Applied Statistics and Data Science (ASDS)

COURSES

ASDS 5301. STATISTICAL THEORY AND APPLICATIONS. 3 Hours.
An introduction to statistical theory and applications using R/SAS software. Topics may include descriptive statistics, numeracy, statistical concepts in estimation and testing, basic principles of design of experiments, analysis of variance, analysis of covariance, and some nonparametric statistical methods. Prerequisite: MATH 3330.

ASDS 5302. PRINCIPLE OF DATA SCIENCE. 3 Hours.
An introduction to the end-to-end process of going from unstructured, messy data to knowledge and actionable insights. Provides a broad overview of what data science means and systems and tools commonly used for data science and illustrates the principles of data science through several case studies, including business, marketing, medical imaging, and biology, among others. Prerequisite: MATH 3330.

ASDS 5303. STATISTICAL AND SCIENTIFIC COMPUTING I. 3 Hours.
Statistical and scientific computing focusing on programming in Python or R. Concepts in statistical computing include Monte Carlo techniques, Jackknife, Bootstrapping, Cross-Validation, and others. Concepts in scientific computing include computational and randomized linear algebra, principal component analysis, and applications. Prerequisite: MATH 3330.

ASDS 5304. APPLIED MULTIVARIATE STATISTICAL ANALYSIS. 3 Hours.
Statistical analysis for data collected in several variables. Topics include sampling from the multivariate normal distribution, multivariate analysis of variance, discriminant analysis, principal components, and factor analysis. Basic knowledge of programming in SAS or R is required. Prerequisite: ASDS 5301, ASDS 5302, ASDS 5303.

ASDS 5305. DEEP LEARNING AND ARTIFICIAL NEURAL NETWORKS. 3 Hours.
Introduction to artificial neural networks and deep learning. Topics include DL basics, history, and introduction to Pytorch; Perceptrons and their inspiration from neuroscience; Gradients, Autograd, designing deep linear systems; Training and testing neural networks, backpropagation, stochastic gradient descent; Applying Optimization and regularization to neural nets; Convolutional Neural Networks (ConvNets/CNNs); Recurrent Neural Networks (RNNs); and Modern Networks and their applications. Prerequisite: ASDS 5301, ASDS 5302, ASDS 5303.

ASDS 5306. APPLIED TIME SERIES ANALYSIS IN DATA ANALYTICS. 3 Hours.
An introduction to the theory and applications of time series modeling with an emphasis on modeling and forecasting using the software. Topics include stationarity and autocorrelation, autoregressive, moving average, ARMA and ARIMA; forecasting and estimation; spectral analysis. Computational implementation in R. Basic programming skills is preferred. Prerequisite: ASDS 5301, ASDS 5302, ASDS 5303.

ASDS 6301. ADVANCE REGRESSION ANALYSIS. 3 Hours.
Topics include multiple linear regression, ordinary and generalized least squares, partial and multiple correlation, regression diagnostics, collinearity, model building, and nonlinear regression. The course provides an extended introduction to the computer package widely used for statistical analysis. Basic knowledge of programming is required. Prerequisite: MATH 3330.

ASDS 6302. MACHINE LEARNING WITH APPLICATIONS. 3 Hours.
Topics include but are not limited to supervised learning methods: linear model, generalized linear model, logistic regression, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), nearest neighbor classifier, support vector machines, tree-based methods (decision tree, random forest, XGBoost), and neural networks; and unsupervised learning methods: clustering, principal component analysis, and independent component analysis. The course provides an extended introduction to tools widely used for statistical machine learning. Basic programming skills are preferred. Prerequisite: MATH 3330.

ASDS 6303. DATA MINING WITH INFORMATION VISUALIZATION. 3 Hours.
Introduction to statistical pattern recognition. The main topics include Bayes decision theory, discriminant functions, maximum likelihood estimation, PCA, LDA, semi-supervised kernel learning, and graph embedding. This course will discuss some applications of data mining in different application fields, such as business, marketing, medical imaging, biology. Prerequisite: MATH 3330.

ASDS 6304. OPTIMIZATION AND BIG DATA ANALYTICS. 3 Hours.
Topics include an introduction to big data analysis, real-world applications of data science, linear system solutions, linear programming, duality theory, convex sets, convex functions, optimality conditions, unconstrained Optimization, constraint optimization, conjugate direction methods, alternating direction method of multipliers, classification/regression models and algorithms, dimensionality reduction for visualization and projects on real data. Prerequisite: ASDS 5301, ASDS 5302, ASDS 5303.

ASDS 6305. STATISTICAL AND SCIENTIFIC COMPUTING II. 3 Hours.
Advanced topics in statistical and scientific computing, emphasizing complex data analysis, such as high-dimensional and functional data. Topics include Expectation-Maximization (EM), Stochastic and Monte Carlo EM, Metropolis-Hastings algorithm, Gibbs sampling, functional principal component analysis, tensors, and tensor decompositions. Prerequisite: ASDS 5301, ASDS 5302, ASDS 5303.

ASDS 6306. INTERNSHIP/CAPSTONE RESEARCH PROJECT. 3 Hours.
The capstone project aims to give students a "hands-on" experience in analyzing interesting datasets with cutting-edge techniques of their interest. While students are encouraged to get feedback and mentoring from faculty, they should work as independently as possible with their teams. Progress will be monitored by completing the following quarterly milestones: (1) Acquisition of dataset, data pre-processing, and clear statement/justification of the proposed analysis. (2) Preliminary results and discussion of the plausibility of the results, (3) Refinement of codes, analysis, and results. (4) Submission of a written 5+ page report including introduction, background, methods, results, discussion, and conclusion. (5) Final oral presentation chaired by an ASDS faculty who read the written report. The project team members are expected to answer project-related questions from faculty and peers satisfactorily. Prerequisite: ASDS 5301, ASDS 5302, ASDS 5303, ASDS 6301, ASDS 6302, ASDS 6303.