T.C.

GEBZE TECHNICAL UNIVERSITY

FACULTY OF ENGINEERING

DEPARTMENT OF COMPUTER ENGINEERING

MINIMUM K-CHINESE POSTMAN PROBLEM

AKİF KARTAL

SUPERVISOR
PROF. DR. DİDEM GÖZÜPEK

GEBZE
2022

T.C.
GEBZE TECHNICAL UNIVERSITY
FACULTY OF ENGINEERING
COMPUTER ENGINEERING DEPARTMENT

MINIMUM K-CHINESE POSTMAN PROBLEM

AKİF KARTAL

SUPERVISOR
PROF. DR. DİDEM GÖZÜPEK

2022
GEBZE

This study has been accepted as an Undergraduate Graduation Project in the Department of Computer Engineering on 16/06/2022 by the following jury.

JURY

Member
(Supervisor)   :   PROF. DR. DİDEM GÖZÜPEK

Member   :   YRD.DOÇ.DR. ZAFEİRAKİS ZAFEİRAKOPOULOS

ABSTRACT

In this project, we present a heuristic and exhaustive search algorithm for the minimum k-Chinese postman problem. We considered the minimum k-Chinese postman problem is, given a multigraph $G = (V, E)$ initial vertex $s \in V$ length $l(e) \in N$ for each $e \in E$ the *minimum k-Chinese postman problem* is to find k tours(cycles) such that each containing the initial vertex *s* and each edge of the graph has been traversed at least once and the most expensive tour is minimized. This problem is NP-hard and we tried to solve it with a polynomial-time algorithm. For this purpose, we created one polynomial-time algorithm and one exponential-time algorithm and we made a complexity analysis for these algorithms. After creating algorithms we compare them with different parameters by looking at the results and running time. We saw that when the k value is increasing, the heuristic algorithm produces better results in a very short time.

Keywords: heuristic, exhaustive search, NP-hard, polynomial-time, exponential-time, parameters, running time, complexity analysis

ÖZET

Bu projede minimum k-Chinese postman problemi için sezgisel ve kapsamlı arama algoritması sunuyoruz. Minimum k-Chinese postman probleminin tanımı şu şekildedir. Verilen bir multigrafda s adında bir başlangıç noktası vardır ayrıca her 2 nokta arasındaki yol için bir yol uzunluğu vardır. Bu probleme göre bizim amacımız, bu graf üzerinde öyle bir k tane tur ya da dolaşım bulacağız ki her bir turda başlangıç noktası olacak ve graf'taki her bir yoldan en az bir kere geçilmiş olacaktır. Bizim amacımız bu k tane turdaki en büyük uzunluğu sahip turun uzunluğunu en aza indirmektir. Bu problem bir NP-hard problemdir ve biz bu problemi bir polinom zamanlı algoritma ile çözmeye çalıştık. Bu amaç için bir tane polinom zamanlı bir tanede üstel zamanlı olmak üzere iki tane algoritma geliştirdik ve bu algoritmaların karmaşıklık analizini yaptık. Bu algoritmaları oluşturduktan sonra farklı parametrelerle test edip birbirleri ile sonuç ve çalışma süresi bakımından karşılardık ve sonuç olarak gördük ki k değerini artırdıkça çok kısa bir sürede sezgisel algoritma daha iyi sonuçlar buluyor.

Anahtar Kelimeler: sezgisel, kapsamlı arama, NP-hard, polinom zamanlı, üstel zamanlı, çalışma süresi, karmaşıklık analizi

ACKNOWLEDGEMENT

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation	:	Explanation
GUI	:	Graphical User Interface
CPP	:	Chinese Postman Problem
MIN	:	Minimum
MAX	:	Maximum

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

A graph is a data structure composed of a set of objects (nodes) equipped with connections (edges) among them. Graphs can be directed if the connections are oriented from one node to another (e.g. Alice owes money to Bob), or undirected if the orientation is irrelevant and the connections just represent relationships (e.g. Alice and Bob are friends). A graph is said to be complete if all nodes are connected to each other. A directed graph with no loops is said to be acyclic. A few practical examples of graphs are friendship networks (e.g. on social media), genealogical (family) trees, molecules, particles produced at the Large Hadron Collider, and a company's organizational chart.[1]

