

ECE 750 T17: Data and Knowledge Modeling and Analysis, Fall 2014

Instructor: Mohamed Kamel, EIT 4136,x 35761, mkamel@uwaterloo.ca

Course Objective:

Engineers encounter data in many of their tasks. Whether the sources of this data may be experiments, databases, computer files, or the Internet, there is a dire need for effective methods to model and analyze the data and extract useful knowledge and information from it. This course aims to provide engineering graduate students with essential knowledge of data representation, grouping, mining and knowledge discovery.

References

1. Margaret Dunham, *Data Mining Introductory and Advanced Topics*, ISBN: 0130888923, Prentice Hall, 2003.
2. Jiawei Han, Micheline Kamber & Jian Pei, *Data Mining: Concepts and Techniques*, 3rd ed, Morgan Kaufmann Publishers, May 2011.
3. R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification (2nd ed.)*, John Wiley and Sons, 2001.
4. A. K. Jain and R.C. Dubes, *Algorithms for Clustering Data*, ISBN: 0-13-022278-x, Prentice Hall, 1988.
5. P. Cohen, *Empirical Methods for Artificial Intelligence*, ISBN:0-262-03225-2, MIT Press, 1995.

Evaluation:

35% Assignments

30% Exams

35% Project: (5% Proposal, 10 % presentation, 20% report)

Major Topics:

1. Data types, sources, nature, scales and distributions
2. Data representations, transformation, dimensionality reduction and normalization
3. Classification: Statistical based, Distance based, Decision based.
4. Clustering: Partitional, Hierarchical, Model and Density based, others.
5. Retrieval and Mining: Similarity measures and matching techniques.
6. Knowledge discovery in data: Association rules mining, web mining, text mining..
7. Performance measures and tools: Statistical Analysis, Validity and Assessment Measures.