ABSTRACT

Travelling is an integral part life. Every new itineraries incorporates new emotions from a traveller. People travel to relax and rejuvenate whereas sometimes to forget the past or to kill the recent mishaps or even sometimes to learn from a new realm. Travelling is requisite for living. Nevertheless, single travelling single or alone is an expensive endeavour. Passionate travellers with keen interests crave to congregate and desire to locomote together to explore the world which eventually furnishes more vigour during the travel and yield notable savings in their travel budget. This project '**Sahayaatri'** proposes a mobile application where a traveller is able to find the travellers with same sentiments and mentality like him or herself. Machine Learning based recommendation is to be used in this project which will be effectively suggesting the keen companions to a traveller to travel with. This Machine Learning powered project will be able to predict the next travelling destination based on the recent itineraries of that traveller. Backend of this web system is to be constructed with Java Spring MVC. MySQL will be the database to store the data generated by the travellers.

TABLE OF CONTENTS

ABSTRACT i

TABLE OF CONTENTS ii

LIST OF FIGURES iii

Chapter 1: INTRODUCTION 1

1.1 Background 1

1.2 Problem Statement 2

1.3 Objectives 2

Chapter 2: LITERATURE REVIEW 3

Chapter 3: FEASIBILITY STUDY 7

3.1 Economic Feasibility: 7

3.2 Resource Feasibility: 7

3.3 Technical Feasibility: 7

Chapter 4: PROJECT METHODOLOGY 8

4.1 System Diagram 8

4.1.1 Block Diagram of proposed system 8

4.1.2 Users Perspective 9

4.2 Development model 10

4.3 Data Collection Techniques 11

Chapter 5: IMPLEMENTATION PLAN 12

5.1 Schedule (Gantt chart) 12

5.2 Software Requirements 12

Chapter 6: Expected Outcomes 13

References 14

LIST OF FIGURES

Figure 1 Block Diagram of Proposed System…….8

Figure 2 Users Perspective……..……..…….……9

Figure 3 Development Model…….……..…….…....……10

Figure 4 Gantt Chart……..……..…..……...12

Chapter 1: INTRODUCTION

1.1 Background

Amalgamation of Artificial Intelligence with the tourism industry can have colossal prospect. Unnoticed realm of the tourism sector are revealing due to the marvels of Artificial Intelligence. Tourism industry is contributing 10.4% to the world economy in 2018 and projected to grow to an estimated 3.6% average every year over the next decade (World Travel and Tourism Council, 2018). The number of tourists worldwide has increased rapidly. Over the same 10-year period, Southeast Asia is expected to be the fastest-growing region regarding travel and tourism's contribution to a country's or a region's Gross Domestic Product (GDP). Nepal has huge possibilities of tourism development as we are rich in biodiversity, natural beauty, culture and hospitality.

With the boom in tourism over the last decade, information sources play an important role for tourists when making decisions and selecting destinations. The Internet is now considered to be the tourist's main information source for information on products and services. However, the sheer volume of data on the Internet has made it difficult for tourists to process information, whether in pre-trip planning or when making choices during travel. The travel-planning problem is highly complex, time-consuming, and dynamic as there are many factors involved in the decision-making process. Some of the factors involved in travel-planning include travel budget, number of nights one intends to stay at a given destination, food quality, the number of individuals travelling, transport mode, leisure activities, weather etc.

Recently, tourism has benefited substantially from Information and Communications Technology (ICT), and especially from Internet technology and its applications. Decision support tools, also known as Recommendation Systems (RSs), have been developed to address these concerns. In the tourism field, they are referred to as Tourism Recommendation Systems (TRSs). Tourists and tourism providers can search, select, compare and make decisions almost instantly, and more efficiently than ever. Due to the enormous amount of heterogeneous information available on the Internet and through other information sources, TRSs can act as information filters. Selecting appropriate tourist services to match user preferences is one of the most complex tasks a tourist faces when planning a visit to an unfamiliar city. Even though

search engines provide lists of tourism services, tourists are still overwhelmed with the information on offer. TRSs can be utilised extensively as a means of reducing information overload for tourists. Finding an appropriate group for the trip is also a major problem. TRS can be utilized to solve these issues too.

