Vak: Praktijkopdracht thema 9 — credits: 6

Vakcode	BFVH3TH9
Naam	Praktijkopdracht thema 9
Studiejaar	2023-2024
ECTS credits	6
Taal	Nederlands
Coördinator	M.A. Noback
Werkvormen	Patient Behandeling
Toetsen	Th. 9 - Praktijk - Overige toetsing

Leeruitkomsten

After successfully completing the course, students will be able to

- Perform an exploratory data analysis (EDA) on a dataset that serves as input data for machine learning;
  - Identify outliers and treat these correctly
  - Identify correlated attributes
  - Apply appropriate data transformations
  - Create visualizations that help understand the data and assess its quality
- Carry out Machine Learning research:
  - Intelligently select, apply, optimize and evaluate ML algorithms;
  - Carry out a quality control on the aspects of code and algorithm performance using appropriate measures of quality;
  - Design and perform benchmarks to compare ML applications;
- Publishing a ML algorithm to the scientific community by wrapping it into a Command-line Java application and publishing it on a well-documented git repository;
- Communicate research results by means of work discussions, presentations, and a scientific report – focusing primarily on the sections Results and Discussion;
- Translate the end result of her own project into a proposal for a subsequent research project, i.e. in one of the bioinformatics minors;
- Display a well-founded critical attitude towards own and others' results as well as own professional functioning

Inhoud

In this quarter project you will get to know Data Mining (DM), also called Machine Learning (ML). These two terms are not quite the same but are often used as if they are. By the end of the course, you will know what the difference is.

You will work individually on a ML problem of your own choice, but one or more other students will analyze the same dataset. This will make troubleshooting easier and opens the possibility of peer review (analysis and code), which is also an important skill to master: providing help and critical feedback to colleagues. The research projects will either come from the public domain, or from the Hanze or one of its partners.

General organization Although the central theme is Machine Learning, several other topics and/or competences will be trained as well – in particular research skills. Therefore, the learning outcomes for the course are the following.

Assessment criteria

1. Day-to-day professional functioning (25%)
   - Critical attitude towards own work and that of others
   - Peer review (giving and receiving!)
2. Research log (Markdown) containing description of daily work - primarily dealing with the research phase (25%)
   - Reproducibility
   - Readability
3. Wrapper application of final algorithm (25%)
   - Functionality
   - Code quality
   - Documentation + Readme.md
   - Git usage
4. Report, focusing on the sections Results and Discussion and also assessing reporting technique (25%) The report should discuss
   - data exploration, data cleaning and transformations
   - Strategy and methodology of ML algorithm benchmarking and selection
   - Strategy and methodology of ML algorithm performance testing and optimization focusing on aspects of speed, memory usage, performance (accuracy, sensitivity etc)
   - Evaluation of final ML algorithm
   - Suggestions for future work

Opgenomen in opleiding(en)

Bio-Informatica

School(s)

Instituut voor Life Science & Technology

share your talent. move the world.

De ECTS onderwijscatalogus van de Hanzehogeschool Groningen wordt met de grootst mogelijke zorg samengesteld. Het is echter mogelijk dat de inhoud van de catalogus -en de daarin vervatte informatie- verouderd, incompleet of onjuist is. Aan de inhoud van de catalogus kunnen dan ook geen rechten worden ontleend.