

APPLIED DATA SCIENCE (ADS)

ADS 521: Foundations of Data Science (3 credits)

Serving as an introduction course to the ADS program, ADS521 examines the history of data science, its status as a scientific and applied discipline in a modern-day world, and surveys all the important topics covered in the courses in the program and many of their applications to everyday life. Part of the course will also serve as review of mathematics and basic programming knowledge.

ADS 524: Data Analytics (3 credits)

This course covers the foundation of statistical analysis of data. While necessary concepts and theory are covered, emphasis is put on the analysis of real-world data. Topics include exploratory data analysis—organizing, displaying and describing data, summarizing data, discrete and continuous distributions, data sampling and hypothesis testing (including confidence intervals), statistical inference, ANOVA analysis. Students are encouraged to use Excel and R for assignments and projects.

ADS 525: Machine Learning I (3 credits)

ADS525 and ADS526 serve as an introduction to the common themes in data mining and machine learning, covering a wide range of various data problems and their solutions, with focus on hands-on applications instead of theory. ADS525 covers topics including data preprocessing, learning methods such as decision trees, random forest, Naïve Bayes, k-Means, data reduction, shrinkage methods, principle component analysis and discriminant analysis, all with hands-on application utilizing statistical packages and programming languages.

ADS 526: Machine Learning II (3 credits)

Continuing from ADS525, ADS526 covers additional important data mining methods such as bagging and boosting techniques, neural networks, clustering and ensemble methods.

Prerequisite: ADS 525

ADS 534: Statistical Modeling (3 credits)

This application-focused course focuses on regression analysis including linear, multiple linear and logistic regression models, with detailed discussions of model formulation, model inference, and model interpretation. Programming languages such as a SAS will be utilized.

Prerequisite: ADS 524 or ADS 535

ADS 535: Probability and Statistical Inference (3 credits)

This course provides an introduction to probability and foundational concepts in statistics. Topics include key concepts in probability theories, random variables and probability, common distributions, expected values, variance, covariance, central limit theorems, sampling, estimation of parameters and hypothesis testing. Methods include nonparametric method, design of experiment, Bayesian analysis, and resampling methods such as Bootstrap and Jackknife. The course will use R or similar programming language.

ADS 635: Data Mining I (3 credits)

Statistical analysis and data mining has been recognized as one of the “Hottest Skills” on LinkedIn year after year. The growing complexity and size of data has given rise to unprecedented demands and challenges for the field. Moreover, the mastery of various methods, the selection and application of appropriate techniques is equally important as the effective presentation and interpretation of findings. The objective of ADS635 and ADS636 is to modernize student training to better suit these demands. Topics for ADS635 will focus on supervised learning, including feature selection, discriminant analysis, regularization methods, ensemble methods, support vector machines and model assessment, all with hands-on application utilizing statistical packages and programming languages.

ADS 636: Data Mining II (3 credits)

This course will cover additional topics in data mining with a focus on unsupervised learning. Topics include association rules, clustering methods, self-organizing maps, recommender systems, ensemble methods, dimension reduction and probabilistic graphical models. Special emphasis will be on data with high-dimensionality or massive sample size. Concepts will be reinforced with hands-on applications that utilize statistical packages and programming languages.

ADS 637: Data Exploration and Visualization (3 credits)

This course is an introduction to data visualization. It includes data preprocessing and focuses on specific tools and techniques necessary to visualize complex data. Data visualization topics covered include design principles, perception, color, statistical graphs, maps, trees and networks, data visualization tools, and other topics as appropriate. Visualization tools may include Tableau, Python, and R, etc. The course introduces the techniques necessary to successfully implement visualization projects using the programming languages studied.

ADS 638: Database Systems (3 credits)

It is increasingly important for data scientists to understand various database models and their associated data access methods. This course covers both the fundamental concepts of database systems and associated tools. Topics include conceptual data modeling, database design and normalization, database implementation and the use of SQL for data definition, manipulation, and query processing. The course also includes a survey of techniques for handling non-relational data models, massive datasets, and unstructured data, including data warehousing, in-memory databases, NewSQL, NoSQL, and Hadoop.

ADS 652: Text Mining (3 credits)

Mining high-quality information from text has become critical to many industries. Starting from basic natural language processing techniques and document representation, to text categorization and clustering, sentiment analysis and text-based prediction, this course serves as a comprehensive introduction to the topic. Relevant tool-kits will be utilized and case studies from various industries will be examined.

Prerequisite: ADS 524 or ADS 535. ADS 525 and ADS 526 OR ADS 635 and ADS 636 are recommended.

ADS 653: Generative AI: Foundations and Applications (3 credits)

This course explores the fundamentals of Generative AI, its underlying architectures, practical applications, and ethical considerations. Students will gain hands-on experience in building and fine-tuning generative models, including Large Language Models (LLMs), image generators, and creative AI systems.

ADS 654: Deep Learning (3 credits)

Many recent breakthroughs in artificial intelligence have been made possible by deep learning, a branch of machine learning concerned with the development and application of modern neural networks. This is an advanced course that builds upon the knowledge of probability, statistics, linear algebra, optimization and basic neural networks. Topics include convolutional and recurrent network structures, deep unsupervised and reinforcement learning, and applications to problem domains such as speech recognition and computer vision.

Prerequisite: ADS 524 or ADS 535, ADS534. ADS 525 and ADS 526 OR ADS 635 and ADS 636 are recommended.

ADS 670: Case Analysis Capstone (3 credits)

This is a project-oriented course at the end of the program. Students will demonstrate their competence in the theory and practice learned from the program through the whole process of a complex data analysis project, including data collection, exploration, preparation, analysis, interpretation and presentation. The project can be either relevant to students' experience or aspired field, accompanied by a final essay in which students reflect upon the goals of the program and their personal goals, demonstrate how they met these goals, and what work supports their arguments.