



Data Science

Summer Semester 2023

Mini-Projects Kick-Off

Prof. Dr.-Ing. habil oec. Wolfgang Maaß
Chair of Information Systems and
Computer Science
Scientific Director at DFKI



Organization

- **Start:** 26.05.2023 – **End:** 07.07.2023
- **Submission: (use templates)**
 - Report: pdf format (5-7 pages)
 - Video presentation: mp4 format (5-7 mins)
 - Declaration of Consent
- **Grading Scheme:**
 - Report + Code: 50%
 - Video: 50%
- **Examination:**
 - Mini-projects (40%) + Written Exam (60%)
- **Note:**
 - Projects that do not meet the formal requirements will be graded as 5.0.
 - Participants who drop from the mini-project after the start will be graded 5.0 for the mini-project. This grade for the mini-project will remain for future participation in this module.

Registration Deadline:
12.05.2023



Kick-off Meeting:
26.05.2023



Submission Deadline:
07.07.2023

Exam Track "1.0"

- Students can achieve grade 1.0 in their written examination (60%) if they fulfil the following criteria:
 - **Personal letter:** To be directly sent to Prof. Wolfgang Maaß (wolfgang.maass@dfki.de) with Subject line: **"DS 23: Personal Letter"**
 - <https://cultureofbelief.com/2018/11/05/giving-yourself-an-a/>
 - **WITS Paper:** Submit a research in progress paper based on their mini-project work according to the guidelines of WITS 2023 (<https://witsconf.org/wits2023-call-for-papers/>)
- **Grading Scheme:**
 - Written Exam: 1.0
- **Registration:**
 - <https://forms.gle/H7CnxUWwMjtZ3bL5A> **(until 30.05.2023)**
- **Deadlines:**
 - Personal letter → **30.05.2023**
 - WITS Paper (max 10 pages, WITS 2023 template) → **07.07.2023**

Groups

Group Nr.	Topic Nr.	Participants
1	1	Madhurima Ghosh, Paranjay Gupta, Senjuti Dutta, Aqsa Nazir, Larisa Ivanova
2	1	Fahim Ahmed Shakil, Julian Klein, Sayed Muddashir Hossain, Pai LIU, Clara Wan Ching Ho
3	1	Ho-Hsuan Wang, Kate Rebecca Belcher, Nicholas Jennings, Myeongju Lee, William LaCroix
4	2	Harun Bajric, Mohammad Khaled Moselmany, Jose Javier Gonzalez, Nico Steinacker-Olsztyn, Riffat Sharmin
5	2	Noor Ahmed, Amna Kayani, Srushti Pravin Bhavsar, Salih Talha Akgün, Bahadır Erkan
6	2	Hewan Shrestha, Theophilus Aidoo, Angela John, Hamayoon Behmanush, Irem Begum Gunduz
7	3	Akansh Maurya, Bantam Khan Baloch, User Butt, Saira Sohail Anwari, Cicy Kuriakose Agnes
8	3	Sophia Wiedmann, Aarushi Singhal, Philipp Jonas Hawlitschek, Vitalii Hirak, Mehrad Zamani
9	3	Jiaqi Fan, Xingyue Hu, Di Wu, Chongyi Liao
10	4	Moritz Heckmann, Jonathan Müller, Robin Wiesen, Lukas Klein

Groups

Group Nr.	Topic Nr.	Participants
11	4	Soumya Shaw, Sneha Chetani, Vaishnavi Mall, Yasin Esfandiari, Shahrzad Kharazmi
12	4	Simon Benenson, Daniel Becker, Johanna girndt, Jack Weyen
13	5	Zixuan Liu, ChihYing Huang, Hüseyin Aleçakır, Ruveyda Betül Bahçeci
14	5	Sai Vignesh Kannan, Kshema Maria George, Prashanth Pombala, Marcel Schubert
15	5	Yukun Jiang, Yage Zhang, Helena Rodrigues Menezes de O. Vaz, Yu-Ci Chen
16	6	Giuliano Rasper, Hussein Sheaib, Patrick Reck, Lukas Himbert, Hilmar Müden
17	6	Nicola Josef Müller, Robert Leist, Paul Eichler, Tobias Recktenwald, Abdulhameed Nazari
18	6	Janik Keller, Cornelius Denninger, Yannick Schording, Margarete Steudter, Zain Khalil
19	6	Marc Altmeyer, Luisa Geiger, Andreas Rouvalis, Yusser Al Ghussin, Gregory Shook

