Retrieved August 25, 2021, from https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada

Rajpurkar, M. S. (2015). *Book Recommendation System*. *1*(11), 314–316.

Singh, R. H., Maurya, S., Tripathi, T., Narula, T., & Srivastav, G. (2020). Movie Recommendation System using Cosine Similarity and KNN. *International Journal of Engineering and Advanced Technology*, *9*(5), 556–559. https://doi.org/10.35940/ijeat.e9666.069520

*Surprise · A Python scikit for recommender systems.* (n.d.). Retrieved August 26, 2021, from http://surpriselib.com/

Virk, H. K., Singh, E. M., & Singh, E. A. (2015). Analysis and Design of Hybrid Online Movie Recommender System. *International Journal of Innovations in Engineering and Technology (IJIET)*, *5*(2), 159–163.

Weng, S. S., & Liu, M. J. (2004). Personalized product recommendation in e-commerce. *Proceedings - 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service, EEE 2004*, 413–420. https://doi.org/10.1109/eee.2004.1287340

*What does RMSE really mean?. Root Mean Square Error (RMSE) is a… | by James Moody | Towards Data Science*. (n.d.). Retrieved August 26, 2021, from https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e

# Appendix

## Plagiarism Report[[1]](#footnote-1)

## Publications in a Journal/Conference Presented/White Paper[[2]](#footnote-2)

## Any Additional Details

1. Turnitn report to be attached from the University. [↑](#footnote-ref-1)
2. URL of the white paper/Paper published in a Journal/Paper presented in a Conference/Certificates to be provided. [↑](#footnote-ref-2)