Katz
Katz School
of Science and Health

M.S. in Artificial Intelligence Course Descriptions

AIM 5001 Data Acquisition and Management

Data Acquisition and Management focuses on data structures, data design patterns, algorithms, methods, and best practices for the pre-modeling phases of data science workflows, including problem formulation, gather, analyze, explore, model, and communicate, analytics programming focuses on the gather, analyze, and explore workflow steps. This comprises the “data wrangling” work which is where most data scientists spend most of their time.
Because data science is iterative, this preparatory work informs the modeling phase. Often, the creation and validation of new models require going back for additional data, different data transformations, and exploration of data distributions. In short, every effective data scientist needs to master analytics programming. Course topics include reading from or writing to databases, text files, and the web; shaping data into “tidy” data frames, exploratory data analysis, data imputations, feature engineering, and feature scaling.

Industry Applications: Data acquisition skills enable reliable analytics and modeling pipelines. Companies such as Amazon and Uber rely on data engineers to transform raw transactional and behavioral data into structured datasets for machine learning. Healthcare organizations including UnitedHealth Group integrate clinical, claims, and demographic data to support analytics and compliance reporting. Consumer goods firms like Procter & Gamble apply feature engineering and exploratory analysis to demand forecasting and marketing optimization.

AIM 5002 Computational Statistics and Probability

Most of data science is statistical learning, which requires strong foundational knowledge in probability and statistics. And applying computational methods such as direct simulation, shuffling, bootstrapping, and cross-validation to statistical problems is often more intuitive, and intuitive and can provide solutions where analytical methods would prove computationally intractable. This course introduces students to the statistical analysis of data using modern computational methods and software. Probability, descriptive statistics, inferential statistics and computation methods such as simulations sample distributions, shuffling, bootstrapping, and cross-validation will be covered.

Industry Applications: Computational statistical techniques are fundamental to experimentation, forecasting and risk assessment. Technology firms such as Meta and Netflix use bootstrapping and cross-validation to evaluate product changes and personalization algorithms. Financial institutions including JPMorgan Chase and BlackRock apply simulation-based probability models to quantify market risk and portfolio uncertainty. Public health organizations such as Centers for Disease Control and Prevention rely on probabilistic modeling and simulation to evaluate disease spread and intervention strategies.

AIM 5003 Numerical Methods

Algorithms in machine learning and neural networks are built upon a strong foundation of linear algebra. For example, modern recommendation systems may have sparse matrices with millions of users and millions of items; matrix factorization methods make the underlying calculations tractable say this course builds a foundation of linear algebra concepts such as matrices, determinants, vectors and eigen values. Then it deepens it into data science applications around network analysis and logistic algorithms. In addition, some multi-variate calculus and graph theory topics are covered.

Industry Applications: Numerical methods make large-scale analytics computationally feasible. Financial data providers like Bloomberg rely on matrix operations and optimization for real-time analytics. Industrial firms such as Siemens use numerical optimization and network analysis for systems modeling. Recommendation systems across technology and retail sectors depend on matrix factorization techniques to handle massive user-item datasets.

AIM 5004 Predictive Models

Prerequisites: Data Acquisition and Management; Computational Statistics and Probability

Predictive modeling answers the question, “What will happen next?” Linear regression and logistic regression are foundational predictive modeling methods, used to predict continuous and categorical output, respectively. The main topics covered in this course include simple and multiple linear regression, variable selection and shrinkage methods, binary logistic regression, count regression, weighted least squares, robust regression, generalized least squares, multinomial logistic regression, generalized linear models, panel regression, and nonparametric regression.

Industry Applications: Predictive modeling expertise is widely used in analytics, forecasting, and risk assessment roles. Banks such as Wells Fargo apply regression models to credit risk and regulatory analysis. Healthcare systems like Kaiser Permanente use predictive models to forecast patient outcomes and resource utilization. Marketing teams across industries rely on regression for pricing strategy and campaign effectiveness analysis.

AIM 5005 Machine Learning

Prerequisites: Data Acquisition and Management; Computational Statistics and Probability

In classical programming, answers are obtained from rules and data. In machine learning, rules are obtained from data and answers. The widespread availability and sharing of data, and improvements in computing capacity, processing methods, and algorithms have given machine learning the power to deliver game-changing systems and technologies to organizations that compete on predictive, prescriptive, and/or autonomous analytics. In this course, we’ll look at methods for using, tuning, and comparing machine learning algorithms, based on measures of performance, accuracy, and explainability. We’ll also look at recent advances and trends in automated machine learning.

