Bilkent University

Department of Computer Engineering

# Senior Design Project

*TechRank*

# Project Specifications Report

Serhat Hakkı Akdağ, Alperen Ustaömer, Mehmet Oğuz Göçmen, İlhami Kayacan Kaya, Pelin Elbin Günay

Supervisor: Çiğdem Gündüz Demir
Jury Members: Mustafa Özdal and Selim Aksoy

**Contents**

Project Specifications Report

*TechRank*

## 1   Introduction

Online shopping is one of the rapidly growing trends in the last decade. People think that doing shopping with using websites or shopping apps is easier and less time consuming than doing shopping in the stores. Because, they can search the products they want to buy by using search bar without losing time while finding the products in the store. Customers are surfing on e-commerce websites not only for buying the products online but also, even when people are buying things from stores, they are checking out reviews and comments of products online. In any case, checking out reviews is quite common nowadays, as it is hugely beneficial. However, people do not want to read hundreds of comments due to timing constraints. In addition, it is not always easy to understand general consensus on comments because there are inconsistencies among comments. Because of that, seeing the products' rates that are constituted by analyzing all comments is beneficial for the user. By observing rates about the products, customers can easily decide which product is worth buying.

When we consider the common usage of online shopping, we decided to develop a web application in order to help people to decide and find out which technological product is beneficial and necessary for them by analyzing user comments from trusted e-commerce websites and by rating and ranking them into some related categories. Our application will analyze all comments scattered on web, then it will rank the devices on the same category using criteria. TechRank will also decide on overall rating for the device and overall rating for the producer company of the device.

### 1.1   Description

TechRank is a web application which provides comment-based ratings and rankings for electronic devices such as TVs, computers, mobile phones, audio devices, appliances, etc. Our system will collect reviews on electronic products from the trusted e-commerce websites. It will analyze the comments using Natural Language Processing (NLP) and based on the analysis it will provide following functionality: Maximum 4 criteria for each product type will be pre-selected. For each one of these criteria, TechRank will assign a percentage point by analyzing comments on specific criterion. For example, if the criterion is performance for computer, TechRank will analyze comments on this device that includes negative or positive statements on performance of this device, and determine percentage point. After this, ranking of the device amongst all other devices on its category regarding selected criterion will be determined. This ranking will show the selected products' place on the market based on only

selected criterion. In addition, overall rating for the device will be assessed by averaging percentage points for each of the criteria. And by averaging overall ratings of all products of the producer, overall rating of the producer will be determined. Using all these rankings, users will be able to understand general consensus on products and their specific qualities, and the producers. By this way, users will not get lost among hundreds or thousands of comments in the web, and will be able to understand the rankings of the product they are looking for amongst all of its rivals on the market. Furthermore, users will be able to view all the comments that are used to determine percentage point of the device on specific criterion.

In addition to this core feature, to provide better experience for users many information on selected device will be made available on TechRank. Application will show pictures of the product from different e-commerce website, it will provide technical specifications of the device; users will be able to see all the comment on web for selected device.

## 1.2    Constraints

The constraints for our project are as follows:

### 1.2.1 Economic Constraints

For the website, hosting service and domain name are required that both of them cost a fair amount per year. For data storing and comment analyzing, AWS will be used. Price information for this service will be determined after understanding how much data will be stored and need to be maintained. For issue tracking a private repository will be opened on GitLab without any fee.

### 1.2.2 Time Constraints

Required Reports of this project must be delivered without any latency. Also, the implementation of the TechRank should be completed as soon as possible to be able to find enough opportunity to review this project again and to eliminate deficiencies.

### 1.2.3 Implementation Constraints

For providing better organization inside our project team, GitLab will be used as issue tracking system in order to efficiently combine all of the work done by each member. In this way, each member in the group will be aware of the general situation of the project and will be able to access the required parts easily.

For the implementation of the Backend part of the project Spring Boot will be used. For the frontend part of the project, JavaScript-based Angular.js framework will be used.

For the data storage of the system we are planning to use S3 (Simple Storage Service) of the Amazon Web Services, since it provides huge amount of free data storage. Also, during the implementation of this system, we need tools that are providing data retrieval and cloud

computing for TechRank services. In order to benefit from cloud computing, EC2 service of the AWS will be used. For data retrieval, AWS Lambda functions will be used. In addition to these services, in the case of some problems, Python packages can be used.

Database of the application will be provided by AWS RDS (Relational Database Service) based on MySQL and an additional NoSQL database, Mongo DB. However, if real time data views and a document database will be necessary, we could use Firebase real time database.

### 1.2.4 Social Constraints

There will be no direct social interaction between users in the TechRank. However, there will be a suggestions and advices part for our web page where users can communicate with us by writing their suggestions to us about how we should improve our web page and our service to users.

