Syllabus

Monday, April 10, 2023 10:08 PM

Welcome to Applied Machine Learning in Python, taught by Kevyn Collins-Thompson !

This course will introduce the learner to applied machine learning, focusing more on the techniques and methods than on the statistics behind these methods. The course will start with a discussion of how machine learning is different than descriptive statistics, and introduce the scikit learn toolkit through a tutorial. The issue of dimensionality of data will be discussed, and the task of clustering data, as well as evaluating those clusters, will be tackled. Supervised approaches for creating predictive models will be described, and learners will be able to apply the scikit learn predictive modelling methods while understanding process issues related to data generalizability (e.g. cross validation, overfitting). The course will end with a look at more advanced techniques, such as building ensembles, and practical limitations of predictive models. By the end of this course, students will be able to identify the difference between a supervised (classification) and unsupervised (clustering) technique, identify which technique they need to apply for a particular dataset and need, engineer features to meet that need, and write python code to carry out an analysis. This course should be taken after Introduction to Data Science in Python and Applied Plotting, Charting & Data Representation in Python and before Applied Text Mining in Python and Applied Social Analysis in Python.

Participation Strategies

Engaged learning looks different for everybody. In this course, we hope you will define your own measures of success and engage with the material in a way that best suits your needs. We recognize and celebrate the diverse ways learners engage in courses. As you go through this course, we hope you will reflect on your unique skills, needs, and aspirations, and engage in the course material in a way that aligns with your own goals. While the course provides time estimates for completion, you should feel empowered to engage in the material in whatever ways make sense to you.

Course Schedule & Grading Policy

The lectures will provide you with some guidance for completing assignments, but you will need to take initiative and look beyond assignment instructions in order to be successful. You'll need to know how to ask questions in the discussion forums of your peers, and seek out new information through web searches and Stack Overflow. Be sure to also check out the Additional Resources.

If you are not sure what kind of output is required, or think there is a need for more clarity, please head to the course discussion forums. Note that some assignments and in video quizzes may not be mobile friendly.

Some assignments allow you to download and view your fellow learner's code and/or data. If you want to look at the learner's code, we recommend that you open it through the Jupyter notebook system on the Coursera platform as that will be more secure. Please ensure that all data you share is publicly available, since you will be sharing these data with other learners.

Week One: 20%

In Module One, you will be introduced to basic machine learning concepts, tasks, and workflow using an example classification problem based on the K-nearest neighbors method, and implemented using the scikit-learn library. This week's assignment has you work through the process of loading and examining a dataset, training a k-nearest neighbors classifier on the dataset, and then evaluating the accuracy of the classifier and using it to classify new data.

- Quiz: Module 1 Quiz (5%)
- Programming Assignment: Assignment 1 (15%)

Week Two: 25%

In Module Two, you will delve into a wider variety of supervised learning methods for both classification and regression, learning about the connection between model complexity and generalization performance, the importance of proper feature scaling, and how to control model complexity by applying techniques like regularization to avoid overfitting. In addition to k-nearest neighbors, this week covers linear regression (least-squares, ridge, lasso, and polynomial regression), logistic regression, support vector machines, decision trees, and the use of cross-validation for model evaluation. For this week's assignment, you'll explore the relationship between model complexity and generalization performance, by looking at the effect of key parameters on the accuracy of different classification and regression models.

- Quiz: Module 2 Quiz (5%)
- Programming Assignment: Assignment 2 (20%)

Week Three: 25%

In Module Three, you will cover evaluation and model selection methods that you can use to help understand and optimize the performance of your machine learning models. For this week's assignment, you will train a classifier to detect fraudulent financial transactions, analyze its performance with different evaluation metrics, and then optimize the classifier's performance based on different evaluation metrics, depending on the goals of the detection task (e.g. to minimize false positives vs false negatives).

- Quiz: Module 3 Quiz (5%)
- Programming Assignment: Assignment 3 (20%)

Week Four: 30%

In **Module Four**, you will cover more advanced supervised learning methods that include ensembles of trees (random forests, gradient boosted trees), and neural networks (with an optional summary on deep learning). You will also learn about the critical problem of data leakage in machine learning and how to detect and avoid it. The final assignment brings everything together: you will design features for, and build your own classifier on, a prediction problem on a

complex real-world dataset.