To find the most efficient way of traversing an entire graph is a widely spread problem in today's society. For a snow truck to plow all snow on every street in a town in the minimal consumed time is only one of a vast amount of applications of this problem. Finding solutions for this problems would lead to both a financial and an environmental improvement to companies and cities all over the world.[2]

1.1. Project Definition

In this project, We tried to solve Minimum k-Chinese Postman Problem which is given a multigraph $G = (V, E)$ initial vertex $s \in V$ length $l(e) \in N$ for each $e \in E$ the *minimum k-Chinese postman problem* is to find *k* tours(cycles) such that each containing the initial vertex *s* and each edge of the graph has been traversed at least once and the most expensive tour is minimized.

To solve this problem, we have implemented 2 different algorithms and evaluated them in different perspectives.

The problem has following inputs.

- *s*, initial vertex

- *k*, given positive number

- *l(e)*, length for each edge

- *n*, number of vertices(nodes)

- *e*, number of edges
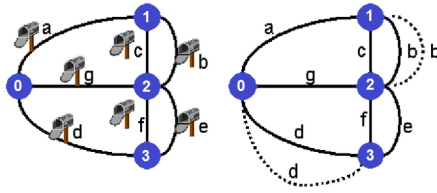
(a) Simple Multigraph Example   (b) Postmans

1.2. The Goal of the Project

Making reason for this project is to implement a heuristic algorithm for the Minimum k-Chinese Postman Problem from Literature.

In Literature, there are a reasonable number of algorithms for this problem but finding implemented one is nearly impossible because these types of algorithms are NP-hard.

Also, People can see how to solve these types of problems, how to make a complexity analysis, and how to make a comparison between different algorithms.

Lastly, after publishing this project people can use these solutions and improve them.

2. Project Design and Details

In order to make this project, we need to consider and determine details about this project.

2.1. Project Design Plan

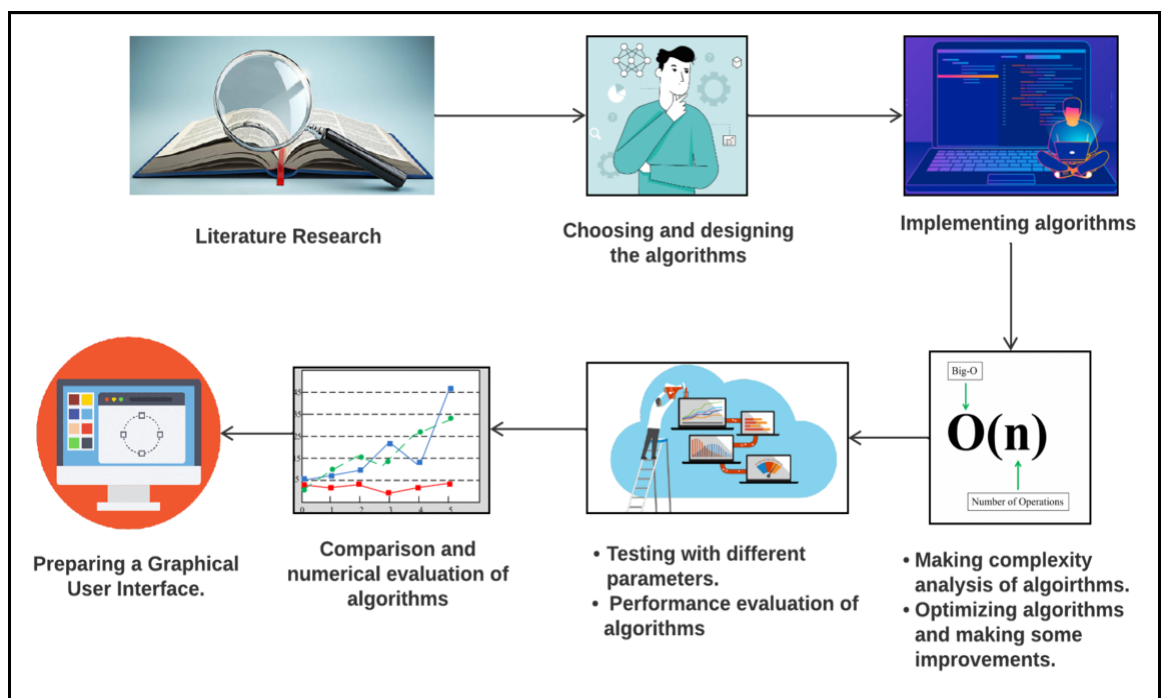In the following image you can see the project design plan in a good manner.

