Course Description (see also alma)

Recommendation

Prior knowledge of statistics and programming experience (e.g., R or Python) are useful but not required.

Contents

The course covers the fundamentals of machine learning and its applications in various areas of applied psychology, with an emphasis on judgment and decision-making, media and communication, and management and consumer psychology. Through selected case studies and projects, students will learn how to:

- Analyze the effects of social and informational influences on human judgment and decision-making processes,
- Identify sentiment or emotion in language and analyze the spread of misinformation in social networks, and
- Predict consumer behavior, provide personalized recommendations, or cluster brand perceptions.

Depending primarily on the interests of the admitted students, various machine learning methods (e.g., decision trees, cluster analysis, neural networks, large language models), are introduced and their application to psychological research questions is demonstrated.

Learning target

Students will develop a basic understanding of machine learning and its value to applied psychological research. They will also gain practical experience in applying specific machine learning techniques to solve real-world decision problems and enhance their ability to analyze and interpret complex data sets, such as unstructured text data. Students will also acquire the competence to critically reflect on the results of machine learning methods and to evaluate their implications for theory building and psychological research practice.

Coursework

In the FS, we will discuss the foundations of machine learning, its impact on applied psychological research, and key milestones (e.g., large language models), including ethical considerations. Coursework in the FS will consist primarily of reading relevant literature to gain a deeper understanding of the value of applying specific machine learning methods for data-driven decision-making. In the FP, students will apply this knowledge to conduct research projects in small groups. Two presentations, one at the end of the FS and one at the end of the FP, will allow students to share their ideas and insights, and receive input for a short written summary of their research project.

Literature

Jacobucci, R., Grimm, K. J., & Zhang, Z. (2023). *Machine learning for social and behavioral research*. The Guilford Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2nd ed.). Springer US. https://doi.org/10.1007/978-1-0716-1418-1

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An introduction to statistical learning: With applications in Python*. Springer International Publishing. https://doi.org/10.1007/978-3-031-38747-0

The book by Jacobucci et al. can be downloaded by students of the University of Tübingen at https://ebookcentral.proquest.com/lib/unitueb/detail.action?docID=30555833 (outside the university network you may need to use a VPN). The books by James et al. are available for free at: https://www.statlearning.com/

Preliminary Schedule

FS (roughly first half of semester)

- Introduction to topic/method
- Preparation and active discussion of case studies and practical applications
- Solving simple programming assignments in small groups
- Proposal presentation

FP (roughly second half of semester)

- Implementation of group project: Reanalysis of a published study with a set of appropriate ML techniques
    - "No Free Lunch" Theorem (Wolpert, 1996) → "We recommend the use of multiple types of machine learning algorithms, preferably with varying degrees of nonlinearity and interpretation." (Jacobucci et al., 2023)
      → Ideally, one modeling approach per group member ("Eigenständigkeit")
    - Incl. comparison to original results (either using classical statistics or different ML approach)
- Final presentation
- Short written summary about research project
    - About 5 pages/person
    - Formatting and style: APA 7

Semester Overview

Note that the following tentative schedule is subject to change based on the progress in class.

Date	Topics / Work Program	
	FS	**FP**
18.04.2024	Module 1	
25.04.2024	Module 2	
02.05.2024	Module 3	
09.05.2024	*Public Holiday*	
16.05.2024	Module 4	Group Formation
23.05.2024	*Public Holiday*	
30.05.2024	*Spring Break*	
06.06.2024		Project Planning
13.06.2024	Module 5	Project Finalization
20.06.2024	Proposal Presentations	
27.06.2024		Project Work
04.07.2024		Project Work
11.07.2024		Final Presentations
18.07.2024	Summary & Wrap-Up	
25.07.2024	*Buffer*	
31.08.2024	Submission Deadline: Written Summary	

FS Contents

Module 1: Foundations of Machine Learning (ML) for Data-Driven Decision-Making (DDDM) *(James et al., 2021, Chapters 1-4; see also Jacobucci et al., 2023, Part I)*

- Historical context and key milestones
- Importance of data-driven decision-making in research and practice (incl. goal setting, i.e., description vs. explanation vs. prediction)
- Overview of different ML techniques and their applications in psychology and related disciplines
- Fundamentals of data exploration and preprocessing (e.g., data types, feature engineering, missing data and outliers, resampling methods)
- Refresher of R (incl. linear and logistic regression)

Module 2: Supervised Learning and Model Evaluation *(James et al., 2021, Chapters 5-6 & 8-9; see also Jacobucci et al., 2023, Part II)*

- Introduction to other supervised learning algorithms for regression and classification (e.g., support vector machines), including tree-based methods and ensembles (e.g., classification trees, random forests)
- Model evaluation and interpretation (e.g., model fit, imbalanced outcomes), including feature selection (e.g., regularized regression) and model selection (tuning parameters; resampling methods)
- Selected case studies of supervised learning applications in psychology

Module 3: Unsupervised Learning and Dimensionality Reduction *(James et al., 2021, Chapter 12; see also Jacobucci et al., 2023, Chapter 10)*

- Clustering techniques for identifying patterns in psychological data (e.g., k-means, hierarchical clustering)
- Dimensionality reduction techniques (e.g., principal components analysis, multidimensional scaling), including visualization of (high-dimensional) data
- Selected case studies of unsupervised learning applications in psychology

Module 4: Deep Learning in Psychology *(James et al., 2021, Chapter 10; Jacobucci et al., 2023, Chapter 11)*

- Basics of neural networks and deep learning, including Natural Language Processing (NLP) and Generative Artificial Intelligence (GenAI)
- Selected case studies of deep learning in psychology (e.g., image recognition, text mining, social network analysis)
- Challenges and considerations in using deep learning (e.g., stimulus creation, deep fakes) and GenAI (e.g., ChatGPT as participants) in psychological research

Module 5: Ethical and Responsible AI and Future Trends in ML

- Transparency, interpretability, bias, fairness, … in AI and their impact on psychological research and automated decision-making
- Trust in technology and measurement challenges
- Emerging trends in machine learning and their potential impact (i.e., opportunities and challenges) on society and research