Topic 1: Investment Decision

Goal: Long-term Investment Analysis using LLMs

• Task:

- Imagine you are an investor interested to make informed long-term investment decisions.
- Develop a system that assists investors in making long-term investment decisions.
- The system can analyze annual report data, macroeconomic indicators, and industry trends to identify stocks with strong growth potential.
- The system should provide recommendations for buying and holding them over an extended period.
- **Approach at least one company** to get a feedback.
- Collect or gather the suitable data for this task.
- Perform exploratory analysis by applying data science pipeline.
- Develop a model based on large language models.
- Deploy the final prototype of your model for the end-user.

• Decision:

- Recommendation whether to buy or hold the stock for an extended period.



Topic 2: Purchase Decision

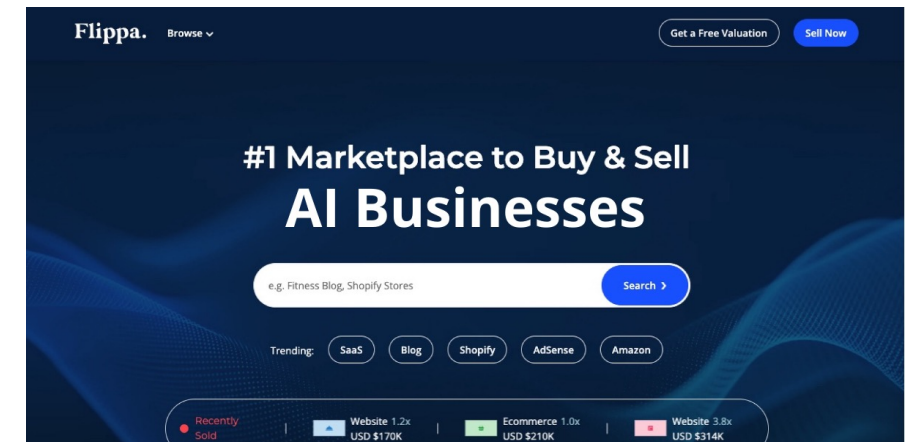
Goal: Development of Faithful System for Software Purchase

- **Task:**

- Imagine you are product manager in SME and interested in buying a software solution.
- Develop a system that calculates the total cost of ownership for different software options.
- The system should analyze licensing fees, implementation costs, maintenance expenses, and potential cost savings to assist organizations in making cost-effective software purchasing decisions.
- The system can also evaluate the user experience of software by analyzing user feedback, reviews, and usability testing data.
- **Approach at least one company** to get a feedback.
- Collect or gather the suitable data for this task.
- Perform exploratory analysis by applying data science pipeline.
- Develop a model based on language models.
- Deploy the final prototype of your model for the end-user.

- **Decision:**

- Recommendation whether to buy the software for the organization.



Topic 3: Sustainability

Goal: Optimizing EV Charging Infrastructure

- **Tasks:**

- Imagine you are responsible for planning the EV charging infrastructure.
- Develop a system that optimizes the placement and utilization of electric vehicle (EV) charging stations.
- The system can analyze historical charging patterns, traffic data, and user behaviour to identify optimal locations for charging stations, reduce congestion, and ensure efficient use of renewable energy sources.
- Collect or gather the suitable data for this task.
- Perform exploratory analysis by applying data science pipeline.
- Develop a model based on machine learning.
- Deploy the final prototype of your model for the end-user.

