Fundamentals of Operationalizing AI

Instructor: Professor Anand Rao

Course Description

Artificial Intelligence (AI) and Generative AI (GenAI) are revolutionizing industries by enhancing efficiency, decision-making, and innovation. However, many organizations encounter significant obstacles in realizing and scaling AI's full potential. Recent surveys reveal that nearly 75% of companies struggle to achieve and scale value from AI initiatives, with approximately 70% of challenges stemming from people and process issues, 20% from technology problems, and only 10% related to AI algorithms. Furthermore, over 50% of executives express concerns about integrating GenAI into their operations, citing issues such as limited traceability and irreproducibility of outcomes, which could lead to suboptimal or even illegal decision-making.

This course is designed to tackle these challenges head-on, providing graduate-level students with a comprehensive understanding of the AI lifecycle. Students will learn how to navigate the complex process of identifying which business tasks should be automated through AI and which decisions should be augmented using AI. The course introduces practical frameworks essential for making these strategic decisions and successfully implementing AI solutions.

Throughout the course, students will engage deeply with each stage of the AI lifecycle. They will learn to identify and prioritize high-impact AI use cases, conduct thorough cost-benefit analyses, and design strategic roadmaps aimed at maximizing return on investment (ROI). The curriculum blends theoretical knowledge with hands-on experience using industry-standard tools such as Jupyter Lab, Docker, Kubernetes, Kubeflow, Kafka, and Evidently. These tools are critical for overcoming the common pitfalls associated with AI deployment and preparing students to scale AI systems in real-world environments.

The course's practical orientation is further enhanced through case studies that serve as a foundation for class discussions. These case studies provide students with the opportunity to analyze real-world AI applications, assess the challenges involved, and understand the decision-making processes behind successful implementations. Additionally, two guest lectures from seasoned industry practitioners will offer firsthand insights into the practical challenges of AI deployment across various sectors.

A strong emphasis is placed on governance and trust, equipping students with the knowledge to develop ethical, transparent, and effective AI systems. Students will learn how to integrate AI into organizational processes, assess talent and skill gaps, and create strategies to build the necessary capabilities for sustained AI-driven innovation.

This course is essential for students aspiring to careers as AI engineers, AI analysts, or AI governance experts. It is equally invaluable for business and technology students who wish to understand how to manage the development and deployment of AI systems. By the end of the course, students will possess a well-rounded, practical understanding of AI system management, enabling them to lead AI-driven projects and drive innovation across industries.

Learning Objectives

Students should be able to:

1. **AI Lifecycle Mastery:** Develop a thorough understanding of the AI system lifecycle, including the identification of business tasks for automation or augmentation, and effectively manage data preparation, model development, deployment, and maintenance to ensure alignment with organizational goals and industry best practices.
2. **Strategic Value Realization:** Identify and assess business needs, conduct detailed cost-benefit analyses, and design strategic roadmaps that prioritize high-impact AI initiatives. Develop AI models that deliver maximum ROI and are closely aligned with business objectives.
3. **Effective AI Operationalization:** Design and implement strategies to successfully deploy, monitor, and maintain AI models at scale, with a focus on overcoming common deployment challenges, ensuring system reliability, building trust, and fostering continuous improvement.
4. **Ethical Governance and Risk Management:** Establish and enforce comprehensive governance frameworks that promote ethical, transparent, and effective AI operations. Address potential risks proactively while fostering stakeholder trust and confidence in AI systems.
5. **Organizational Integration and Skill Building:** Evaluate organizational talent and process needs, and create actionable strategic plans to address skill gaps. Integrate AI seamlessly into business processes and build the necessary capabilities to sustain AI-driven innovation within the organization.
6. **Practical Tool Proficiency:** Gain hands-on experience with key AI lifecycle tools—Jupyter Lab, Docker, Kubernetes, Kubeflow, Kafka, and Evidently—preparing you to effectively scale and deploy AI systems in real-world environments.