Figure 2.1: Project Design Plan

2.2. Project Requirements

- Making literature research and understanding the problem.

- Choosing and designing algorithms.

- Implementing both heuristic and exhausted search algorithms.

- Making complexity analysis of the algorithms.

- Testing with different parameters and performance evaluation of algorithms.

- Comparison and numerical evaluation of algorithms.

- Showing the comparison average results on the charts.

- Preparing a GUI and running algorithm on that GUI.

2.2.1. Literature Research

In order to solve this problem we have to make a deep research and reading in literature. For Minimum k-CPP, I have read following articles.

Author	Article
Dino Ahr, Gerhard Reinelt	New heuristics and lower bounds for the min-max k-chinese postman problem[3]
Kaj Holmberg	Heuristics for the weighted k-Chinese/rural postman problem with a hint of fixed costs with applications to urban snow removal[4]
G. Gutin, G. Muciaccia	Parameterized Complexity of k-Chinese Postman Problem[5]
Anton Hölscher	A Cycle-Trade Heuristic for the Weighted k-Chinese Postman Problem [2]
Dino Ahr, Gerhard Reinelt	A tabu search algorithm for the min–max k-Chinese postman problem[6]

Table 2.1: Found and Read Articles while Researching

2.2.2. Tools and Technologies

In order to make this project following tools and technologies have been used.

Figure 2.2: Tools and Technologies

1. **Python 3.9:** This is used as as programming language to implement algorithms and gui.

2. **Windows 10:** This is used as operating system.

3. **PyCharm IDE:** This is used to as development environment.

4. **igraph:** This python library is used to generate and draw graphs.

5. **PyQt5, Qt Designer:** These are used to make graphical user interface.

6. **Git and Github:** These are used to keep source code.

3. Heuristic Algorithm

In this module, I tried to determine whether a dormitory comment is positive or negative. In order to do this, I have to use some Natural Language Processing methods. In my solution, in order to create word vectors, I have used TF-IDF(term frequency-inverse document frequency) vectorizer and Logistic Regression as a Classifier from python scikit-learn library. In order to complete this module following steps have been applied.

1. Find a good dataset

2. Clear data and remove noise in the dataset

3. Split train and test data

4. Create and train a model by using TF-IDF vectorizer and Logistic Regression

5. Test that model accuracy

6. Export and deploy the model

Next we will see these steps in detailed way.

3.1. Algorithm Steps

As we know our comments will be in Turkish. But there was no good dataset in Turkish for dormitory comments all we have are movie reviews or product reviews in Turkish. Therefore while I was searching for a good dataset I have found the this dataset.[1] In this English dataset, we have 35038 hotel comments dataset with the response(positive or negative). For this module, I will use this dataset by translating Turkish comments into English.

3.2. Algorithm implementation Details

After finding the dataset, I can create and train our model. Before that, I need to clear and remove noise in the data. In the following code pieces, we can see these steps.

---

[1]https://www.kaggle.com/anu0012/hotel-review?select=train.csv

```
1 from sklearn.feature_extraction.text import TfidfVectorizer
2 from sklearn.linear_model import LogisticRegression
3 from sklearn.pipeline import Pipeline
4
5 tvec = TfidfVectorizer(ngram_range=(1, 1))
6 clf2 = LogisticRegression()
7
8 model = Pipeline([('vectorizer',tvec)
9                 ,('classifier',clf2)])
10
11 model.fit(attribute_train, target_train)
```

Listing 3.1: Creating Model

In following code piece, I will test my model by using Confusion Matrix. In this test my test data will be %10 of train data. For more information check chapter **??**.

```
1 from sklearn.metrics import confusion_matrix
2 from sklearn.metrics import accuracy_score
3
4 verdict = model.predict(attribute_test)
5 confusion_matrix(verdict, target_test)
6
7 print("Accuracy : ", accuracy_score(verdict, target_test))
```