### 1.2.5 Sustainability Constraints

TechRank will keep comments up-to-date and web servers of TechRank will be updated incessantly without allowing problems on them. A cron job will take care of this updating process. In order to be able to keep the clients' interests, products that exist on the database of TechRank will be updated frequently. By this way users will have easy access to comments and evaluations about the latest technology products. Also, criteria that decided for every single product type should be updated according to these technological developments.

In addition, in order to reach data from online shopping web pages in a sustainable way, we should update our online listeners according to the changings on these webpages to be able to reach these data whenever it required.  By this way we will prevent possible data loss because of the updates on shopping web pages.

### 1.2.6 Language Constraints

The main language of TechRank will be English due to the fact that English is very commonly used for writing comments for products on popular shopping web pages from all around the world. By this way we have access to a lot of data written in English. For now, we do not have any plan about adding other languages to TechRank since it requires changings not only on interface but also all of the algorithms. However, after completing our implementation, this application can also be translated to other languages with extra work.

### 1.3    Professional and Ethical Issues

Developing a yearlong project requires being an excellent group worker and communicating in a professional manner such that every member will act respectfully to each other. Being a part of the developer team will require from members to manage their time accordingly such that every team member will be equally responsible from every stage of this

project. This way there won't be any irrevocable conflicts between team members that will disturb the ambiance of developer team.

Looking to application process, so much data will be scraped and crawled and then analyzed through NLP. When the analysis process is started, location data and account data will also be available for analyzing. However, this sensitive data will not be analyzed or used in any way. Moreover, since all comments for a product will be displayed; there may be some comments which contain insensitive words. While analyzing a comment; if system comes across such words, that comment will not be displayed to user. Also, analysis of these results will not be shared with third parties.

In sentiment analysis process with NLP; we will pre-determine maximum of 4 categories to put products into, this process will be manual. However, during sentiment analysis process there won't be any human factors that interfere in the decisions of our analysis algorithms. Moreover, during a user's session in our application, we will not try to analyze the behavior of user and rank the products we display by applying user-specific ratings.

Lastly, during the development and deployment processes of the project, the team will exhibit the standards of Code of Ethics [1] for Engineers.

## 2. Requirements

Functional and non-functional requirement for TechRank are as follows:

### 2.1 Functional Requirements

- User will be able to see the overall rating score and ranking of a device that he/she is looking for. The ranking of the device is based on the overall score that the device has among the devices that belong to the same category with it, such as laptops, headphones, desktops, cameras, etc. By means of clicking to overall score of the current device that the user is looking at, user can navigate to the ranking list of the devices that belongs to the same category of the current device, and then see the details of the devices in the ranking list by clicking and navigating to their product details page.

- User will be able to see the rating score and ranking of the device under predefined criteria categories such as durability, usability, reliability etc. Each criterion has a ranking list constructed according to the rating scores of the devices. User is able to observe these rating lists by means of clicking each criterion of the current device and navigate to the ranking list that

corresponds to chosen criterion. Then, user can navigate to product details pages of the devices by means of clicking them from the rating list.

- User will be able to see the overall rating of the brand that the device he/she is looking for belongs to. By means of clicking to the rating of the brand on the current device's page, user can navigate to the rating list of the brands that has been constructed by the overall scores of the brands. Overall brand ratings will be done according to the overall ratings of the devices that belong to that particular brand. User will be able to see the overall rating of the brand that the product it is looking for belongs to.

- User will be able to see the comments about the device under each criterion section by means of clicking to each criterion. In addition to the rating list of the criterion, there are also the comments that caused a rating score of that device on chosen criterion.

- User will be able to search any technological product. However, if a product does not contain enough comments, it will not be displayed to user to prevent ambiguity.

- User will be able to see pictures of the device that are collected from different sources and navigate through them while browsing a product.

- User will be able to view technical specifications of the device.

- User will be able to view all the comments about product collected from different sources on web.

- User will be able to give feedback to the ratings that have been made for the device that he/she is looking at.

## 2.2 Non-functional Requirements

- System will be function as a web application.

- System will require internet connection.

- System will function 24/7.

- Data will be updated within some period of time constantly.

- Gathering data for the products will be conducted periodically, after collecting comment data, sentiment analysis with NLP will be initiated only once for these periods. Starting and ending times for these periods will be predefined.

- The query responses will be fast enough (around a second or two) so that user will not wait too long to see details about a particular query that has been made through search bar.

- Web user interface will be user friendly to navigate people to their desired pages easily.

- User feedback will be collected about the ratings, so that system will be optimized to act more quality and user-centered.

**References**

[1]     "Code of Ethics," NSPE Code of Ethics for Engineers | National Society of Professional Engineers. [Online]. Available: https://www.nspe.org/resources/ethics/code-ethics. [Accessed: 15-Oct-2018].