

Course Syllabus: Data Mining and Ethical Dialogues

Instructor Information:

- Instructor: Mohammad Naderi Dehkordi
- Office: Tome Scientific Building – Room 242
- Email: naderidm@dickinson.edu

Course Overview: This course provides an in-depth exploration of data mining techniques, combined with a critical examination of the ethical implications and dialogue practices associated with these technologies. Students will develop technical skills in data mining and engage in structured dialogues about data privacy, algorithmic bias, and the broader societal impacts of data mining. The course integrates advanced technical topics with dialogue techniques learned from the Dialogues Across Differences (DXD) workshop.

Learning Objectives:

- Master fundamental and advanced techniques in data mining, including data preprocessing, classification, clustering, and association analysis.
- Analyze and discuss the ethical implications and social responsibilities related to data mining practices.
- Develop and apply dialogue skills to facilitate interdisciplinary discussions on data ethics.
- Implement and evaluate data mining projects using ethical frameworks and dialogue strategies.

Course Materials:

- Textbook: "Introduction to Data Mining" by Pang-Ning Tan, Michael Steinbach, and Vipin Kumar.
- Supplementary Readings:
 - "Weapons of Math Destruction" by Cathy O'Neil
 - "Algorithms of Oppression" by Safiya Umoja Noble
 - Selected articles and case studies on data ethics and algorithmic bias (provided in class).
- Additional Resources:
 - DXD workshop materials on dialogue techniques
 - Technical papers and case studies on ethical data mining practices

Weekly Schedule:

Week 1-2: Introduction to Data Mining and Ethical Dialogue

- Topics:
 - Overview of data mining concepts and applications
 - Introduction to the ethical implications of data mining
 - Importance of dialogue in addressing ethical issues
- Dialogue Skills:
 - Connect before content
 - Community agreement and setting norms for dialogue
- Technical Component:
 - Introduction to data mining tools and platforms (e.g., Python, R, Weka)
- Assignments:
 - Reading: Chapter 1 of the textbook
 - Technical Exercise: Basic data exploration and visualization
 - Reflection Journal: Initial thoughts on ethical implications of data mining

Week 3-4: Data Preprocessing and Cleaning

- Topics:
 - Data quality, preprocessing techniques, and feature engineering
 - Handling missing values, noise, and outliers
- Dialogue Skills:
 - Holding the container
 - Neurophysiology of safety in technical discussions
- Technical Component:
 - Advanced preprocessing techniques (e.g., feature scaling, dimensionality reduction)
- Assignments:
 - Reading: Chapter 2 of the textbook
 - Technical Exercise: Data cleaning and preprocessing tasks using Python libraries

- Group Discussion: Ethical considerations in data handling

Week 5-6: Classification Methods

- Topics:
 - Decision trees, k-nearest neighbors, support vector machines, and ensemble methods
 - Model evaluation metrics and techniques (e.g., ROC curves, confusion matrices)
- Dialogue Skills:
 - Full-spectrum listening
 - Structured go-arounds for technical decision-making
- Technical Component:
 - Implementation of classification algorithms using Scikit-learn and TensorFlow
- Assignments:
 - Reading: Chapter 4 of the textbook
 - Technical Exercise: Implementing and evaluating classification algorithms
 - Group Discussion: Ethical implications of algorithmic decision-making

Week 7-8: Clustering Techniques

- Topics:
 - K-means, hierarchical clustering, DBSCAN, and advanced clustering algorithms
 - Applications and limitations of clustering methods
- Dialogue Skills:
 - Questions of understanding
 - Reflective structured dialogue for algorithmic interpretations
- Technical Component:
 - Advanced clustering methods and their applications
- Assignments:
 - Reading: Chapter 8 of the textbook
 - Technical Exercise: Implementing clustering algorithms with real datasets
 - Midterm Paper: Ethical considerations and applications of clustering

Week 9-10: Association Analysis

- Topics:
 - Apriori and FP-Growth algorithms, association rules mining
 - Market basket analysis and its implications
- Dialogue Skills:
 - Group Work: Differences in capacities and skills
 - Group Work: Identifying and addressing biases in data
- Technical Component:
 - Implementing association analysis using Apache Spark and other big data tools
- Assignments:
 - Reading: Chapter 6 of the textbook
 - Technical Exercise: Implementing and analyzing association rules
 - Group Project: Case study on ethical implications of association analysis

Week 11-12: Ethical Implications in Data Mining

- Topics:
 - Case studies on data privacy, algorithmic bias, and ethical decision-making
 - Discussion of "Weapons of Math Destruction" and "Algorithms of Oppression"
- Dialogue Skills:
 - Group Work: Pros and cons of ethical proposals
 - Ethical decision-making frameworks in technology
- Technical Component:
 - Ethical frameworks and tools for assessing data mining projects
- Assignments:
 - Reading: "Weapons of Math Destruction" and "Algorithms of Oppression"
 - Reflection Journal: Insights from readings and case studies
 - Technical Exercise: Applying ethical frameworks to data mining projects

Week 13-14: Dialogue in Data Mining

- Topics:

- Role-playing exercises on controversial data mining topics
 - Strategies for interdisciplinary communication and conflict resolution
- Dialogue Skills:
 - Empathetic communication
 - Techniques for facilitating complex dialogues
- Technical Component:
 - Design and implementation of dialogue-driven project evaluations
- Assignments:
 - Role-playing Exercise: Simulated dialogues on ethical data mining issues
 - Preparation for final project presentations

Week 15: Final Presentations and Reflections

- Topics:
 - Presentation of final projects and integration of technical and ethical considerations
 - Course wrap-up and reflections on dialogue and technical learning
- Dialogue Skills:
 - Reflective practice and feedback
- Assignments:
 - Final Project Presentation: Addressing a real-world ethical issue in data mining
 - Reflection Essay: Personal growth and learning in dialogue and data mining

Assessments:

- Technical Problem Sets and Exercises: 30%
- Participation in Dialogue Sessions: 20%
- Reflection Journals: 10%
- Midterm Paper: 10%
- Group Project: 15%
- Final Project and Presentation: 15%

Additional Support:

- Office Hours: Weekly office hours for additional support on technical and dialogue-related queries.
- Guest Lectures: Sessions with experts in data ethics and dialogue facilitation.
- Technical Workshops: Additional sessions on advanced data mining tools and techniques