Industry Applications: Machine learning (ML) techniques drive predictive and prescriptive analytics across industries by leveraging statistical patterns in data to automate complex decisions. Financial institutions use ML for credit scoring, underwriting, and fraud detection; retailers and logistics firms apply forecasting and optimization for demand and inventory management; and healthcare organizations employ clustering and survival analysis for patient risk and resource planning. Emerging applications include generative and agentic AI systems such as Claude Code, ChatGPT, and Gemini, embodied AI in robots, drones, humanoids, and autonomous vehicles, and vision-language models and coding agents that integrate perception, reasoning, and adaptive decision-making.

AIM 5006 Artificial Intelligence

Prerequisites: Data Acquisition and Management; Computational Statistics and Probability

Artificial Intelligence (AI) is an interdisciplinary field, integrating knowledge and methods from computer science, mathematics, philosophy, psychology, economics, neuroscience, linguistics, and biology. Intelligent agents mimic cognitive functions to implement intelligent behaviors such as perception, reasoning, communication, and acting in symbolic and computational models. AI is used in a wide range of narrow applications, from medical diagnosis to speech recognition to bot control.

The autonomous single, multiple, and adversarial agents that students build in this course will support fully observable and partially observable decisions in both deterministic and stochastic environments. Topics covered include search, constraint satisfaction, Markov decision processes, planning, knowledge representation, reasoning under uncertainty, graphical models, and reinforcement learning. The techniques and technologies mastered here will provide the foundational knowledge for the ongoing study and application of AI in other applications across practice areas.

Industry Applications: Professional practice in artificial intelligence focuses on the design and deployment of AI systems. Applications span large-scale optimization at companies like Google and Microsoft, generative and reasoning models such as ChatGPT, Claude Code, and Gemini, and embodied AI in robots, drones, humanoids, and autonomous vehicles. In aerospace, defense, and healthcare, organizations like Boeing, Lockheed Martin, and Mayo Clinic apply planning, constraint satisfaction, and probabilistic reasoning for mission planning and clinical decision support. Professionals collaborate closely with domain experts, systems engineers, and risk or compliance teams to ensure reliable and ethical AI deployment. Advanced applications include text and image generators, and Agentic AI systems that are goal-oriented and proactive, capable of planning and executing multi-step tasks autonomously in domains like software development agents, research copilots, autonomous vehicles, robotic systems, and enterprise workflow automation.

AIM 5007 Neural Networks and Deep Learning

Prerequisites: Machine Learning

Data scientists have been able to leverage better algorithms on faster hardware optimized with graphical processing units to deliver improved performance and accuracy in whole classes of applications that had been previously commercially unviable. The biggest beneficiaries are applications that require unstructured data, such as audio and/or video processing. Deep neural networks have also provided gains for other complex applications, from recommendation systems to natural language processing. This course builds on the concepts in machine learning to train multi-layered neural networks.

Industry Applications: Deep learning powers high-dimensional pattern recognition and autonomous perception across sectors by leveraging GPU-accelerated architectures to derive insights from complex data. In healthcare, convolutional neural networks automate anomaly detection in MRI, CT, and X-ray imaging. Transformers and large language models such as ChatGPT, Claude Code, and Gemini enable advanced NLP, code generation, multimodal reasoning, and conversational systems. Emerging embodied and agentic AI applications include vision-language models, coding agents, and autonomous robots, drones, humanoids, and vehicles. Deep learning also drives retrieval-augmented generation (RAG) with fine-tuned code assistants, next-generation search, marketing, and creative content production using generative and diffusion models.

AIM 5008 Capstone for Artificial Intelligence and Machine Learning

The Capstone integrates prior coursework, research, colloquia, and professional experience, and provides the opportunity to synthesize theory with practice in an applied project, thesis, approved internship, or equivalent activity. Examples include developing an AI application or methodology, publishing a research paper at a peer-reviewed conference, or creating a startup company through YU’s Innovation lab-though students may propose other related projects based on their interests. The Capstone will include four components: a brief proposal and project schedule; the main deliverable (e.g., thesis, conference paper, working system with analysis/code/data); and a final presentation to the student and faculty body. Faculty will provide students with mentorship and feedback at each stage of the work.

Industry Applications: Capstone projects replicate professional practice in consulting, research, and product development, producing deliverables comparable to those in corporate R&D labs, healthcare analytics teams, financial institutions, and startups. Advanced applications include large language and video models, Claude Code, ChatGPT, Gemini, embodied and agentic AI systems, vision-language models, coding agents, RAG architectures, and next-generation autonomous and generative technologies.

