MSE 404/MSE504 Course Syllabus

**Faculty Contact Information:**

Instructor: Prof. Sergei V. Kalinin,

Office: 314 IAMM

E-mail: [sergei2@utk.edu](mailto:sergei2@utk.edu)

Teaching Assistant: Aditya Raghavan, [araghav4@vols.utk.edu](mailto:araghav4@vols.utk.edu)

**Instructor Availability:**

Please don’t hesitate to email me with updates, questions, or concerns. I will typically respond within 24 hours during the week and 48 hours on the weekend (unless on travel). I will notify you if I will be out of town and if connection issues may delay a response. Do not hesitate to contact Aditya with all matters pertaining to the course.

**Meeting Time:** 10:20 am - 11:10 am MWF, Ferris Hall 502

The lectures and materials will be posted on Canvas and at GitHub: <https://github.com/SergeiVKalinin/MSE_Fall_2025>

**Office Hours:**

Friday 1:30 - 3:00 PM are preferred for 1:1 meetings to discuss any course related item. Additional time can be found upon request. Please schedule via email.

**Course Description:**

ML and Artificial Intelligence (AI) have become the defining features of our era, revolutionizing domains from medicine to autonomous driving. In the past decade alone, ML has rapidly advanced dominating *in silico* applications such as image analysis, recommender systems, and the development of large language models. Now, envision the forthcoming decade as the era of ML's impact on real-world challenges. With ML, we can unlock new frontiers in materials and process optimization, pave the way for groundbreaking materials discovery, and delve into the intricacies of physics at the nanoscale. ML agents are becoming a part of the future workforce of automated and hybrid labs and manufacturing facilities.

This course is designed to equip you with the fundamental knowledge of ML methods while delving deep into the realm of practical applications. By enrolling in this course, you will embark on an exciting exploration of ML's potential in solving real-world scenarios. Discover how ML can revolutionize Materials Science, unraveling complex problems and uncovering innovative solutions. Join this course today and be at the forefront of this transformative field, where the convergence of ML and Materials Science promises a future brimming with limitless possibilities. The instructor (Kalinin) has worked on ML applications in imaging and materials science for >15 years at ORNL, and spent a year at Amazon (special projects), thus bringing a wealth of experience in ML applications for real-world problems.

**Course Outline**

1. **Introduction**
   * Machine Learning for Materials
   * History of ML and Scientific Data
   * Python Ecosystem, Data, and LLMs
2. **Classification**
   * Representing the World, Classification, and Decision Trees
   * Decision Trees, Forests, and Flowers
   * Simple Perceptron, Adaline, and Logistic Regression
   * More Classifiers and Performance Metrics (ROC & AUC)
   * Training Classifiers: Ensembles and Boosting
   * How Materials Are Discovered and Made
3. **Clustering**
   * Fundamentals of Clustering
   * Clustering on Imaging and Spectroscopic Data
4. **Dimensionality Reduction**
   * Linear Dimensionality Reduction Methods
   * PCA, LDA, QDA for Spectra and Images
   * Linear Dimensionality Reduction for Images
   * Other Techniques: CCA, ICA, …
5. **Decision Making**
   * A/B Testing and Bandits (2 lectures)
   * Contextual Bandits and ML Agents (2 lectures)
6. **Multistep Decision Processes**
   * Tree based methods: from search to A\*
   * Heuristics in Multistep Decisions
   * Dynamic Programming
7. **Reinforcement Learning**
   * Markov Processes
   * Reinforcement Learning Fundamentals (2 lectures)
8. **Deep Learning**
   * Multilayer Perceptron
   * Deep Convolutional Neural Networks (2 lectures)
   * Variational Autoencoders
   * Invariant Variational Autoencoders (2 lectures)
   * Explainable ML (2 lectures)
9. **Causality**
   * Causal Inference Methods (2 lectures)
10. **NLP and LLMs**
    * Natural Language Processing (NLP)

**The case studies throughout the course will be:**

* Applying ML to imaging and characterization from post-acquisition data analysis to automated microscopes
* Materials synthesis and characterization: from automated tools and workflow design to automated labs and user facilities
* Learning materials informatics: Materials Project, JARVIS, AFlow
* Theory dive: QM9 data set
* LLMs for text analytics and data mining

**Prerequisites:**

To be successful in this course you will need a general background in materials science, chemistry, or condensed matter physics. Homeworks, midterms, and finals will allow flexibility in choosing the problems that match your domain expertise. Python or similar programming experience, while not essential, will be extremely useful. Students without any prior programming experience should expect to spend extra time outside of class learning basic skills (but ChatGPT will help!).

**Student Learning Outcomes:**

1. This course aims to provide students with the skills needed to move from data to decisions. Understanding of Materials Science and Machine Learning: Students should expect to gain a solid understanding of materials science, machine learning, and the intersection of the two. They will learn how machine learning can drive the discovery and optimization of materials, which is essential in numerous industries, including technology, manufacturing, energy, and more.
2. Proficiency in Machine Learning Techniques: The course is designed to provide students with basic knowledge of principles of various machine learning techniques, including supervised, unsupervised, and active learning. These skills are essential for careers that involve data analysis, prediction modeling, and artificial intelligence.
3. Practical Application of Machine Learning: Students will learn how to apply machine learning techniques to real-world materials science problems. This includes using these techniques for property prediction, imaging and characterization, and process optimization. These practical applications prepare students for careers where they will need to apply theoretical knowledge to solve practical problems.
4. Experience with Advanced Machine Learning Concepts: The course covers advanced topics such as linear dimensionality reduction methods, deep convolutional neural networks, variational autoencoders, and generative models. This knowledge can help students be at the forefront of the AI field, making them valuable assets to companies investing in these areas.
5. Preparedness for the Future of Science and Industry: With insights into automated labs, large language models in scientific workflows, and federated tools and workflows, students will be prepared for the future of industry. These skills are increasingly important as companies automate processes and incorporate AI into their workflows. This knowledge can help students stand out in the job market and be prepared for the careers of the future.

