Graduate Certificate in Agricultural Data Science

List of suggested courses (updated 23 September 2021)

A Graduate Certificate is similar to a minor and is designed to complement a student's major program of study. This program is open only to enrolled, degree-seeking graduate students at the University of Georgia. The curriculum is highly interdisciplinary and **requires a total of 16 credit hours across four areas**. Two required core courses integrate analytical approaches and case studies from a range of disciplines. A capstone seminar features intramural and invited speakers across disciplines and industries. Elective courses are drawn from four different colleges and two Institutes. The curriculum has been designed such that prerequisites for core and elective courses will not be prohibitive for students coming from a range of different disciplines and graduate majors.

Certificate Areas:

- Area 1: Agricultural Data Science Core (6 credits): Two required courses covering the foundations in descriptive and predictive analytics in the agri-food sciences and providing context for and integration among more specialized data science elective courses.
- Area 2: Analytical Foundations (3 credits): More specialized elective courses in the foundations of data science: programming, data management, statistics, econometrics, and/or data mining.
- Area 3: Analytical Applications (at least 6 credits): Elective courses from a range of applications including precision agriculture, geographic information science, imaging and sensing, agricultural statistics, bioinformatics, and consumer analytics, among others.
- Area 4: Seminar in Agricultural Data Science (1 credit): Interdisciplinary seminar course featuring UGA and external (industry and academia) speakers highlighting diverse applications in agricultural analytics.

AREA 1: AGRICULTURAL DATA SCIENCE CORE

Both courses are required:

INFO 8000 – Foundations of Informatics for Research and Practice (3 credits)

This interdisciplinary course provides instruction and exposure to the theory, tools, and techniques that connect data to information, knowledge, and decisions. Students will gain the knowledge and skills necessary to deeply engage in the increasingly interdisciplinary, data-driven, security-focused industrial and research enterprises as they complete practical analytical tasks and projects.

AESC 6100 - Applied Agricultural Data Science (3 credits)

This course will cover a variety of modern approaches for analyzing and interpreting data types commonly encountered in the agri-food sciences (including but not limited to variable selection and transformation, decision trees, neural networks, regression models, combination of models, and text mining).

AREA 2: ANALYTICAL FOUNDATIONS (Programming, Statistics, Data Mining, Data Management)
Take 3 credits:

AAEC 6610 – Quantitative Techniques in Agricultural Economics (3 hours)

Basic quantitative techniques in agricultural economic theory, emphasizing basic models used in the study of prices, marketing, and production.

AAEC 6630 or 6630E - Quantitative Tools for Agribusiness Management (3 credits)

Quantitative methods for agribusiness management focused on seven topics, including statistical tests, regression, forecasting, linear programming, non-linear optimization, multi-criteria decision making, and simulation models. These tools are introduced in lecture and then put to practical use in the computer lab using SAS and Excel.

BINF 8006 - Programming and Data Structures for Informatics (4 credits)

An intensive introduction to fundamental concepts in programming and data structures and their application to everyday use in informatics analyses. Hands-on exercises will emphasize problem-solving and writing code to collect, analyze, and present results.

CSCI 6360 - Data Science II (4 credits)

An introduction to advanced analytics techniques in data science, including random forests, semi-supervised learning, spectral analytics, randomized algorithms, and just-in-time compilers. Distributed and out-of-core processing.

CSCI 6370 – Database Management (4 Credits)

The theory and practice of database management. Topics to be covered include efficient file access techniques, the relational data model as well as other data models, query languages, database design using entity-relationship diagrams and normalization theory, query optimization, and transaction processing.

CSCI(STAT) 6375 - Foundations of Data Science (4 credits)

Introduction to the life cycle of data starting from data collection to cleaning, management, storage, sorting, provenance, visualization, and analysis. A rigorous overview of methods for text mining, image processing, linear models, and scientific computing. Core concepts in supervised and unsupervised analytics, dimensionality reduction, and data visualization will be explored in depth.

STAT 6365 - Modern Statistical Programming (3 credits)

Statistical analysis and data manipulation in R and Python. Implementation of SQL. Topics include data input/output; data formats and types; data management; functions for statistical modeling; introduction to algorithms; flow control and program design; and programs for complex data manipulation and analysis. Additional topics may include MATLAB and parallel computing.