TRSs can help assist tourists to travel independently to an unfamiliar city, especially as regards searching, selecting and comparing tourism services. Not only can TRSs help travellers when planning their trip, but also during and after a trip, thanks to mobile and wireless communication. A well-developed TRS can suggest appropriate tourism services to tourists without interfering with their privacy and suggest other travel-related products to them.

1.2 Problem Statement

- Finding people who want to travel with the like-minded people is cumbersome.
- Some travel to relax and rejuvenate whereas some travel to forget the past or to escape some tragedy. However, searching people having same sentiments and emotions to travel with is strenuous.
- Getting a credible, secured and reliable mediator to create respective group is not possible.

1.3 Objectives

- To create a web system that finds the travel companions for a traveller who is willing to travel with like-minded people as him/herself.
- To provide a mobile platform (app), based on recent activity of a traveller, which predicts the next destination that a traveller may travel and suggesting the complete itinerary to that specific user(traveller).
- To create a system where a traveller can easily find the Travel Agencies that deliver the same packages (destinations, days of stay, locomotion ways) in which a traveller(user) is interested.

Chapter 2: LITERATURE REVIEW

The face of the travel industry has undergone significant changes over the past decade. Much of these changes are attributed to developments in the information technology, services and applications coupled with the increasing competitive environment. The various sections in the travel industry are affected; the airlines as suppliers, travel agencies as service providers, and the consumers of such services i.e. the travellers, with the changes to the global distribution channels brought about by the generalised use of the internet and related information and 138 communication technologies. There are positive impacts on the travel industry such as improvement in the distribution channels and disintermediation. However, there are also negative impacts on the industry such as price transparency and reduced customer loyalty [1]. Along with this paradigm shift, is the changing consumer purchasing habits when it comes to buying airline tickets, they have the ready option of buying online rather than through the traditional travel agencies which perform that services for the consumers such bookings, issuance of tickets and collecting payments on behalf of the airlines. According to Yoon et.al. (2006), the future of travel agencies may be uncertain. This area of interest has drawn the researcher to examine the underlying factors that affect consumers' attitude and intention toward online airline ticketing.

Traveller decision-making is adaptive process whereby there is a dynamic and an interactive adjustment within their surrounding environment at the destination.

According to Fotis et.al, Social media has an important role in a tourist experience, for they: allow access to more information sources; stimulate the feeling of belonging to a virtual tourist community; and they also promote the storytelling that happens after the trip[2].

Similarly, Airbnb, Inc. is a privately held global company headquartered in San Francisco that operates an online marketplace and hospitality service which is accessible via its websites and mobile apps[3]. Members can use the service to arrange or offer lodging, primarily homestays, or tourism experiences. The company does not own any of the real estate listings, nor does it host events; as a broker, it

receives commissions from every booking. The insights of the technology used in Airbnb are described below[4].

Framework - Ruby on Rails. RoR framework is well known for its awesome capabilities that speed up development and, as a result, reduce costs and TTM (time to market).

JavaScript framework - React. A JavaScript UI library, React, is a flexible and efficient solution for building sleek user interfaces.

Web server - Nginx. Nginx is a powerful HTTP and proxy server that speeds up content delivery, ensures Airbnb's security and scalability. Not only Airbnb reaps the benefits of Nginx; tech industry giants like Instagram, Netflix, and Zappos use NGinx too.

Key-value storage - Redis. Redis provides a scalable cache infrastructure and a key/value database.

Cloud storage - Amazon S3, EBS. To store user data including millions of user pictures, Amazon resorts to Amazon services.

Cloud hosting - Amazon EC2. Amazon EC2 is an efficient tool that distributes the incoming traffic and doesn't let Airbnb's system go down during sudden traffic spikes or any unexpected traffic fluctuations.

Cloud database - Amazon RDS. Amazon keeps its data in an Amazon's cloud relational database. Earlier, Amazon used MySQL databases, but switched to Amazon RDS to simplify administration and other routine tasks.

Big Data tools - Presto, Druid, Airpal. Airbnb possesses a tremendous volume of user data, so they use various instruments to store, process, analyze, and manage that data.

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to progressively improve their performance on a specific task. Machine learning algorithms build a mathematical model of sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task.[1][2]:2 Machine learning algorithms are used in the applications of email filtering, detection of network intruders, and computer vision,

where it is infeasible to develop an algorithm of specific instructions for performing the task. Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a field of study within machine learning, and focuses on exploratory data analysis through unsupervised learning.[3][4] In its application across business problems, machine learning is also referred to as predictive analytics.

