FH CAMPUS WIEN
UNIVERSITY OF APPLIED SCIENCES

Data Science and Visualisation Techniques applied on Bus Search Requests and its Correlating Booking Data

Subtitle

Bachelor Thesis

Submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science in Engineering

to the University of Applied Sciences FH Campus Wien
Bachelor Degree Program: Computer Science and Digital Communications

Author:

first name surname

Student identification number:

Number

Supervisor:

Title first name surname

Date:

dd.mm.yyyy

Declaration of authorship:

I declare that this Bachelor Thesis has been written by myself. I have not used any other than the listed sources, nor have I received any unauthorized help.

I hereby certify that I have not submitted this Bachelor Thesis in any form (to a reviewer for assessment) either in Austria or abroad.

Furthermore, I assure that the (printed and electronic) copies I have submitted are identical.

Date: Signature:

Abstract

(E.g. “This thesis investigates...”)

Kurzfassung

(Z.B. “Diese Arbeit untersucht...”)

List of Abbreviations

ARP	Address Resolution Protocol
GPRS	General Packet Radio Service
GSM	Global System for Mobile communication
WLAN	Wireless Local Area Network

Key Terms

GSM

Mobilfunk

Zugriffsverfahren

Contents

1 Introduction

2 Info

2.0.1 Research Question

How to apply Data Science and Visualization Techniques on Bus Search Requests and its Correlating Booking Data to create an Analytical Dashboard.

2.0.2 General

Basic Idea/Content:

- Explain available dataset, data structure, how data is gathered
- Explain what techniques are used to clean the base dataset
- Interdisciplinary - explain why certain KPIs or models are choosen and are applied onto the dataset
- Prediction Model - explain the technologies, methods etc. used to create a ML based prediction model (To improve Yield Management). Maybe two models, Supervised Learning and unsupervised learning
- Data Clustering and other KPIs + applied statistical models (e.g. Clustering, LR), Heatmaps etc.
- Visualisation techniques used to display results of the applied statistical models
- Maybe? Short chapter about technical setup of the Dashboard

In general i would appreciate some general feedback what else could/should be descriped within this thesis. I think the Prediction Model will be the main aspect of this thesis (How it is created + which techniques are used, how is the performance when comparing prediction to actual booking data etc.)

2.0.3 Lit Research

[1] - provides also a lot of useful references to other papers that can be used
[2] - ML
[3], [4], [5], [6] - Tensorflow, ML etc.
[7], [8] - interdisciplinary to provide context which KPI's etc are chosen etc.

3 Available Dataset

general Infos

3.1 Data Origin

Where does the data come from (how it is gathered etc, important to explain the necessary data cleansing), how is it stored, for which purpose is it used.

3.2 Data Structure

Avaiable attributes within the dataset, which attributes can be used for statistical purposes etc.

3.3 Data Cleansing

Which parameters are applied to the Dataset in order to clean it

3.4 Possible Analytical Strageties that can be applied

Explain what can be applied to the dataset, what procsses could be improved by analysing the data.

Improve Yield Managment (Prediction Model, ML)

4 Prediction Model

General Infos and solutions. Explain what the model should do.

4.1 Technical Setup

What technologies are used etc.

4.2 The Model

how the model is created, which attributes are used etc.

4.3 Supervised Learning - Training Data

Explain supervised learning approach for the model

4.4 Unsupervised Learning

explain unsupervised learning apporach for the model

4.5 Comparison - Supervised, Unsupervised

4.6 Model accuracy

Having a look at the model performance accuracy (comparing predictions of the model with already available data) , explain potential twerks that have been applied to the model itself to achieve a higher level of accuracy.

5 Analytical Dashboard

5.1 Technical Setup

Explain the basic setup and used technologies for used for the analytical web based dashbaord

5.2 Applied Statistical Models

explain which attributes also provide additional information that can be gathered from the dataset, which models were applied (algorithms)

5.3 Visualisation techniques

which plots etc (and why) are used to display the gathered information

Bibliography

[1] S. V. Mahadevkar, B. Khemani, S. Patil, K. Kotecha, D. R. Vora, A. Abraham, and L. A. Gabralla, "A review on machine learning styles in computer vision—techniques and future directions," *IEEE Access*, vol. 10, pp. 107 293–107 329, 2022. 2

[2] K. Pahwa and N. Agarwal, "Stock market analysis using supervised machine learning," in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 2019, pp. 197–200. 2

[3] A. GÃ©ron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow : concepts, tools, and techniques to build intelligent systems*, third edition. ed., 2022. 2

[4] A. Géron, "Praxiseinstieg machine learning mit scikit-learn und tensorflow : Konzepte, tools und techniken für intelligente systeme," 2018. 2

[5] P. R. Gulhane and T. S. Pradeep Kumar, "Tensorflow based website click through rate (ctr) prediction using heat maps," in *2018 International Conference on Recent Trends in Advance Computing (ICRTAC)*, 2018, pp. 97–102. 2

[6] A. Dehghan-Banadaki, T. Taufik, and A. Feliachi, "Big data analytics in a day-ahead electricity price forecasting using tensorflow in restructured power systems," in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2018, pp. 1065–1069. 2

[7] A. Wannes and S. A. Ghannouchi, "Kpi-based approach for business process improvement," *Procedia Computer Science*, vol. 164, pp. 265–270, 2019, cENTERIS 2019 - International Conference on ENTERprise Information Systems / ProjMAN 2019 - International Conference on Project MANagement / HCist 2019 - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050919322215 2

[8] P. R. Martins de Andrade and S. Sadaoui, "Improving business decision making based on kpi management system," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 1280–1285. 2

List of Figures

List of Tables

Appendix

(Hier können Schaltpläne, Programme usw. eingefügt werden.)