Listing 3.2: Testing Model with Confusion Matrix

3.2.1. Sorting the edges

In order to use this model in my mobile application I need deploy it to the public so that my application can use it. In order to do this first I need to export my model into a file by using python **pickle**[1] library.

```
1 import pickle
2 path='/content/drive/MyDrive/Colab Notebooks/model2/'
3 pickle.dump(model, open(path+'model.pkl','wb'))
```

Listing 3.3: Exporting Model Using Pickle into a File

---

[1]https://docs.python.org/3/library/pickle.html

After exporting our model I can use that model in our API. In the following code piece we can see how I am using the exported model file to integrate with my mobile application.

3.2.2. Creating the Closed Walks

```python
import pickle
from flask import Flask
from flask import jsonify
from flask import request
from googletrans import

app = Flask(__name__)
model = pickle.load(open('model.pkl', 'rb'))
translator = Translator()

@app.route("/api/v1/getResult", methods=['POST'])
def getresult():
    request_data = request.get_json()
    comment = request_data['comment']
    com_en = translator.translate(comment, dest="en").text
    com_en = com_en.replace("country", "dormitory")
    com_en = [com_en]
    result = model.predict(com_en)
    return jsonify({'result': result[0]})
```

Listing 3.4: Use of Exported Model to Evaluate Translated Comment in API

3.2.3. Adding Dummy Tours

In order to implement this sentiment analysis module, I have benefited from these resources.[7]–[12]

3.2.4. Merging Tours

4. Exhaustive Search Algorithm

In order to store data, we need a database in this project. But the important thing is this database should be **reachable** from anywhere which means it shouldn't be a local database on any computer. Therefore cloud databases is good for this project. In-state of the art **Google firebase platform** is mostly used for this purpose because it has powerful options and is free for the starter package. Firebase works with two different databases. The Real-Time Database is the original Firebase database product, and Cloud Firestore is a new and improved version of the Real-Time Database.[1] In this project, we have used the Cloud Firestore database of Firebase.

4.1. Algorithm Steps

One of the requirements of this project collecting data about dormitories that are around Gebze Technical University. In order to do this, I have collected data by calling each dormitory and talking with the responsible person about that dormitory. In this process, **some of the people avoided giving information about that dormitory.** Therefore there can be some missing and wrong information about dormitories.

4.2. Algorithm implementation Details

Cloud Firestore is a NoSQL document database that lets you easily store, sync, and query data for your mobile and web apps at a global scale.[2]

4.2.1. Finding All Cycles

Below we have the tables and relations between them in the figure 4.1. As a relation, we have one-to-one and one-to-many relations between tables. Also in tables, most data types are the **string to manage** data easily in the application.

We can see real data in the figure 4.2. On the left side, we have 4 database tables on the center, we have each separate data in table with its unique id on the right side, we have the context of the data.