AIM 5011 Natural Language Processing

Prerequisites: Machine Learning; Neural Networks and Deep Learning

Natural Language Processing lives at the intersection of machine learning, artificial intelligence, and linguistics. It is the key to unlocking vast amounts of human-generated, unstructured data. The increased availability of corpuses of text data, the wide availability of cheap distributed systems, improvements in neural network algorithms, and increased access to graphical processing units (GPUs) have improved the performance and accuracy of entire families of once computationally intractable problems, making these commercially feasible. This course explores a series of text and voice processing use cases, including sentiment analysis and topic modeling. It is the key to unlocking vast amounts of human-generated, unstructured data. Along the way, students gain experience working with supervised and unsupervised methods using both machine learning algorithms and deep neural networks.

Industry Applications: Natural language processing capabilities are deployed in search, document analysis, conversational systems, and domain-specific applications such as sentiment analysis, semantic search, summarization, information extraction, and question answering. Technology companies like IBM and Google integrate NLP and modern large language models (LLMs) such as GPT-style transformers, Claude-like code assistants, ChatGPT-class models, and Gemini-style multimodal systems into enterprise and consumer products for translation, code generation, and content creation. In non-technology domains, including law firms, healthcare providers, and financial institutions, NLP supports document review, compliance monitoring, clinical text analysis, and knowledge access via retrieval-augmented generation (RAG) and next-generation search. Emerging embodied and agentic AI applications leverage vision-language models, coding agents, and autonomous decision systems to connect language with perception and action across robots, drones, humanoids, and autonomous vehicles.

AIM 5012 Data Visualization

Data scientists depend on data visualizations for their own exploratory analysis to support their modeling decisions--the mind can process visual information faster than numbers. Data visualization is also important to inform—and often to persuade—other people about what can be inferred from the data. These explanatory visualizations often require higher production values, interactivity, and guiding text. In this course, students apply the concepts, methods, and best practices of data visualization to create reproducible, code-based exploration and explanatory data visualizations.

Industry Applications: Data visualization skills are integrated into analytics, reporting, and communication roles across sectors. Organizations such as Tableau and McKinsey & Company rely on visualization to support executive decision-making and client communication. In journalism and public institutions, including The New York Times, visualization professionals translate complex data into accessible narratives. These roles are typically embedded within analytics or communications teams.

AIM 5014 Special Topics (1–3 credits)

Current Special Topics include:

— AI and the Future of Work and Collaboration

This course explores how to think, create, and collaborate in a world shaped by generative AI. Students confront what's at stake: the risk of "same-ness," the erosion of personal voice, and the challenge of maintaining authorship in an AI-assisted world. Through hands-on exercises and critical reflection, you will develop core habits—questioning, iteration, and productive disagreement—that lay the foundation for intentional, human-centered work. Students will experiment with AI as both creative partner and constructive critic, building fluency while defending their own ideas and aesthetic choices.

— AI for Extended Reality: Digital Twins and Spatial Intelligence

Course Description: This project-driven special topics course introduces AI methods that power modern extended reality (XR) experiences, with an emphasis on digital twins and spatial intelligence. Students will build an end-to-end XR prototype that integrates perception (e.g., scene understanding, object detection/tracking), 3D representation (meshes, point clouds, neural fields), and interactive intelligence (LLM/VLM-driven agents, multimodal interfaces). The course bridges theory and implementation through weekly milestones and a final demo, covering practical pipelines for real-time XR: sensing → world modeling → simulation/interaction → evaluation. Students finish with a XR/digital twin project suitable for research or industry.

— AI for Safety and Security

With high-visibility safety and security failures resulting from the use of Agentic AI, such as the recent AWS outage and a SEV1 security incident at Meta, companies with AI solutions such as Amazon, Anthropic, Google, Meta, Microsoft, and OpenAI are increasingly focused on the building of "AI Defense Stacks," such as Claude Mythos, to protect against both external hackers and internal system failures. This course provides a comprehensive introduction to the security challenges of modern AI systems. We examine how vulnerabilities can be introduced during system architecture design, model development, training, and deployment and how bad actors can exploit these vulnerabilities via prompt injection, adversarial inputs, data poisoning, and model extraction. Students will also learn about common AI misapplications and the risks introduced by multi-agent collaboration. Alongside these threats, students will examine emerging defenses such as secure architectures, verifiable training, fingerprinting, and prompt-level protections, gaining a deeper understanding of how to assess and improve AI system security.

— Artificial General Intelligence

This is an advanced course which provides the opportunity to learn special interest courses on emerging theory, phenomena, and technologies in the field of artificial intelligence, machine learning, data science, and big data.

— Foundations and Practice of Effective Coding with AI

Course Description: This course provides graduate students with both theoretical foundations and practical methodologies for Effective Coding with AI. Students will study the core architecture and mechanisms underlying AI models, including transformers, self-attention, tokenization, embedding representations, and autoregressive generation, to develop a principled understanding of how AI coding assistants (e.g., Claude, ChatGPT, GitHub Copilot, Cursor) produce, evaluate, and refine code. The course further covers key paradigms such as prompt engineering, retrieval-augmented generation (RAG), fine-tuning, and agentic workflows, equipping students to critically assess model capabilities, diagnose failure modes, and select appropriate strategies for diverse development tasks.