Web scraping a web page involves fetching it and extracting from it.[1][2] Fetching is the downloading of a page (which a browser does when you view the page). Therefore, web crawling is a main component of web scraping, to fetch pages for later processing. Once fetched, then extraction can take place. The content of a page may be parsed, searched, reformatted, its data copied into a spreadsheet, and so on. Web scrapers typically take something out of a page, to make use of it for another purpose somewhere else. An example would be to find and copy names and phone numbers, or companies and their URLs, to a list (contact scraping).

Web scraping is used for contact scraping, and as a component of applications used for web indexing, web mining and data mining, online price change monitoring and price comparison, product review scraping (to watch the competition), gathering real estate listings, weather data monitoring, website change detection, research, tracking online presence and reputation, web mashup and, web data integration.

Web pages are built using text-based mark-up languages (HTML and XHTML), and frequently contain a wealth of useful data in text form. However, most web pages are designed for human end-users and not for ease of automated use. Because of this, tool kits that scrape web content were created. A web scraper is an Application Programming Interface (API) to extract data from a web site. Companies like Amazon AWS and Google provide web scraping tools, services and public data available free of cost to end users.

Chapter 3: FEASIBILITY STUDY

3.1 Economic Feasibility:

This project requires the online money payment system. From the technical grounds in Nepal, digital payment companies like Khalti and eSewa are such providers who have the price of quotations of Rs 25,000 and Rs 20,000 respectively to get their authenticated API token or simply to call their API's from any web-app.

3.2 Resource Feasibility:

It can be said that almost all people travel in some points of their life. Tourism Ministry of Nepal published the statistical report showing 9, 40,218 tourists travelled to Nepal staying 12.6 days on average in 2017. It manifests that collecting a traveller's data and their would-be companion's data is feasible in our system.

3.3 Technical Feasibility:

Even major functionalities like web-scrapping and machine learning based recommendations can be developed to perfection with common technologies. The libraries, frameworks and the databases to be used in this project are open-sourced, whose entire features are free to use. Therefore, this project is considered technically feasible.