---

[1]https://blog.back4app.com/firebase-vs-firestore/
[2]https://firebase.google.com/products/firestore
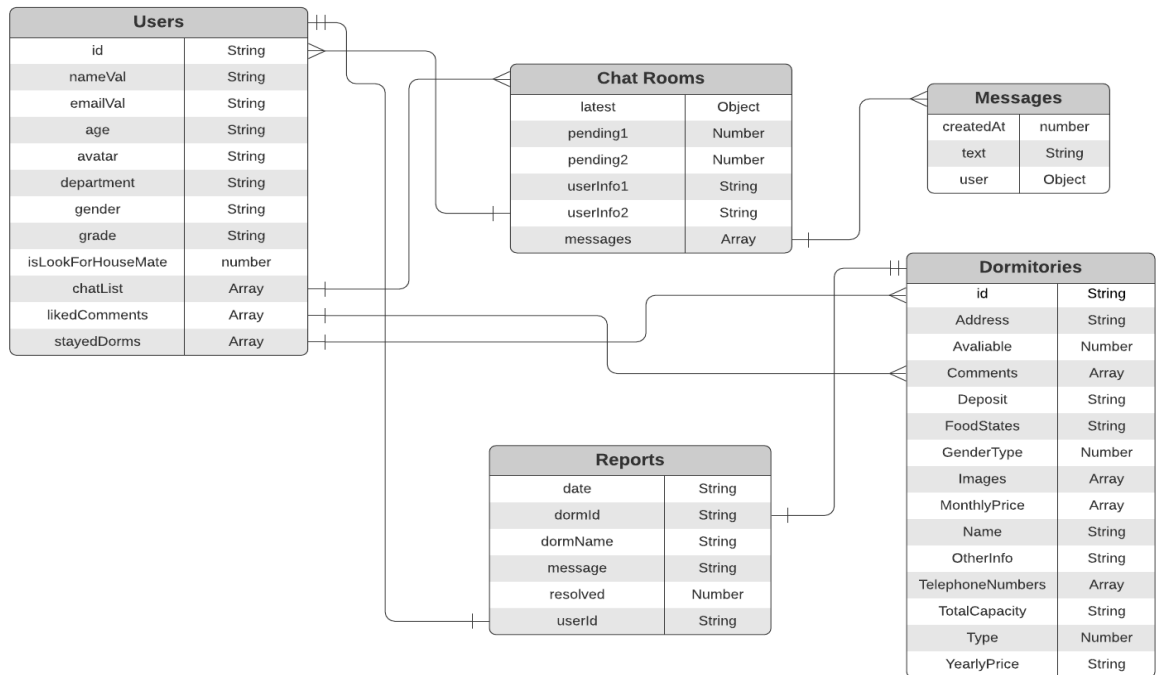
Figure 4.1: Table Relations in Database

Figure 4.2: Real Data and Tables in Firestore Database

4.2.2. Finding All k Combinations

4.2.3. Finding All Proper Cycles

4.2.4. Choosing the Optimal Cycle

5. Complexity Analysis of Algorithms

Making the mobile application is the most involved part of this project and to make mobile application I have used react native technology which is one of the heavily used technology in business.

One of the critical part of react native is **you need to make everything with code** which means **there is no drag and drop** feature for the components such as buttons etc.

Figure 5.1: Mobile Application Logo

5.1. Complexity Analysis of Heuristic Algorithm

Next, we will see each screen and features of the application one by one.

5.2. Complexity Analysis of Exhaustive Search Alg.

6. Performance Evaluation of Algorithms

Making the mobile application is the most involved part of this project and to make mobile application I have used react native technology which is one of the heavily used technology in business.

One of the critical part of react native is **you need to make everything with code** which means **there is no drag and drop** feature for the components such as buttons etc.

Figure 6.1: Mobile Application Logo

6.1. Performance Evaluation of Heuristic Algorithm

Next, we will see each screen and features of the application one by one.

6.2. Performance Evaluation of Exhaustive Search Alg.

7. Comparison and Numerical Evaluation

Making the mobile application is the most involved part of this project and to make mobile application I have used react native technology which is one of the heavily used technology in business.

One of the critical part of react native is **you need to make everything with code** which means **there is no drag and drop** feature for the components such as buttons etc.

Figure 7.1: Mobile Application Logo

7.1. Running Time Comparison

Next, we will see each screen and features of the application one by one.

7.2. Maximum Length Comparison

8. Graphical User Interface

Making the mobile application is the most involved part of this project and to make mobile application I have used react native technology which is one of the heavily used technology in business.

One of the critical part of react native is **you need to make everything with code** which means **there is no drag and drop** feature for the components such as buttons etc.

Figure 8.1: Mobile Application Logo

8.1. User Interface Design and Results

Next, we will see each screen and features of the application one by one.

9. Success Criteria

For this project, we have determined 4 success criteria. These are;

1. At least 8 different screens in the mobile application.

2. Mobile Application size will be less than 100 MB.

3. At least %60 sentiment analysis accuracy.

4. At most 3.5 seconds (3500 ms.) backend(database) response time.

Next, we will see that how I have accomplished these criteria one by one.

9.1. Criterion 1

- At least 8 different screens in the mobile application.

