

Technical University of Braunschweig
Institute of
Business Information Systems

Exposé for a bachelor thesis:

**Topic: Implementation of Traveling Salesman Problem for taxi route
optimization**

With a use of CRISP-DM Model to process data for TSP

Presented by:

Aymen Ben Aicha
Emsstraße 2B
38120 Braunschweig
E-Mail: a.ben-aicha@tu-braunschweig.de

Matriculation number: 5031286
Bachelor in Business Information Systems
5. Semester
Supervisor: Mr. Felix Spühler

Date: 19.03.2022

Table of contents

Problematic.....	2
Objectives	3
Theoretical foundations	4
Provisional outline	5
Timeline	6
Bibliography	7

Problematic

The Taxi & Limousine Commission is a taxi firm based in New York active in the passenger's transport sector. The firm is looking to optimize its pick-up time and location to achieve higher profits and occupancy rates. To work on this problem, different datasets are available, containing needed informations and locations. The suggested process will be to work on the given dataset, retaining consistent data by applying data processing methods and fulfilling all the CRISP-DM steps.

After validating the dataset, the challenge will be to implement an algorithm that optimizes the taxi route and gives the best awaited output, by finding the best path while setting convenient timestamps and picking the right departure and arrival locations. This optimizing algorithm is known as the problem of routing optimization and here particularly is the Traveling Salesman Problem more suitable.

Objectives

The main goal of this Bachelor thesis is to determine which road should be selected at which time while taking into considerations which factors. Implemented will be the Traveling- Salesman- Problem to achieve optimal results, while working on data pre-processing processes.

The following points are the expected goals of the work:

- Setting a clear business objective
- Preparing validated and relatable dataset to work on
- Implement TSP- Algorithm on the defined context
- Discuss the solution and its limitations
- Documentation

Theoretical foundations

Given a set of locations, trip start and end time and fare amount for each possible trip, a mathematical model should be taken into consideration to get an optimized tour planning.

The Traveling Salesman Problem (TSP) is a mathematical way to describe the desire to find the optimal route visiting all of the given places. The used solution could be classified as symmetric(sTSP), asymmetric(aTSP) and multi(mTSP) Traveling Salesman Problem.

After studying the given data set, the right category of TSP should be considered.

In this thesis we will mainly discuss many theoretical aspects of the TSP and look for the best way to find the optimal route and implement it, depending on the available data.

Provisional outline

1. Introduction
 - 1.1 Motivation
 - 1.2 Objectives
 - 1.3 Structure of the work
 - 1.4 Methodical Approach
2. CRISP-DM implementation
 - 2.1 Model definition
 - 2.2 Business objectives
 - 2.3 Data understanding
 - 2.4 Data Preparation
 - 2.5 Modeling
 - 2.6 Evaluation
3. Solution Approach
 - 3.1 Motivation
 - 3.2 Traveling Salesman Problem
 - 3.3 Approximation Procedure
4. Case Study
5. Conclusion
 - 5.1 Conclusion
 - 5.2 Limitation
 - 5.3 Final conclusion and outlook
6. Bibliography
- Appendix

Timeline

Duration: 12 Weeks (15.04.2022–15.07.2022)

until 21.04.: Business and Data Understanding

until 12.05.: Data Preparation, modeling and evaluation

until 15.05: Register the thesis (prüfungsamt)

until 26.05.: Deployment and Mathematical Formulation of the problem

until 02.06.: Optimizing Algorithm implementation

until 09.06.: Algorithm Feedback + Refinement

until 30.06.: Writing Thesis

until 07.07.: Thesis Feedback + Refinement

until 12.07.: Printing

until 14.07.: Submission

Bibliography

Jorge Nunes, Luís Matos, and António Trigo, Taxi Pick-Ups Route Optimization

Using Genetic Algorithms. <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.866.1780&rep=rep1&type=pdf>

Florentin D. Hildebrandt Marlin W. Ulmer (2021). Supervised Learning for Arrival

Time Estimations in Restaurant Meal Delivery. https://www.researchgate.net/publication/343988622_Supervised_Learning_for_Arrival_Time_Estimations_in_Restaurant_Meal_Delivery

What is CRISP DM? <https://www.datascience-pm.com/crisp-dm-2/>