Chapter 4: PROJECT METHODOLOGY

4.1 System Diagram

4.1.1 Block Diagram of proposed system

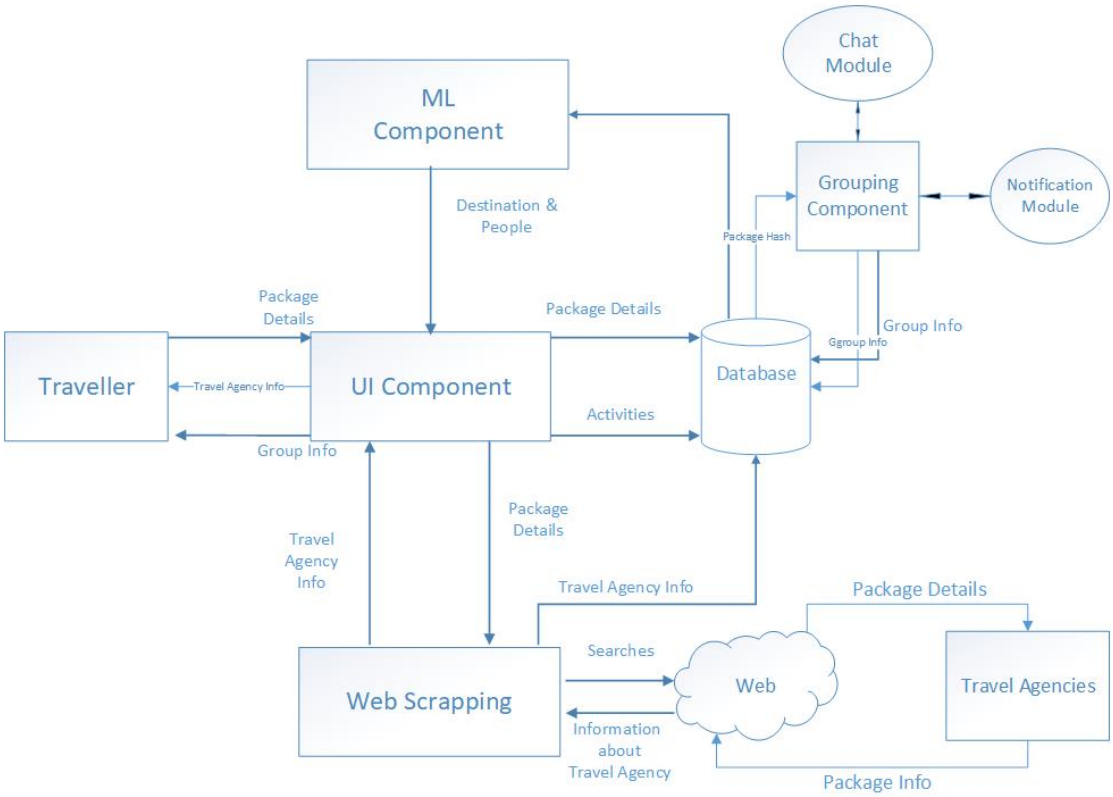

Figure 1: Block Diagram of proposed system

4.1.2 Users Perspective

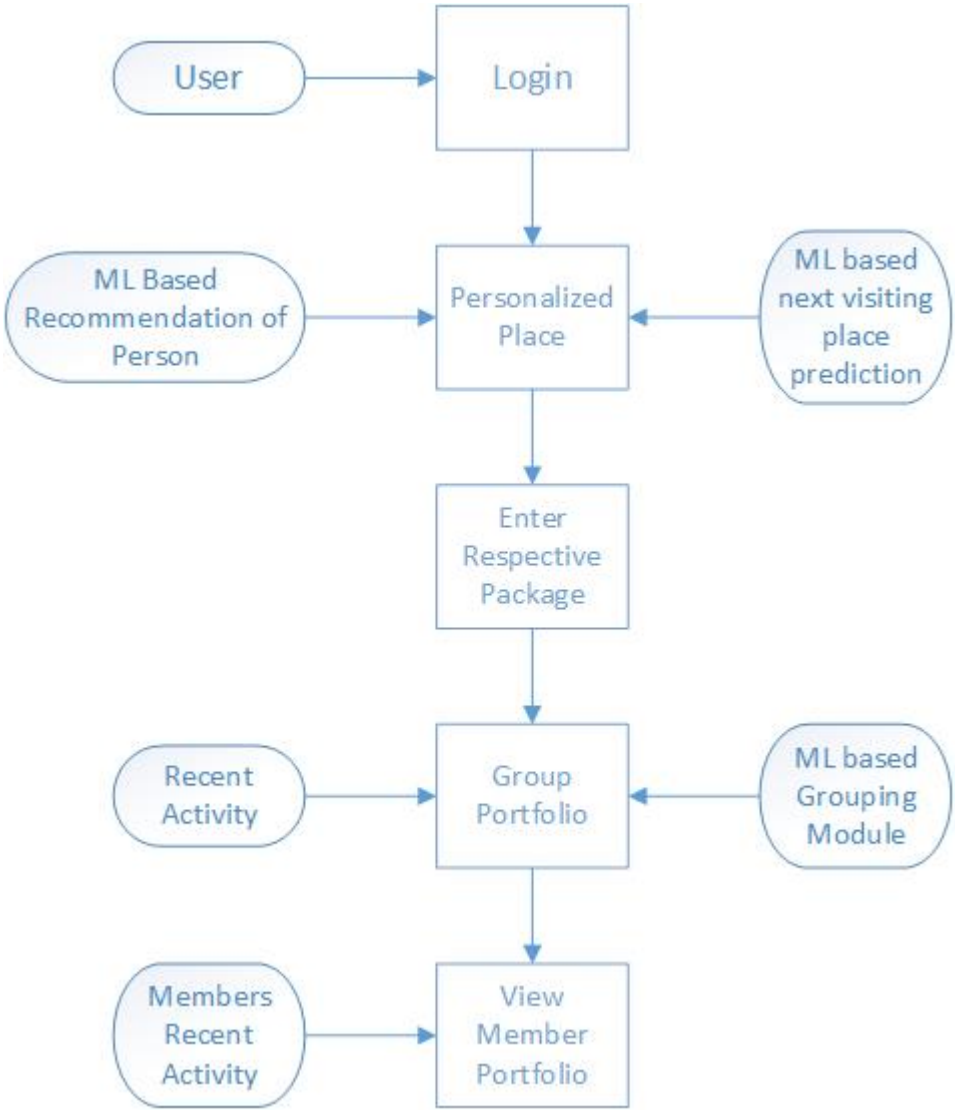

Figure 2: Users perspective of proposed system

4.2 Development model

The development model that we are going to use is a Prototyping model. The Prototyping Model is a systems development method (SDM) in which a prototype (an early approximation of a final system or product) is built, tested, and then reworked as necessary until an acceptable prototype is finally achieved from which the complete system or product can now be developed. This model works best in scenarios where not all of the project requirements are known in detail ahead of time. It is an iterative, trial-and-error process that takes place between the developers and the users.