I have accomplished this criterion successfully. In the mobile application, we have **14 different** screens. These are start, login, sign up, forgot password, dormitory list, housemate list, user information, chat, chat list, user profile, error reports list, dormitory information, error reporting, and dormitory comments. For more information check chapter 8.1.
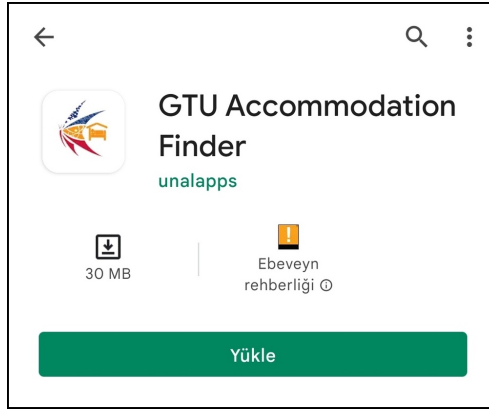
9.2. Criterion 2

- Mobile Application size will be less than 100 MB.

I have accomplished this criterion successfully. In order to show this, I have uploaded my mobile application on the google play store.[1] In the google play store, **my application download size is 29,93 MB**.

Check the application size in the following images.

---

[1]https://play.google.com/store/apps/details?id=com.cse495.GTUAccommodationFinder

(a) Mobile Application on Google Play

(b) Mobile Application Details on Google Play

Figure 9.1: Details about Application Size in Google Play Store

In following graph, which is taken from google play console[1] we can see my total application size. according to Google play reports my total application size is **33.7 MB.**
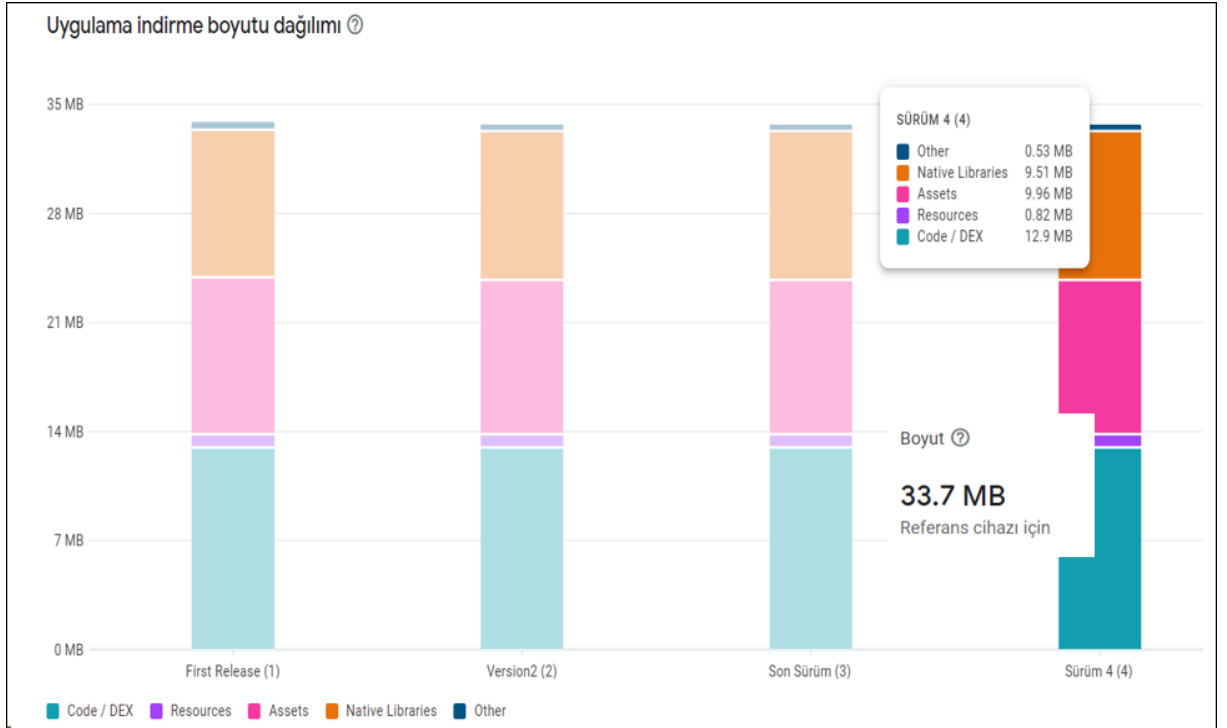

Figure 9.2: Mobile Application Size Graph on Google Play Console

---

[1]https://play.google.com/console/about/

9.3. Criterion 3

• At least %60 sentiment analysis accuracy.