AREA 3: ANALYTICAL APPLICATIONS

Take at least 6 credits:

AAEC 6620 - Applied Econometrics (3 credits)

Standard and advanced econometric techniques are applied to topics in agriculture and resource economics. Techniques include models for cross-section data, such as pooled regressions, limited dependent variable models, random and fixed effects models for panel data, and forecasting and volatility models for time series data. Students will conduct statistical analyses and model evaluation.

AAEC 8610 - Advanced Econometric Applications (3 hours)

Development and use of econometric techniques. Emphasis on the application of maximum likelihood estimation using MATLAB. The analysis of categorical and survival data, multiple equation regression models, simultaneous systems, and multivariate time series. The treatment of models with discrete and limited dependent variables in a panel data context.

ADSC 8000 – Statistical Genetics and Bioinformatics with Application to Animal Agriculture (3 credits)

Currently used methods and techniques in the field of molecular genetics and bioinformatics with applications in livestock and poultry species. Major emphasis on Genome wide association studies (GWAS), and genomic selection (GS). Lab sessions include the simulation and analysis of genomic data within the context of GWAS and GS. Students will have the opportunity to develop their own code or to use software provided by the instructor.

ADSC 8120-8120L – Advanced Statistical Methods in Animal Breeding (3 credits)

Latest statistical and computing procedures used in animal breeding and genetics with special emphasis on Bayesian statistics, Markov chain Monte Carlo techniques and methods useful for analysis of molecular marker information.

BINF(MIBO)(BCMB) 8211 - Advanced Methods for Biological Data Analyses (3 credits)

Advanced strategies and methodologies for large-scale data analyses in support of genomics, transcriptomics, proteomics, and studies of biological pathways and networks. Topics include gene finding, genomic rearrangements, microarray data analyses, protein function inference, protein-protein interaction prediction, and pathway and network prediction. Major data mining tools will be covered for each topic.

BINF 8441 - Statistics for Bioinformatics (3 credits)

Introductory statistics for students in the life sciences, including probability, discrete and continuous random variables, distributions, expectations, maximum likelihood, Bayesian inference, hypothesis testing, and linear regression. These topics will be mixed with applications of the statistical concepts to biological data. Statistical inference and real data analysis are implemented in R.

BINF 8980-8980D - Case Studies in Systems Biology (4 credits)

Shared research experience in systems biology. Each semester the research case study will be either on the biological clock, hot-pathogen interactions, or marine metagenomics. Project will include genomics experiments involving microfluidics, network identification, and genomic

analysis. Emphasis will be placed on transformative research accomplished on the clock, host-pathogen system, or marine ecosystem.

CRSS 6030/6030L - Sensors in Precision Agriculture (3 credits)

An in-depth examination of sensing mechanisms for precision agricultural practices. Classroom and lab exercises will in general examine sensing technologies, commercial instruments and applications, hands-on testing, and data collection in real-world precision agricultural applications.

CRSS 6050 - Improving Nutrient and Energy Efficiency with Geographic Information Systems (4 credits)

Students will apply GIS, GPS, and remote sensing principles in agricultural applications; emphasizing hands-on experience working with producers, vendors, and researchers collecting data, developing spatial databases, analyzing data, and communicating findings. Students will explore agriculture as a landscape complex and evaluate regulatory requirements, conservation opportunities, applied technologies, and market factors.

CRSS 6060-6060L - Advanced Topics in Precision Agriculture (3 credits; Tifton Campus)

Concepts and analytical techniques used in precision agriculture to make management decisions such as geostatistics to analyze georeferenced data, development of management zones, integration of sensors with real-time control systems, and big data analytics. Lab exercises will provide experiential learning of topics covered during lectures.

CRSS(PBGG) 8010 - Research Methods and Design for Crop Science (3 credits)

Foundational methods to design and conduct effective field trials for plant research.

CSCI 6380 – Data Mining (4 credits)

A broad introduction to data mining methods and an exploration of research problems in data mining and its applications in complex real-world domains. Approaches include association and classification rule learning, tree learning, neural network and Bayesian methods, support vector machines, clustering, and ensemble learning.

CSCI 6330 - Artificial Intelligence and the Web (4 credits)

The application of artificial intelligence methodologies and algorithms to problems involving the world wide web. Introduction to problem-solving, knowledge representation, learning, and reasoning techniques and exploration of how they are applied to enable information provisioning, social networking, and service provisioning on the web.

CSCI 6850 - Biomedical Image Analysis (4 credits)

Introduction to the standard approaches to biomedical image analysis, including basic concepts of biomedical imaging, basic algorithms, principles of software systems, and their applications. Biomedical image analysis software tools will be used in hands-on projects.