**Value Proposition:**

1. You are interested in ML and AI and would like to try it hands-on on real world problems from materials science, chemistry, condensed matter physics, and microscopy
2. Learn the basics of the ML methods and build upon this knowledge - from simple principal component analysis to large language models and causal analysis.
3. Explore how ML is being adopted by industry - from IT leaders such as Amazon, Google, and Meta to instrumental, chemical, and materials companies.
4. Learn why the next decade of ML will be transition from purely in-silico to real-world materials and device applications, and be a part of this transition
5. And learn to work backwards from real-world problems to solution.

**Learning Environment:**

The class will be delivered as in-person lectures and recorded by Zoom. The Jupyter notebooks, code libraries, and videos will be provided. Weekly programming exercises will be assigned via Google Colabs (please submit to [sergei2vk@gmail.com](mailto:sergei2vk@gmail.com) ) and those students wishing to interact with the instructor in person should attend office hours (please set by e-mail [sergei2@utk.edu](mailto:sergei2@utk.edu) ).

**Use of ChatGPT:**

Strongly encouraged both for programming and written assignments. However, the students have to be aware of the limitations of the generative models.

**Canvas:**

All course details, assignments, lecture notes and announcements will be available on Canvas. You are required to be aware of anything posted to the course website. Please update your canvas notification settings.

**Reference Materials:**

I will provide copies of lecture notes, presentations, and Colabs on GitHub and Canvas. There is no specific textbook for the course and we will take material from a variety of sources including:

* Andrew Bird et al, *Python Workshop*, <https://www.packtpub.com/product/the-python-workshop/9781839218859>
* Sebastian Raschka, *Machine Learning with PyTorch and Scikit-Learn*, <https://subscription.packtpub.com/book/data/9781801819312/1>
* Rowel Atienza, *Advanced Deep Learning with TensorFlow 2 and Keras - Second Edition*, <https://www.packtpub.com/product/advanced-deep-learning-with-tensorflow-2-and-keras-second-edition/9781838821654>
* (Optional) [Alaa Khamis](https://medium.com/@alaakhamis?source=post_page---byline--ec1feb6b5044---------------------------------------), *Optimization Algorithms: AI Techniques for Design, Planning, and Control Problems,* <https://www.manning.com/books/optimization-algorithms>
* (Optional) Peter Norvig, *Artificial Intelligence: A Modern Approach, Global Edition*, <https://www.amazon.com/Artificial-Intelligence-Modern-Approach-Global/dp/1292401133>

**Grading & Policies:**

* Homework assignments 40%
* Mid-terms (2) 30%
* Final Project & Presentation 30%

Course grades will be assigned on the following basis:

95.0 - 100.0 A

90.0 – 95.0 A-

85.0 - 89.99 B+

80.0 - 84.99 B

75.0 - 79.99 C+

70.0 - 74.99 C

60.0 - 69.99 D

00.0 - 59.99 F

This course grade basis may (at the instructor’s discretion) be shifted uniformly down, should the overall performance of the class require it. This course grade basis will not be shifted up. (That is, if an exam proves to be too hard and the average is low, an 89% may make an A. However, if an exam proves to be too easy and the average is high, a 90% will always make an A.)

**Assignments:**

* All assignments will be submitted as Google Colabs to [sergei2vk@gmail.com](mailto:sergei2vk@gmail.com) (using Gmail allows to avoid lengthy authorization process)
* Late assignments will be accepted without penalty; however, keep in mind that the level of problems is increasing, and the UT final submission day is fixed
* Students can work together to solve homework assignments. However, each student must turn in his/her own work in his/her own notebook.
* Instances of plagiarism will be addressed as stipulated by University guidelines. The use of ChatGPT or any other LLMs does not constitute plagiarism in this course. Please do not force me to have to deal with plagiarism cases. Remember, you are here to learn.

**Important Dates:**

The final project will be due on our scheduled final exam date.

**Religious Holidays:**

Students have the right to practice the religion of their choice. If you need to miss class to observe a religious holiday, please submit the dates of your absence to me in writing via email by the end of the second full week of classes. You will be permitted to make up work within a mutually agreed-upon time.

**Statement on Civility & Community:**

The Department of Materials Science and Engineering at the University of Tennessee is committed to creating an environment that welcomes all people, regardless of their identities. We value the diversity that enriches our department. We understand the importance of free and open dialogue that includes the free exchange of ideas. We do not tolerate uncivil speech or any form of discourse that infringes on others’ rights to express themselves, or has a negative impact on their education, or work environment. We actively promote an environment of collegiality and an atmosphere of mutual respect and civility. We understand that respect includes being considerate of others’ feelings, circumstances, and their individuality. We recognize the necessity of a civil community in realizing the potential of individuals in teaching, learning, research, and service. We believe these values extend beyond the department into our work within physics regionally, nationally, and internationally, as well as work and studies in the university, and the broader community. We encourage all members of the department to intervene and report any incidents involving bigotry, or that violate the university code of conduct.

**COVID-19:**

The best source of information for information related to the University of Tennessee’s response to the COVID-19 pandemic can be found online at: <https://www.utk.edu/coronavirus/>.