

Syllabus

IS 5310-01: Machine Learning for Business

Winter, 2020

I. Instructor

Ben Kim, Ph.D.
Professor of Information Systems

Mailing Address:

Department of Management, Albers School of Business and Economics, Seattle
University, 901 12th Ave., Seattle, WA, 98122, USA
Phone: 206-296-2806, E-mail: bkim@seattleu.edu
Office: Pigott 417
Office Hours: By appointment

II. Class

Time: 6:00 pm – 8:40 pm on Wednesday
Room: PIGT 328

III. Course Description

This course introduces data mining or knowledge discovery to provide business intelligence by analyzing massive amounts of data to find interesting patterns that can be used to guide decision making and predict future behavior. Data mining is an academic discipline developed by collective efforts from machine learning, statistics, database systems, and others. The areas where data mining can be used include marketing, finance, auditing, security, and others. Topics include Bayesian classification, decision trees, clustering, association, neural networks, and others. Students are expected to analyze real-world data in business for intelligent decision making.

We use the following software packages for practice and assignments:

- R/R Studio
- Microsoft Visual Studio
- Microsoft SQL Server

IV. Readings

- *Data Mining: Concepts and Techniques, 3rd Edition* by Han, Kamber, and Pei, Morgan Kaufmann, ISBN: 9780123814791
- <https://cran.r-project.org/manuals.html>
- Lecture Notes – Available at Canvas

V. Learning Objectives

- Understanding the fundamentals of data mining algorithms such Bayes Theorem, Association, Clustering, Decision Trees, and others
- Understanding how to use data mining algorithms to analyze data using software systems such as R/R Studio, Visual Studio, and SQL Server
- Understanding how data mining projects, including data extraction, transformation, and loading (ETL), can be applied to solve real world problems
- Understanding how data mining or big data projects assist client organizations to achieve their goals

VI. Requirements and Grading

1. Weight of Each Requirement

| | |
|-------------------------|------|
| Individual Assignments: | 20 % |
| Group Assignment: | 10 % |
| Midterm Exam: | 30 % |
| Final Exam: | 30 % |
| Class Participation: | 10 % |

2. Grading Policy

All the assignments should be submitted at the due time of the due date. Late papers may be accepted but subjected to penalty of ten percent of the possible points every calendar day. After four days after the due date, the score will be zero.

VII. University Policies

Academic Resources

- Library and Learning Commons (<http://www.seattleu.edu/learningcommons/>)
(This includes: Learning Assistance Programs, Research [Library] Services, Writing Center, Math Lab)
- Academic Integrity Tutorial (found on Canvas and SU Online)

Academic Policies on Registrar website (<https://www.seattleu.edu/redhawk-axis/academic-policies/>)

- Academic Integrity Policy
- Academic Grading Grievance Policy
- Professional Conduct Policy (only for those professional programs to which it applies)

Notice for students concerning Disabilities

If you have, or think you may have, a disability (including an ‘invisible disability’ such as a learning disability, a chronic health problem, or a mental health condition) that interferes with your performance as a student in this class, you are encouraged to arrange support services and/or accommodations through Disabilities Services staff located in **Loyola 100, (206) 296-5740**. Disability-based adjustments to course expectations can be arranged only through this process.

Office of Institutional Equity

Title IX of the Education Amendments of 1972 (Title IX) prohibits discrimination based on sex in educational programs or activities that receive Federal financial assistance. This prohibition includes sexual misconduct, which encompasses sexual harassment and sexual violence. Seattle U remains committed to providing a safe and equitable learning, living, and working environment. Seattle U offers emergency, medical, and other support resources, as well as assistance with safety and support measures, to community members who have experienced or been impacted by sexual misconduct.

Seattle U requires all faculty and staff to notify the University’s Title IX Coordinator if they become aware of any incident of sexual misconduct experienced by a student.

For more information, please visit <https://www.seattleu.edu/equity/>. If you have any questions or concerns, you may also directly contact the Title IX Coordinator in the Office of Institutional Equity (**email:** oi@seattleu.edu; **phone:** 206.296.2824) University Resources and Policies

Tentative Class Schedule

- Week 1
(1/8/2020) Introduction to Data Mining and Getting to Know Data
Text: Chapter 1 and 2
- Supervised Learning vs Unsupervised Learning
 - Types of Data (nominal, ordinal, binary, numeric, discrete, continuous)
 - SQL, DMX (Data Mining Extensions), Microsoft Visual Studio, and R
- Week 2
(1/15/2020) Online Class
R – data types, variables, data structures, reading data from database systems, and basic statistics
- Week 3
(1/22/2020) Naïve Bayesian Classification
Text: Chapter 8
- Bayes' Theorem
 - Laplacian Correction
- Week 4
(1/29/2020) Decision Tree Induction/Random Forest
Text: Chapter 8
- ID3 (C4.5), CART
 - Information Gain, Entropy, Gini Index
 - Rule-Based Classification
- Week 5
(2/5/2020) Neural Networks
Text: Chapter 9.2
- Backpropagation
 - Activation functions
 - Data Transformation using Normalization
- Week 6
(2/12/2020) Midterm Exam
- Week 7
(2/19/2020) Frequent Patterns, Association, and Correlations (Market Basket Analysis)
Text: Chapter 6
- Frequent Itemsets, Support, Confidence
 - Downward Closure Property, Apriori Algorithm
 - Confidence, Lift, and Importance of a Rule
- Week 8
(2/26/2020) Clustering and Sequence Clustering
Text: Chapter 10
- K-Means Algorithm
 - EM (Expectation Maximization) Algorithm
 - Hierarchical Methods

- Week 9 Evaluation of Models
(3/4/2020) Text: Chapter 8
- Confusion (Classification) Matrix
 - Evaluation Metrics: Accuracy, Error Rates, Sensitivity, and Specificity
 - Lift Chart
 - Prediction Errors
- Week 10 Paper Presentations and the Future of Machine Learning in Business
(3/11/2020)
- Week 11 Final Exam
(3/18/2020)

Bio of Ben Kim



Professor Ben Kim received his Ph.D., MBA, and BA from the University of Minnesota, University of Washington, and Seoul National University of Korea, respectively. Professor Kim's teaching and research areas include corporate data management, data mining, e-business, and global business management. He has published articles in *International Journal of E-Business Research*, *Journal of Systems Management*, *Data Base Management*, *Expert Systems with Applications*, *Management Decision*, *Logistics Information Management*, and many others, as well as numerous conference proceedings. Professor Kim has several chapters published in books such as *Successful Software Reengineering*, *High-Performance Web Databases*, and *Data Management Handbook*. He has conducted seminars and taught various executive programs on corporate information systems and strategies for government agencies and businesses of the United States, Korea, Japan, and Europe. He held a position of Research Fellow at J D Edwards in Denver, Colorado. At Seattle University, he is a tenured full professor and was a recipient of Genevieve Albers Professorship. Internationally, Professor Kim is a Fellow of the Japan Society of Information and Management and also holds a Visiting Professorship at Heilbronn University of Germany.