FANR 6750-6750D - Experimental Methods in Forestry and Natural Resources Research (4 credits)

Statistical procedures and computer software to collect, analyze, and interpret forest resources research data.

FANR 8400 - Advanced Spatial Analysis for Natural Resources (3 credits)

Advanced theory and applications of spatial information technology and spatial analysis techniques in natural resources. Focus will be on addressing realistic problems within the field of natural resources, including in student's own research area.

FHCE 7050 - Consumer Analytics and Research Methods II (3 credits)
Advanced research methods with an emphasis on applied consumer research techniques, interpretation, and dissemination. Through the use of modern analytic tools and diverse quantitative methods, students learn to integrate design, measurement, sampling, data management, and analytic techniques found in applied consumer analytic settings such as business, government, and non-profit organizations.

FORS 6760-6760L – Quantitative Models for Forest Resources Managers (3 credits)
Model forms used to simulate tree and forest stand development as well as models used to simulate the growth of various wildlife and fish species. Parameter estimation methods and model evaluation included.

FORS 7690-7690L – Applied Geographic Information Systems (GIS) for Forest Resource Management (3 credits)
Geographic Information Systems (GIS) methods and techniques to solve management problems faced by forestry professionals. Focus on the collection, organization, and analysis of spatial and tabular information with an emphasis on big data (FIA, RPA, census, Landsat) and their use in the decision-making process.

GEOG(CRSS) 6375 - GIS Applications in Agriculture (4 credits)
An exploration of the uses of Geographic Information Science (GIS) technology in agricultural applications. Basics of Global Positioning Systems (GPS) for location reference, GIS for field investigation, and remote sensing for crop and soil evaluation will be explored. Construction of GIS databases for precision farming and watershed management applications will be required.

PATH 8310-8310L – Epidemiology of Plant Diseases (3 credits)
Factors altering the course of disease epidemics in plant populations. Techniques for qualitative and quantitative measures of such factors will be emphasized.

PBIO(BIOL)(BINF) 6550 - Bioinformatics Applications (3 credits)
State-of-the-art computational analyses of genome, DNA, RNA, and protein sequences will be presented, including programs for analyzing these data and the underlying analysis methods. Topics include sequence and structure databases; sequence assembly; sequence alignment; evolutionary analyses; gene function prediction; genome annotation; and applications for medical, agricultural and environmental genomics.

PBIO(BINF)(FANR) 6700 - Computational Plant Science (3 credits)
Introduces computational techniques to students who are new to programming or do not regularly program using examples from plant science. In doing so, the course introduces basic simulation techniques and imaging techniques that can be specialized and further developed in higher level graduate courses.

PBIO(PATH) 8250 - Experimental Design and Analysis in Field Plant Biology (3 credits)

Students will learn how to design effective, informed, experimental, and observational studies for field plant biology. The course will also cover statistical analyses and software to interpret and visualize those data.

STAT 6230 - Applied Regression Analysis (3 credits)

Applied methods in regression analysis. Topics include univariate linear regression, techniques of multiple regression and model building, ANOVA as regression analysis, analysis of covariance, model selection and diagnostic checking techniques, nonlinear regression, and logistic regression.

STAT 8200 - Design of Experiments for Research Workers (3 credits)

Methods for constructing and analyzing designed experiments are considered. Concepts of experimental unit, randomization, blocking, replication, and orthogonal contrasts are introduced. Designs include completely randomized design, randomized complete block design, Latin squares design, split-plot design, repeated measures design, and factorial and fractional factorial designs.

STAT 8250 – Multivariate Methods

An introduction to the methodology of multivariate statistics for quantitatively-oriented students from various disciplines who have training in regression and analysis of variance. Topics include the multivariate normal distribution, one and two population inference on population mean vectors, MANOVA, principal component analysis, factor analysis, discrimination, classification, and canonical correlation.

AREA 4: SEMINAR IN AGRICULTURAL DATA SCIENCE

This course is required:

AESC 8150 – Seminar in Agricultural Data Science (1 credit)

Capstone seminar course featuring UGA and external speakers (industry and academia) highlighting diverse applications in agricultural analytics.

More detailed information about all courses, including the semesters when they are offered, is available in the UGA Bulletin (http://bulletin.uga.edu/). For more information about the Certificate, please visit https://site.caes.uga.edu/agdatascience/