- **Decision:**

- Recommendation on how to reduce congestion to promote sustainability.



Topic 4: Tracking Sports

Goal: Platform for Tracking Sports Performance Data

- **Tasks:**

- Imagine you are an athletic trainer in the sports organization.
- Develop a system that analyzes sports performance data of non-elite athletes (e.g., player statistics, training data, match results) to identify strengths, weaknesses, and areas for improvement.
- The system can provide personalized recommendations and training plans to optimize player performance and team strategies.
- Collect and integrate tracking and performance data from various sources.
- Identify strengths, weaknesses, and areas for improvement based on the analysis.
- Design a personalized recommendation system for individual players and team strategies.
- Deploy the final prototype of your model for the end-user.

- **Decision:**

- Recommendations on training strategies to improve player performance.



Topic 5: Logistics

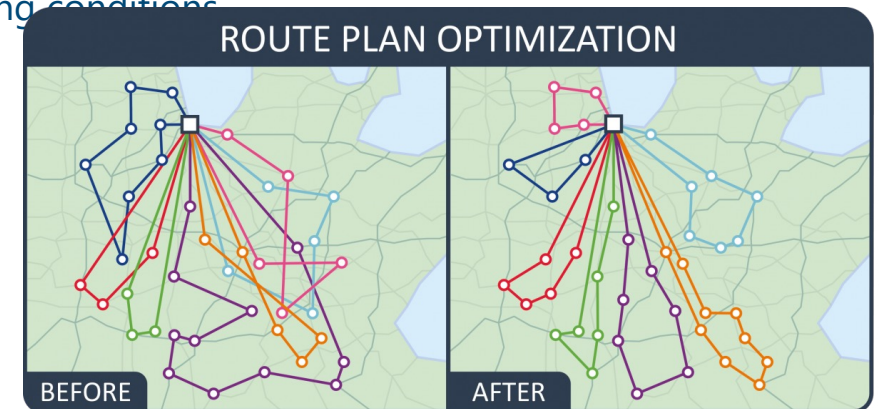
Goal: Transportation Route Optimization using Graph-based model

- **Tasks:**

- Imagine you are logistic manager in company.
- Develop a system that optimizes transportation routes considering factors such as traffic, weather conditions, delivery priorities, and vehicle capacities.
- The system can dynamically adjust routes in real-time to minimize costs, reduce delivery time, and improve overall logistics efficiency.
- Collect and integrate data on traffic conditions, weather forecasts, delivery priorities, and vehicle capacities.
- Develop a model based on graph networks that can analyze and process the data to generate optimal transportation routes.
- Implement real-time data integration to dynamically adjust routes based on changing conditions.
- Deploy the final prototype of your model for the end-user.

- **Decision:**

- Recommendation on optimal transportation route to improve logistics efficiency.



Topic 6: Market Segmentation

Goal: Clustering Analysis on AI Consultancy Services in Saarland

- **Tasks:**

- Imagine you are CEO of AI startup.
- Develop a system that clusters companies based on their industry, products, customer base, and market positioning.
- This can help businesses identify market segments, target their marketing efforts effectively, and make informed business decisions.
- Collect and integrate data on companies' competitors, customer base, and market positioning.
- Validate the system using real-world data and compare the segmentation results against established market segments.
- Implement visualization tools to present the clustered market segments in a user-friendly and interactive manner.
- Deploy the final prototype of your model for the end-user.

- **Decision:**

- Recommendation on the optimal location for setting up the consultancy.



Prof. Dr.-Ing. habil oec. Wolfgang Maaß

Professor of Information Systems

Faculty of Empirical Human Sciences and Economics

Professor of Computer Science (co-opted)

Faculty of Mathematics and Computer Science

Scientific Director

Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI)



bit.ly/wmaassLI



wolfgang.maass@dfki.de



bit.ly/issenwm



https://www.instagram.com/maass_unisb/