German Research Center for Artificial Intelligence

Project: Data Science Use Case (DLMDSPDSUC01)

Nghia Duong-Trung[1]
[1] German Research Center for Artificial Intelligence (DFKI GmbH)
Internal Use @ IU International University of Applied Science, Berlin Campus

SPONSORED BY THE

Federal Ministry
of Education
and Research

disclaimer

German Research Center for Artificial Intelligence

I used the DFKI slide template for a rapid production of my talk. It does not imply that DFKI or Federal Ministry of Education and Research has funded or related to the content. The slide is for reference only.

Who I Am

German Research Center for Artificial Intelligence

- Name: Nghia Duong-Trung
- Senior Researcher @ The German Research Center for Artificial Intelligence
  - BMBF Projects: `https://milki-psy.de/`, `https://tech4comp.de/`
- Freelance Lecturer @ IU
- PostDoc @ Technische Universität Berlin, Germany
  - BMBF Project: `https://kiwi-biolab.de/`
- PhD @ University of Hildesheim, Germany
- Profile: `https://sites.google.com/ismll.de/duongtrungnghia/`

Study Schedule

German Research Center for Artificial Intelligence

	Date	Time	Title
	16.01.2023	18:00 - 20:15	Project: Data Science Use Case - MSE_BER_DLMDSPDSUC01_2022_WS_Q1_MADS-120
	23.01.2023	18:00 - 20:15	Project: Data Science Use Case - MSE_BER_DLMDSPDSUC01_2022_WS_Q1_MADS-120
	30.01.2023	18:00 - 20:15	Project: Data Science Use Case - MSE_BER_DLMDSPDSUC01_2022_WS_Q1_MADS-120
	13.02.2023	18:00 - 20:15	Project: Data Science Use Case - MSE_BER_DLMDSPDSUC01_2022_WS_Q1_MADS-120
	27.02.2023	18:00 - 20:15	Project: Data Science Use Case - MSE_BER_DLMDSPDSUC01_2022_WS_Q1_MADS-120
	06.03.2023	18:00 - 20:15	Project: Data Science Use Case - MSE_BER_DLMDSPDSUC01_2022_WS_Q1_MADS-120

IU: New from Q12023

German Research Center for Artificial Intelligence

- Check attendance
  - Attendance or partial attendance
  - Excuse note (yes | no)
  - Absence reason (yes | no)
- Regularly take a screenshot in Zoom

Portfolio: definition and objective

German Research Center for Artificial Intelligence

- A practical exercise: written report + practical implementation + product development
  - Computer vision, natural language processing, time series
- An *achievement* you show in your CV
- The portfolio consists of three phases: conception, development/reflection, finalization
- The portfolio is submitted via the 'PebblePad' tool. See a separate user guide on myCampus
- Individual work

Formal guidelines

German Research Center for Artificial Intelligence

- Please refer to the citation guidelines during lectures

- The portfolio has to be submitted via email.
  - Files must always be named according to the following pattern: Name-FirstName_MatrNo_Course_P(hase)-2_S(ubmission). Example: Mustermann-Max_12345678_ProjectDataScienceUseCase_P2_S

Portfolio: conception phase

German Research Center for Artificial Intelligence

- The concept or core idea should be introduced contextually with a short description of the content and the goal(s) as well as the initial motivation
- Sketch your idea using Data Science Workflow Canvas.docx (see `https://github.com/duongtrung/IU-DataScienceCourse`)
- Output: compile all in one PDF file
  1. A brief explanation of the concept in the form of a text, maximum one A4 page + Data Science Workflow Canvas
     - Abstract, reference resources, key citations
  2. Data Science Workflow Canvas (1 page)
- Deadline: until **23:59 PM, 24.01.2023**. Important: This part must be submitted before the development phase.
  - You still can edit it until the final submission of phase 3

Portfolio: development phase

German Research Center for Artificial Intelligence

- Start with Machine Learning Canvas (see `https://github.com/duongtrung/IU-DataScienceCourse`)
- Important factors:
  - Realization of the project/task
  - What are your improvements?
  - Orientation to the goals: how can you judge your project is successful?
  - Consideration of available resources: time, choice of technology, computing machine
  - Project timeline
  - Public product and reproducibility

Portfolio: development phase (cont.)

German Research Center for Artificial Intelligence

- Output 1: compile all in one PDF file
  1. A brief explanation of the concept in the form of a written, ranging 8-10 A4 pages + Machine Learning Canvas
     - Introduction, related work, prerequisite background, proposed solution and experiment, discussion and outlook, conclusion, citations
     - Images and tables are welcome
  2. Machine Learning Canvas (1 page)
- Output 2: a zip file with implementation resources. External link to download the data if it is large.
- Deadline: until **23:59 PM, 06.03.2023**. Important: This part must be submitted before the final phase.
  - You still can edit it until the final submission of phase 3

Portfolio: finalization

German Research Center for Artificial Intelligence

- The final product and/or the final version of the portfolio
  - Have you successfully finished all three phases?
  - Implementation, results
  - If necessary, reflection on your performance
- Output: all in one PDF file
  - A short presentation (powerpoint, latex), maximum 5 slides
- Deadline: until **23:59 PM, 12.03.2023**

Evaluation criteria

German Research Center for Artificial Intelligence

Evaluation criteria	Explanation	Weighting
Problem Solving Techniques	*Capturing the problem *Clear problem definition/objective *Understandable concept	10%
Methodology/Ideas/Procedure	*Appropriate transfer of theories/models *Clear information about the chosen Methodology/Idea/Procedure	20%
Quality of implementation	*Quality of implementation and documentation	40%
Creativity/Correctness	*Creativity of the solution approach *Solution implemented fulfils intended objective	20%
Formal requirements	* Compliance with formal requirements	10%

Format of abstract

German Research Center for Artificial Intelligence

Length	2 pages of text
Paper size	DIN A4
Margins	Top and bottom 2cm; left 2cm; right 2cm
Font	General Text - Arial 11 pt.; Headings - 12 pt., Justify
Line Spacing	1,5
Sentences	Justified; hyphenation
Footnotes	Arial 10 pt., Justify
Paragraphs	According to mental structure - 6 pt. after line break
Affidavit	The affidavit shall be made in electronic form via “myCampus”. No submission of the examination performance is possible before it.

Format of development/reflection follows the same configuration.

Final remark

German Research Center for Artificial Intelligence

- Data Science Use Case vs Machine Learning Use Case
  - How Data Processing Pipelines, Feature Engineering affect the machine learning's outcomes
- If you have any questions regarding the submission of the portfolio, please contact the exam office via mail

Questions?

German Research Center for Artificial Intelligence