Learning Resources

The following textbooks will be used as reference for the topics covered. Each topic will also have selected reading materials. The recommended book for purchase is:

1. [Designing Machine Learning Systems](#) by Chip Huyen, O'Reilly Media, May 2022.
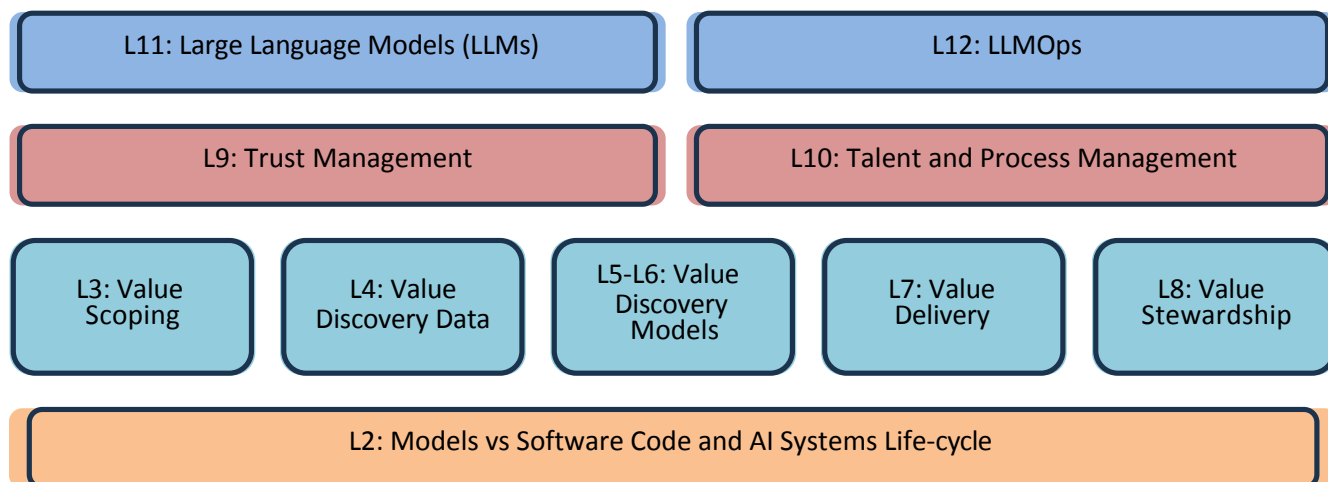
In addition, the following books also cover similar topics to this course.

2. [Operationalizing AI: How to accelerate and scale across people, process, and platforms](#) by John J. Thomas, William Roberts, and Paco Nathan, O'Reilly Media, March 2021.
3. [Operating AI](#) by Ulrika Jagare, Wiley, May 2022.
4. [AI Engineering](#) by Chip Huyen, O'Reilly Media, December 2024.

Assessments

The final course grade will be calculated using the following categories:

Assessment	Percentage of Final Grade
Class Participation	10%
Two Class Quizzes	30%
Individual Assignment	30%
Team Project Presentation	30%
Total	100%

L11: Large Language Models (LLMs)	L12: LLMOps

L9: Trust Management	L10: Talent and Process Management

L3: Value Scoping	L4: Value Discovery Data	L5-L6: Value Discovery Models	L7: Value Delivery	L8: Value Stewardship