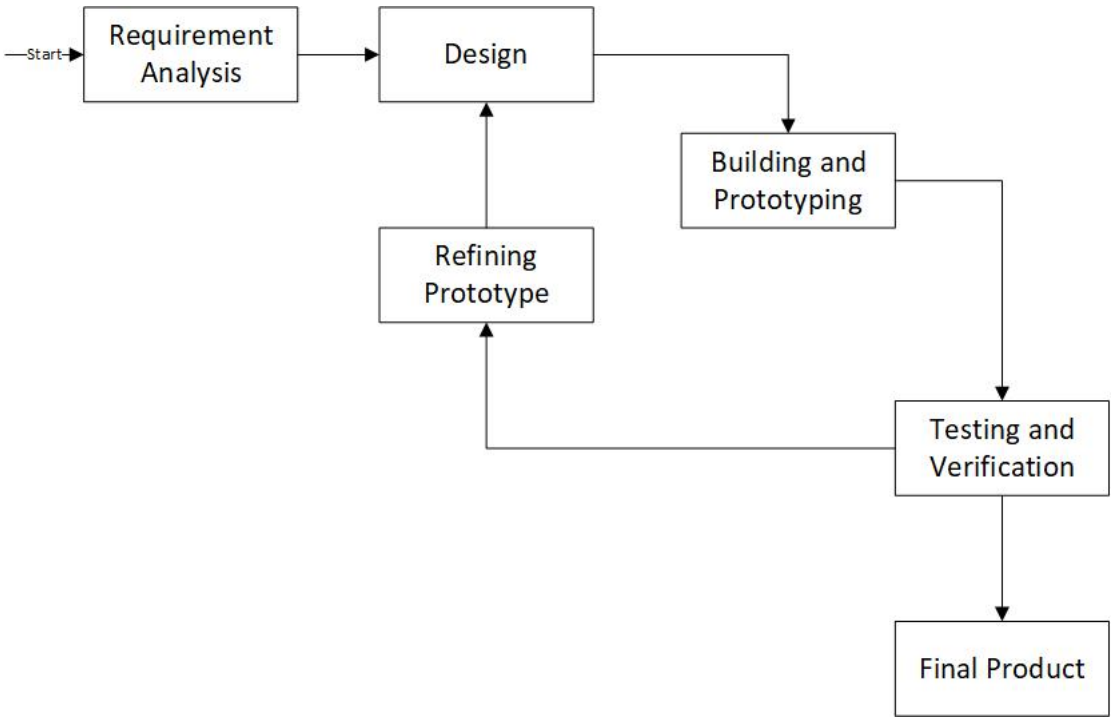

Figure 3: Prototype Model

4.3 Data Collection Techniques

Recommendation systems are one of the most common, easily comprehendible applications of big data and machine learning. Among the most known applications are Amazon's recommendation engine that provides us with a personalized webpage when we visit the site, and Spotify's recommendation list of songs when we listen using their app[5].

For the recommendation, system to fully function data is the most essential part as the data helps for the prediction and recommendation.

Data Collection is the first and most crucial step for building a recommendation engine. The data can be collected by two means: explicitly and implicitly. Explicit data is information that is provided intentionally, i.e. input from the users such as movie ratings. Implicit data is information that is not provided intentionally but gathered from available data streams like search history, clicks, order history, etc. [6]

Initially, with the help of google form, we can collect the primary data from different people of different age groups, for the data set explicitly. Then the data set can be splitted into training and testing set. Splitting the data into training and testing sets is an important part of evaluating predictive and recommendation modeling. The idea of it is to train the recommendation system with the larger portion of the data for training , which may include certain fields (such as age, gender, destinations, etc.) and, accordingly with the help of testing set of data we could be able to test the machine for its accuracy and performance. Since this project has a cold start problem it can be reduced when attribute similarity is taken into account. Cold start problem is not having enough historical interactions. You can encode attributes into binary vector and feed it to recommender.

