Mathematics 428 – Mathematical Foundations in Machine Learning

I. Identification

A. School of Arts and Science
B. Mathematics
C. MAT 428
D. Number has never been used
E. Mathematical Foundations in Machine Learning
F. This course give a broad introduction to machine learning by using the tools of basic knowledge of programming and probability theory, including classification; support vector machines; neural networks; clustering, feature selection, ensemble learning and reinforcement learning. The course will also discuss recent applications of machine learning, such as to computer science, data mining, bioinformatics, and so on.
G. The prerequisite is a grade of C- or better in MAT 221 and MAT372 or equivalent.There are no special conditions.
H. Probable Texts:
Machine Learning: A Probabilistic Perspective, by Kevin P. Murphy. MIT Press 2012.

The Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Trevor Hastie, Robert Tibshirani and Jerome Friedman. Springer, 2017.

Learning From Data: A Short Course, by Yaser S. Abu-Mostafa, Malik Magdon Ismail, and Hsuan-Tien Lin, AMLBook, 2012

I. Prepared by Yulei Pang
J. To be determined

II. Rationale

A. This course is being proposed as a part of undergraduate data science program. This course will cover supervised learning, unsupervised learning and reinforcement learning, which are classified as three broad categories for Machine learning tasks
B. This course will not be cross-listed
C. This course will likely be taken by students in the (anticipated) Data Science program.

III. Course Description

The course is meant to be a continuation of a statistical class, which covered a broad overview of statistical learning. Some theory will be presented, but the class will focus on applying statistical techniques to explore the construction and study of algorithms that can learn from and make predictions on data, rather than following strictly static program instructions. A statistical software package is required.

A. Student Learning Outcomes
   1. To understand what is learning and why it is essential to the design of intelligent machines.
   2. To know how to fit models to data.
   3. To understand numerical computation, statistics and optimization in the context of learning.
   4. To have a good understanding of the problems that arises when dealing with very small and very big data sets, and how to solve them.
   5. To understand the basic mathematics necessary for constructing novel machine learning solutions.
   6. To design and implement various machine learning algorithms in a wide range of real-world applications.
   7. To understand the background on deep learning and be able to implement deep learning models for language, vision, speech, decision-making, and more.

B. Course Outline
   1. Review of Basic Linear Algebra and Probability: (10%)
      Vector spaces
      Matrix theory
      Discrete random variables and common discrete distributions
      Fundamental probability rules
      Bayes rule
      Continuous random variables and common continuous distributions
   2. Classification: (20%)
      Fisher's linear discriminant
      Logistic regression classification
      Naive Bayes
      Support vector machines
      Bayesian neural networks for classification
      Decision Tree
      Classification errors, regularization, measurement
   3. Support vector machines: (12%)
      SVMs for regression
      SVMs for classification
      A probabilistic interpretation of SVMs
   4. Neural Networks: (16%)
      Feed-forward Network Functions
      Network Training
      Error Backpropagation
      The Hessian Matrix
      Regularization in Neural Networks
      Bayesian Neural Networks
      Deep learning
   5. Clustering: (16%)
      Centroid based clustering
      Hierarchy clustering
      Distribution based clustering

6. Feature selection: (10%)
   Evolutionary local selection algorithms
   Feature selection in supervised learning
   Feature selection in unsupervised learning
7. Ensemble learning: (6%)
   Stacking
   Identify mislabeled data
8. Reinforcement learning: (10%)
   Markov Decision Processes (MDP)
   Partially Observable Markov Decision Process (POMDP)

Table for MAT 428 Contact Hours

Learning Activity	Weekly Hours Spent towards course	Total Hours spent for 15 week course	Semester Credits Earned
In-Class Time Lecture / Lab	3.5	52.5	------
Text Reading	1	15	------
Mathematical Homework Assignments	2	30	------
Statistical Project Assignments Using R software	2	30	------
Total Hours	8.5	127.5	3

C. Modes of Instruction
   Mathematics 428 may follow a lecture format with homework assignments. Use of a computer package is required. A computer lab session is recommended. The course may also be offered as a hybrid course, where some sessions meet virtually.

D. Evaluation
   Exams = 50%   Projects = 30% Homework=10% Presentation=10%

Outcome	**Evaluation**
To understand what is learning and why it is essential to the design of intelligent machines.	Evaluated by homework and exams
To know how to fit models to data.	Evaluated by homework and exams
To understand numerical computation, statistics and optimization in the context of learning.	Evaluated by homework and exams

To have a good understanding of the problems those arise when dealing with very small and very big data sets, and how to solve them.	Evaluated by homework, presentation and exams
To understand the basic mathematics necessary for constructing novel machine learning solutions.	Evaluated by homework and exams
To design and implement various machine learning algorithms in a wide range of real-world applications.	Evaluated by homework and exams
To understand the background on deep learning and be able to implement deep learning models for language, vision, speech, decision making, and more.	Evaluated by homework, presentation and exams

In addition to homework, which will be graded, students will have a final project that they will complete to assess the skills. The final project will also be collected and graded.

E. Bibliography (Style used by American Mathematical Society)

Pattern recognition and machine learning, by Bishop, Christopher M. Springer, 2006.

The Elements of Statistical Learning: Data Mining, Inference, and Prediction, by Trevor Hastie, Robert Tibshirani and Jerome Friedman. Springer, 2017.

Feature selection for knowledge discovery and data mining, by Liu, Huan, and Hiroshi Motoda. Springer Science & Business Media, 1998.

Learning From Data: A Short Course, by Yaser S. Abu-Mostafa, Malik Magdon Ismail, and Hsuan-Tien Lin, AMLBook, 2012