    I have accomplished this criterion successfully. In order to show this, I tested my model with 2 different test method. Below we will see this methods one by one.

9.4. Summary of Success Criteria Results

    In the following table, we can see the all success criteria expected and actual results. Note that in the sentiment analysis success criterion we got the expected accuracy but **it is open to improvement**.

Success Criterion	Expected	Actual	Result
At least 8 different screens in the mobile application	At least 8	14	Successful
Mobile Application size will be less than 100 MB	At most 100 MB	33.7 MB	Successful
At least %60 sentiment analysis accuracy	At least %60	%76	Successful
At most 3.5 seconds(3500 ms.) backend response time.	At most 3.5s	1.57s, 1.20s, 2.38s	Successful

Table 9.1: Success Summary of All Results

10. Conclusions

In this project, I have tried to solve and make it easy to find a place to stay while studying in university for GTU Students. As we have seen this problem is mostly a software engineering problem, therefore, I had to take action according to this and I did mostly.

In order to solve this problem I have applied the following steps;

1. Define Requirements

2. Make a design(both visual and architectural)

3. Divide design into modules

4. Code modules one by one

5. Test each module

6. Combine modules

7. Deploy application

8. Maintenance application

As an engineer in the design step, I have considered both visual design and architectural design. My visual design is user-friendly and meets the requirements as expected. On the other hand in the architectural design step, I have chosen the most appropriate technologies for the project and heavily used ones. For example, in order to integrate the sentiment analysis module with this project, I have used a request-response mechanism with HTTP protocol by using python flask technology.

By dividing the big project into modules I have conquered each small piece so that my problems were small to solve and test. After finishing the modules I have created this project.

As a result, while making such a project the key point is being agile about both changes and learning technologies since time is limited and you can't avoid changes.

BIBLIOGRAPHY

[1] https : / / towardsdatascience . com / quick - guide - to - graph - traversal-analysis-1d510a5d05b5.

[2] A. Hölscher, *A cycle-trade heuristic for the weighted k-chinese postman problem*, 2018.

[3] D. Ahr and G. Reinelt, *New heuristics and lower bounds for the min-max k-chinese postman problem*, 2002.

[4] K. Holmberg, *Heuristics for the weighted k-rural postman problem with applications to urban snow removal*, 2015.

[5] A. Y. Gregory Gutin Gabriele Muciaccia, *Parameterized complexity of k-chinese postman problem*, 2014.

[6] D. Ahr, *A tabu search algorithm for the min–max k-chinese postman problem*, 2005.

[7] https://github.com/RaihanAk/Hotel-Review-Sentiment-Analysis_ MachineLearning.

[8] https://github.com/BaharYilmaz/turkce-duygu-analizi.

[9] https://dergipark.org.tr/tr/download/article-file/852974.

[10] https : / / towardsdatascience . com / a - beginners - guide - to - sentiment-analysis-in-python-95e354ea84f6.

[11] https://medium.com/analytics-vidhya/sentiment-analysis-on-amazon-reviews-using-tf-idf-approach-c5ab4c36e7a1.

[12] https : / / machinelearningmastery . com / save - load - machine - learning-models-python-scikit-learn/.

[13] https://reactnativepaper.com/.

[14] https://reactnativeelements.com/.

[15] https://fontawesome.com/v5.15/icons?d=gallery&p=2.

[16] https://github.com/bezkoder/react-firebase-database-crud.

[17] https://www.crowdbotics.com/blog/add-real-time-database-to-react-native-app-with-firestore.

[18] https://www.npmjs.com/package/react-native-image-slider-box.

[19] https://dev.to/chinmaymhatre/how-to-generate-apk-using-react-native-expo-kae.

[20] https://www.veribilimiokulu.com/hata-matrisini-confusion-matrix-yorumlama/.