- Quiz: Module 4 Quiz (5%)
- Programming Assignment: Assignment 4 (25%)

Course Item	Percentage of Final Grade	Passing Threshold
Week 1 Quiz	5%	80%
Week 1 Jupyter Notebook Assignment	15%	
Week 2 Quiz	5%	80%
Week 2 Jupyter Notebook Assignment	20%	
Week 3 Quiz	5%	80%
Week 3 Jupyter Notebook Assignment	20%	
Week 4 Quiz	5%	80%
Week 4 Jupyter Notebook Assignment	25%	

In-Video Questions (IVQs)

In this course, in-video questions or IVQs may appear during lectures to help you learn as well as assess your understanding of the content. IVQs are optional and do not count towards your overall course grade.

Types of in-video questions

Many of the lectures contain in-video questions (IVQs). These questions are presented in a variety of formats. Some will ask you to write or think about a concept from the video. Others will ask for a short answer. Still others may ask you to choose from a multiple-choice list of answers. If an IVQ is a survey or a poll, you will see a summary of responses from other learners after you respond. You can look at the question again later to see new summary data as more of your peers answer.

Some IVQs also contain runnable code blocks. These IVQs allow you to practice the coding concepts during the lecture. In this course, these types of IVQs will usually be directly followed with the solution code.

Ground Rules

We expect everyone to be mindful of what they say and its potential impact on others. The goal is to have respectful discussions that do not violate the community space created for these conversations. Here are some productive ways to engage in this course:

- **Participate:** This is a community. Read what others have written and share your thoughts.
- **Stay curious:** Learn from experts and each other by listening and asking questions, not making assumptions.
- **Keep your passion positive:** When replying to a discussion forum post, respond with thoughts on what was said, not about the person who posted. Avoid using all caps, too many exclamation points, or aggressive language.
- **Acknowledge discomfort:** The topics discussed in this course might be challenging or hard to talk about. Stick with it and remember the benefits of having these tough conversations that surface from multiple perspectives.
  We expect all learners to abide by our full Learner Engagement Policy. We will specifically be monitoring this course for language that could be considered inflammatory, incivil, racist, or otherwise unacceptable for this learning space, and we will remove language deemed such.

Please note that external study groups on applications like WhatsApp are not affiliated or endorsed by the University of Michigan. We strongly discourage joining external groups and instead recommend interacting with your fellow learners within the platform. Please express caution if you do join or post any personal information in these forums or in these groups. These forums are publicly accessible and any information you post may be collected, published, or used in an exploitative manner (scams, etc).

Academic Honesty

All submitted work should be your own and academic dishonesty is not allowed. Academic dishonesty can be defined as:

- Copying answers
- Copying words, ideas, or other materials from another source without giving credit to the original author
- Copying from your peers within the course
- Employing or allowing another person to alter or revise your work, and then submitting the work as your own
  Please don't share or reuse solutions to assignments which is an academic integrity concern. Please do not:

- Share complete assignment code in the course discussion forums
- Upload completed assignments to public websites with the goal of sharing solutions. (You can share your work and ideas for professional purposes though).
- Take a peer's solution and submit it as your own

Course Support

Questions and discussion of course material should take place within the course itself. Please do not contact instructors or teaching assistants off the platform, as responding to individual questions is virtually impossible. We encourage you to direct your questions to [forum], where your question might be answered by a fellow learner or one of our course team members. For technical help please contact the Coursera Learner Help Center or use the support forums.

Accessibility

We are committed to developing accessible learning experiences for the widest possible audience. We recognize that learners with disabilities (including but not limited to visual impairments, hearing impairments, cognitive disabilities, or motor disabilities) might need more specific accessibility-related support to achieve learning goals in this course.

Please use the accessibility feedback form to let us know about any accessibility challenges such as urgent issues that keep you from making progress in the course (e.g., missing or inadequate alt-text, captioning errors).

Diversity, Equity, Inclusion, and Justice

We welcome all learners to this course. People like you are joining from all over the world and we value this diversity. We strive to create a community of mutual respect and trust, where people from all backgrounds, identities and views are valued and heard without the threat of bias, harassment, intimidation, or discrimination. We pay attention to your feedback, how different types of learners experience this course, and aim to make improvements so the course can best serve everyone. We hope you enjoy learning about topics that are important to you.

From <https://www.coursera.org/learn/python-machine-learning/supplement/6EzLY/syllabus>