L2: Models vs Software Code and AI Systems Life-cycle

Course Schedule

Week	Theme/Topic	Learning Outcomes Addressed	Assignments Due
1	L1: Introduction & Overview	Clear understanding of the scope of the course, learning objectives, topics covered, student evaluation, and relevance of course to broader practice. *Case Study 1: Scaling AI*	
	L2: Models vs Software code and AI Systems Life Cycle	Differentiate AI/ML Models and Software Code Identify and Mitigate Consequences of Model-Code Confusion Distinguish Between AI Models and AI Systems Understand the AI System Lifecycle and Tools for Lifecycle Management *Case Study 2: Data Scientists are from Mars & Software Developers are from Venus*	
	Recitation	Introduction to Shell, Anaconda, and Hugging Face environments	(Ungraded) Quiz-L1&L2 (Ungraded) Case Study Assignment 1
2	L3: Value Scoping	Perform needs analysis of businesses matching problem spaces with solution spaces AI Use Case & Ethics Analysis Cost-Benefit Analysis specifically focused on automation and augmentation of AI Model and Roadmap Design *Case Study 3: ROI of AI: Navigating the road to value realization*	
1	L4: Value Discovery - Data	Data Extraction focused on batch and stream processing Data Pre-processing with feature engineering and feature stores	
	Recitation	Introduction to Kafka Stream processing and Faust python package for stream processing	(Ungraded) Quiz-L3&L4 (Ungraded) Case Study Assignment 2 &3

3	L5: Value Discovery - Models	Understand the process of model selection and training based on type and volume of data Analyze key metrics for evaluating model performance, including accuracy, precision, and recall and assessing model effectiveness in meeting business objectives. *Case Study 4: AI transformation journey*	
	L6: Value Discovery – Ensemble Models	Apply ensemble methods for Model Optimization including techniques such as bagging, boosting, and stacking, to improve model accuracy and reliability. Understand when and why to apply each technique based on data and business requirements.	(Graded) Quiz 1
	Recitation	Introduction to Feast, an open- source feature engineering tool and MLFlow	(Ungraded) Quiz-L5&L6 (Ungraded) Case Study Assignment 4
4	L7: Value Delivery	Understand the model deployment process including packaging, environment setup, prediction architectures and validation. Understand how to integrate AI models into business processes, ensuring their smooth adoption and operationalization. *Case Study 5: Engineering stability in the face of unprecedented growth*	
	L8: Value Stewardship	Understand the importance of real-time monitoring for AI systems focusing on data and model drifts t0 ensure ongoing business and operational requirements. Implement continual learning and retraining for Business Relevance.	
	No Recitation		(Ungraded) Quiz-L7&L8 (Ungraded) Case Study Assignment 5
5	Guest Lecture (TBD)	Identify and solve challenges of applying operationalization in the industry.	(Graded) Individual Assignment Due
	L9: Trust Management	Identify and mitigate AI risks, such as bias, security vulnerabilities, and privacy issues in AI systems.	

		Design and apply governance frameworks that ensure continuous monitoring, fairness, transparency, and compliance, adapting policies to emerging risks and regulatory requirements.	
	Recitation	Introduction to Dockers, Kubernetes and containerization of models.	(Ungraded) Quiz-L9 (Graded) Team Assignment (Interim)
6	L10: Talent and Process Management	Structure cross-functional AI teams for Lifecycle Management. Adopt Agile Methodologies for AI development and deployment combining software agile and data science agile methods.	
	Guest Lecture (TBD)	Identify and solve challenges of applying operationalization in the industry.	
	Recitation	Introduction to Evidently, an open-source model monitoring tool.	(Ungraded) Quiz-L10
7	L11: Large Language Model	Analyze role and capabilities of Foundation Models, Generative AI, LLMs, and multi-modal models and their transformative potential. Understand advanced techniques, such as, prompt engineering, Retrieval-Augmented Generation (RAG), and fine-tuning within the Generative AI Lifecycle framework. *Case Study 6: Strategic pivot to Generative AI*	(Graded) Quiz-2
	L12: Large Language Model Operations	Assess generative AI and LLM applications and their potential to generate value. Explore emerging LLM architectures and the LLMOps tools landscape. Understand critical elements of a scalable, reliable, and efficient LLMOps Deployment.	
	Recitation/Final		(Ungraded) Quiz-L11&L12 Final class presentations or Final revision