BIG DATA ANALYTICS Courses

**BDA 411/511. Introduction to Machine Learning I. 3 Credits.**
An introductory course on machine learning. Machine Learning is the science of discovering pattern and structure and making predictions in data sets. It lies at the interface of mathematics, statistics and computer science. The course gives an elementary summary of modern machine learning tools. Topics include regression, decision trees, artificial neural networks, genetic algorithms, clustering, dimension-reduction, learning sets of rules, support vector machines, hidden Markov models, and Bayesian learning. The course will also discuss applications of machine learning that include data mining, bioinformatics, speech recognition, and text and web data processing. Students enrolled are expected to have some ability to write computer programs, some knowledge of probability, statistics and linear algebra. Prerequisites: MATH 312, MATH 316, and STAT 331.

**BDA 413/513. Introduction to Machine Learning II. 3 Credits.**
Due to the rapid development of technology, vast amount of data is being generated in social media, genomics, financial industry, business and health care. We rely increasingly on mathematical and statistical tools to analyze this big data. From the vast array of tools available, in this course students will study the most relevant ones. Topics that will be covered are Logistic Regression, Lasso and Related Methods, Support Vector and Kernel Methodology, Principal Components (SVD) and Variations, Multidimensional Scaling, Boosting, Random Forests, Graphical Models, False Discovery Rates and Permutation Tests. Prerequisite: A grade of C or better in BDA 411.

**BDA 431/531. Modern Statistical Methods for Big Data Analytics. 3 Credits.**
The statistical perspective of data mining is emphasized for the majority of the course. Both applied aspects (programming, problem solving, and data analysis) and theoretical concepts (learning, understanding, and evaluating methodologies) of data mining will be covered. Topics may include data preparation, data cleaning, exploratory data analysis, statistical estimation and prediction, regression modeling, clustering, classification and regression trees. An elementary knowledge of concepts in probability and statistics and basic computing proficiency are required. Prerequisites: A grade of C or better in STAT 331 and STAT 405.

**BDA 432/532. Introduction to Optimization and Inverse Problems. 3 Credits.**
Topics considered include the solution of non-smooth optimization problems arising in data science, including unconstrained and constrained optimization problems, Lagrange multiplier methods, inequality constraints, Kuhn-Tucker conditions, and applications. Also considered are linear and nonlinear inverse problems, regularization of ill-posed problem including singular value decomposition, and Tikhonov regularization methods and sparse regularization methods, inverse eigenvalue problems and applications such as compressed sensing, image reconstruction and machine learning. Prerequisites: MATH 307, MATH 312 and MATH 316.

**BDA 450. Senior Project in Big Data Analytics I. 3 Credits.**
This course introduces students to practical applications of big data analytics. Lecture topics include an overview of the various topics in business, engineering, and government currently using big data analytics. Students will choose a project involving a real world application to explore techniques learned during other course work. Course involves written and oral presentations for students to improve communication and teamwork skills. Prerequisites: A grade of C or better in STAT 331 and STAT 405. Pre- or corequisite: BDA 431.

**BDA 451. Senior Project in Big Data Analytics II. 3 Credits.**
This course allows the student to pursue an in-depth exploration of a project initiated in BDA 450. The course involves written and oral presentations for students to improve communication and teamwork skills. Prerequisites: BDA 450 and permission of instructor.