Grounded in this theoretical framework, students will apply their knowledge through a progressive, project-oriented learning arc: (1) leveraging AI tools to rapidly prototype and deploy a professional portfolio application; (2) building and evaluating a sentiment analysis pipeline that bridges classical

NLP techniques with modern LLM-based approaches; and (3) integrating these components into a full-stack, AI-powered application with real-time inference capabilities. Throughout the course, students will also examine ethical considerations including attribution, bias, intellectual property, and responsible AI use in software engineering. By completion, students will possess both the conceptual depth to understand and adapt to the rapidly evolving AI landscape and the practical skills to design and build their own intelligent applications across a wide range of domains.

— Model Context Protocol for AI Agents
Course Description: This rigorous, project-based course treats AI agency as a formal software engineering discipline, focusing entirely on the Model Context Protocol (MCP) to architect autonomous, secure personal and development AI systems. The curriculum guides students from developing foundational local utility servers to engineering sophisticated, multi-agent systems capable of complex coding, tool usage, skill learning and memory enhancements. A heavy emphasis is placed on context economics, drastically reducing token consumption through dynamic local code execution, agent skill optimization and modern digital security. The course also focuses on teaching students to defend against and mitigate critical vulnerabilities like tool poisoning and indirect prompt injections. The program equips developers with the advanced skills necessary to transition from traditional coding to managing production-ready, highly optimized AI agent workflows in professional environments.

— Modern Intelligent Database Systems
Course Description: The path toward artificial general intelligence (AGI) demands data infrastructure far beyond what traditional database systems provide. This course grounds students in modern database design beginning with relational theory and PostgreSQL and extending to vector, graph, and document stores. Students learn to create pipelines suitable for retrieval-augmented generation (RAG), multimodal search, and real-time inference.

Through weekly milestones, students build a full-stack application backed by a hybrid data architecture. The course examines how schema design, indexing, consistency guarantees, and query optimization shape the performance of intelligent systems, while also exploring the generative architectures (transformers, diffusion models, and multimodal frameworks) these databases support.

By completion, students will be equipped to architect database systems and data pipelines powering the next generation of intelligent applications.

AIM 5015 Deep Reinforcement Learning

This course prepares students for hands-on skills in deep reinforcement learning with Python and the PyTorch framework. Combining theory and practice, the course guides students step by step to analyze, design, implement, and present real-world deep reinforcement learning projects. Topics include Markov Decision Processes, dynamic programming, temporal-difference learning, deep Q-networks, policy gradients, actor-critic methods, and continuous action space algorithms. The course also covers the latest reinforcement learning tools and techniques.

AIM 5016 Computer Vision

This course covers the foundations of computer vision, from image formation and filtering through modern deep learning approaches for recognition, generation, and 3D understanding. Topics include image processing and linear filtering, edge detection, image pyramids, neural networks, convolutional neural networks, transformers, object recognition, generative models, 3D reconstruction, motion estimation, and vision-language models. Students develop both mathematical foundations and practical skills through homework and a final project.

AIM 5017 Neural Computation

Neural computation spans the full arc from biological brains to engineered systems. This course begins with the anatomy and physiology of the human brain, traces the developmental program that wires 86 billion neurons, formalizes the plasticity rules that enable learning, and examines the theories of consciousness that define what it means for a system to be aware. The second half turns outward: connectome-to-function mapping, whole-brain emulation, biological computing with living neurons, brain–computer interfaces, and the specialization of silicon hardware from CPUs through GPUs to model-specific ASICs.

AIM 5999 Independent Study in Artificial Intelligence and Machine Learning

The course provides flexibility to learn more about a topic of interest outside of the formal course setting. The subject should be chosen in consultation with a faculty advisor who acts as the student's supervisor, and with the permission of the program director. The student is required to submit a course contract describing the course of study and its specific learning objectives.

Industry Applications: Independent study aligns with apprenticeship-style models found in research institutions, consulting environments, and specialized industry roles, with mentorship that emphasizes depth, autonomy, and alignment with organizational or domain-specific objectives. Advanced agentic AI systems use large language models and related architectures to plan, decide, and act toward specified goals across applications such as virtual assistants and chatbots, code assistants and autonomous coding agents, research and knowledge copilots with RAG, creative content generation, robotics and autonomous vehicles, drones and humanoid robots, predictive maintenance and operations optimization, clinical decision support, and multi-agent systems that coordinate complex workflows.

Updated: July 2026