

Practical AI and Machine Learning for Materials Science

This intensive 12-hour short course introduces practical artificial intelligence (AI) and machine learning (ML) tools for materials scientists and engineers. Designed for researchers, engineers, and advanced students in ceramics and related fields, the course focuses on how to apply modern data-driven methods to real materials problems such as property prediction, materials discovery, and process optimization.

Rather than emphasizing abstract theory, the course prioritizes hands-on workflows using Python and modern AI tools. Participants will learn how to structure materials data, build predictive models, use large language models (LLMs), and implement optimization strategies. By the end of the course, attendees will be able to prototype their own AI-driven materials workflows and understand when and how these approaches provide value.

Class Dates	Topics / Activities During	Assignments, Notes, Demos
10/13/2026	What is Materials Informatics (and what is not) High-value use cases in ceramics AI/ML workflow: build → measure → learn Materials data types and problem framing	Optional Pre-watch: Introduction / What is Materials Informatics? (playlist lecture) Emphasis: why materials informatics matters (not hype—real use cases) Key takeaway: most ML projects fail due to poor problem formulation, not algorithms Demo: pandas-based dataset exploration from GitHub notebooks Optional: students identify a problem from their own work that could be ML-ready
10/14/2026	Feature engineering for materials Composition-based feature vectors (CBFV) Domain knowledge vs automated features	Optional Pre-watch: feature engineering + CBFV lectures from playlist Key idea: representation matters more than model choice Use GitHub notebook for composition-based feature generation Hands-on: convert raw compositions → feature vectors Note: discuss limitations of composition-only models (no structure, processing)
10/15/2026	Supervised learning (regression + classification) Model selection: linear, random forest, boosting Overfitting, validation, and metrics	Optional Pre-watch: supervised learning lectures (regression/classification) Focus: practical modeling—not math derivations Demo: scikit-learn workflow from GitHub (train/test split, model training) Hands-on: predict a property (e.g., bandgap, modulus, etc.) Key lesson: validation > accuracy (avoid fooling yourself)
10/20/2026	Model interpretation and trust Feature importance and extracting insight Uncertainty and failure modes	Optional Pre-watch: “best practices” lecture (data skepticism, cleaning) Core message: garbage in → garbage out (most important lesson in the course) Hands-on: feature importance + model interpretation

		<p>Demo: extracting physical insight from model outputs Discussion: when NOT to trust a model</p>
10/21/2026	<p>Optimization and materials discovery Bayesian optimization and active learning Efficient experimental design</p>	<p>Optional Pre-watch: optimization / active learning content (GitHub + lecture examples) Focus: how ML actually accelerates discovery, not just prediction Demo: Bayesian optimization loop (build → suggest → update) Hands-on: simulate iterative materials optimization Key concept: exploration vs exploitation in real experiments</p>
10/22/2026	<p>LLMs for materials science (RAG, data extraction) Agents and automation in research workflows End-to-end AI workflow</p>	<p>Optional Pre-watch: LLM / AI tools lectures (from your newer content + demos) Topics: RAG, tool use, agents in scientific workflows Demo: LLM extracting materials data from text / literature Demo: simple agent workflow (data → reasoning → output) Capstone: connect full pipeline (data → model → decision)</p>