Items clustered based on their interaction similarity and attribute similarity are often aligned. You can use neural network to predict interaction similarity from attributes similarity and vice versa. [7]

Chapter 5: IMPLEMENTATION PLAN

5.1 Schedule (Gantt chart)

ID	Task Name	Start	Finish	Duration
1	Project Analysis	12/3/2018	12/14/2018	2w
2	Requirement Analysis	12/12/2018	1/4/2019	3.5w
3	Designing	1/3/2019	3/13/2019	10w
4	Development	1/21/2019	7/12/2019	25w
5	Testing and Verification	5/15/2019	7/29/2019	10.8w
6	Documentation	12/3/2018	7/30/2019	34.4w

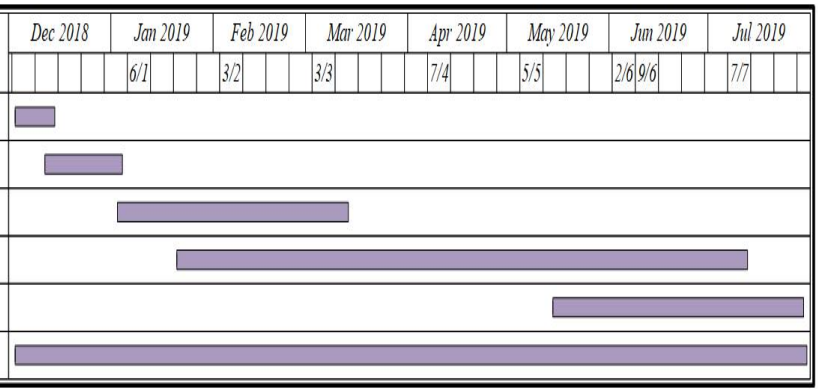

Figure 2: Gantt chart

5.2 Software Requirements

- Java Framework Spring
- MongoDB
- Python
- JavaScript

Chapter 6: Expected Outcomes

- Web system that finds the travel companions for a traveller who is willing to travel with like-minded people as him/herself.
- Based on recent activity of a traveller, which predicts the next destination that a traveller may travel and suggesting the complete itinerary to that specific user (traveller).
- A system where a traveller easily finds the Travel Agencies that deliver the same packages (destinations, days of stay, locomotion ways) in which a traveller (user) is interested.

References

[1]Daniele, R. and Frew, A. (2004), “From intermediaries to market-makers: an analysis of the evolution of e-mediaries”, in Frew, A. (Ed.), Information and Communication Technologies in Tourism 2004, Springer, Wien, pp. 546-57.

[2]Fotis, J., Buhalis, D., & Rossides, N. (2012). Social Media Use and Impact during

the Holiday Travel Planning Process. Information and Communication

Technologies in Tourism 2012. In M. Fuchs, F. Ricci & L. Cantoni (Eds.),

(pp. 13-24): Springer Vienna.

[3] Julia Tokareva, J. (2019). What Technology Stack Does Airbnb Use?. [online] Forbes.com. Available at: https://www.forbes.com/sites/quora/2018/02/20/what-technology-stack-does-airbnb-use/#5fcf73d04025 [Accessed 22 Jan. 2019].

[4] Airbnb Press Room. (2019). About Us - Airbnb Press Room. [online] Available at: https://press.airbnb.com/en-uk/about-us/ [Accessed 22 Jan. 2019].

[5] M. Tjokro, “How to Build a Recommendation System for Purchase Data (Step-by-Step),” medium.com, 15-Oct-2018. [Online]. Available: https://medium.com/datadriveninvestor/how-to-build-a-recommendation-system-for-purchase-data-step-by-step-d6d7a78800b6. [Accessed: 21-Jan-2019].

[6] P. Sharma, “Comprehensive Guide to build Recommendation Engine from scratch,” Analytics Vidhya, 28-Jun-2018. [Online]. Available: https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-recommendation-engine-python/. [Accessed: 21-Jan-2019].

[7] P. Kordík, “Machine Learning for Recommender systems - Part 1 (algorithms, evaluation and cold start),” medium.com, 03-Jun-2018. [Online]. Available: https://medium.com/recombee-blog/machine-learning-for-recommender-systems-part-1-algorithms-evaluation-and-cold-start-6f696683d0ed. [Accessed: 21-Jan-2019].