New Jersey Institute of Technology

Machine Learning in S2022: **CS 675 002** and **CS 675 004**

Instructor Team

- **Instructor:** Assoc. Prof. Dr. Przemyslaw Musialski
  - **Phone:** 973-596-2869
  - **Email:** przemyslaw.musialski@njit.edu (mailto:przemyslaw.musialski@njit.edu)
  - **Homepage:** https://web.njit.edu/~przem/ (https://web.njit.edu/~przem/)
  - **Office**: GITC 4407
  - **Office Hours**: Tu, Th, 11:00am-12:20pm, or by appointment
- **Teaching Assistants**
  - TBD

Communication

This course uses Canvas for announcements and discussion. If you have questions about the class materials or assignments, requests for clarification, or other issues that may interest the class as a whole, post them to the General Discussion Forum. (https://d.docs.live.net/courses/10170/discussion_topics/17225) **If you have any further questions that you are confident do not belong on Canvas, drop me a message using Canvas Messaging System: https://njit.instructure.com/conversations (https://d.docs.live.net/conversations) .**

Do not write personal emails to me, except in emergency cases! U**se Canvas Messages for all inquiries.**

Instruction Delivery

In Spring 2022, the course will be delivered **face-to-face** in the classroom if not announced otherwise. Student attendance is expected.

Course Description

This course is an introduction to machine learning and contains both theory and applications. Students will get exposure to a broad range of machine learning methods and hands-on practice on real data. Topics include Bayesian classification, perceptron, neural networks, logistic regression, support vector machines, decision trees, random forests, boosting, dimensionality reduction, unsupervised learning, regression, and learning new feature spaces.

Prerequisites

Basic probability, linear algebra, computer programming, and graduate or undergraduate senior standing, or approval of the instructor.

Learning Outcomes

By the end of the course, students should be able to:

- Understand the background of supervised and unsupervised machine learning

- Understand a wide variety of learning algorithms
- Understand how to evaluate machine learning models
- Apply the algorithms to real problems, and optimize their parameters.

Reading Material

Reading material will be posted in the schedule on the fly so please check it regularly.

- **Instructions to get technically ready for the coding part of the course.pdf** (https://njit.instructure.com/courses/21154/files/3002211/download?download_frd=1)
- **Introduction to Numpy** (https://njit.instructure.com/courses/21154/files/3002210/download?download_frd=1)
- **Review of Linear Algebra** (https://www.cs.cmu.edu/~zkolter/course/linalg/linalg_notes.pdf)
- **Review of Probability Theory** (http://cs229.stanford.edu/section/cs229-prob.pdf)
- **Course Material on Google Drive** (https://drive.google.com/drive/folders/11HjTIP6YJ4BTwgLyN2vLDQoeuFUNKtPQ?usp=sharing)
- After each lecture, slides and further reading will be posted on CANVAS.

Textbooks

There will be no required textbooks for the class. Some of the class material, however, will be based on content from the following books (none of which you are required to purchase):

1. Richard Duda, Peter Hart and David Stork, Pattern Classification, 2nd ed. John Wiley & Sons, 2001.
2. Tom Mitchell, Machine Learning. McGraw-Hill, 1997. 1st edition.
3. Christopher Bishop, Pattern Recognition and Machine Learning. Springer 2007.
4. Hal Daumé, **A Course in Machine Learning** (http://ciml.info/)
5. Shai Shalev-Shwartz and Shai Ben-David, **Understanding Machine Learning: From Theory to Algorithms** (http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/index.html)
6. Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2, 3rd edition, Raschka, V. Mirjalili, Packt Publishing, ISBN-10: 1789955750
7. Machine Learning, An algorithmic Perspective, 2ndEdition, Stephen Marsland
8. The Elements of Statistical Learning, 2ndEdition, Hastie, R. Tibshirani, J. Friedman
9. **Linear Algebra and Learning from Data** (https://math.mit.edu/~gs/learningfromdata/) (2019), by **Gilbert Strang** (http://math.mit.edu/~gs) (**gilstrang@gmail.com** (mailto:gilstrang@gmail.com))
   ISBN : 978-06921963-8-0 ---> this is my favorite book :)

Linear algebra resources

- **Linear Algebra Review and Reference** (http://cs229.stanford.edu/section/cs229-linalg.pdf) from Stanford
- The **Linear Algebra** (https://www.deeplearningbook.org/contents/linear_algebra.html) chapter in the Deep Learning textbook.
- **Linear Algebra lectures, MIT** (http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010/index.htm)
- **Linear Algebra and Matrices** (http://www.seas.upenn.edu/~jadbabai/ESE504/LAreview.pdf) , a review
- **Review of some elements of linear algebra** (http://www.seas.upenn.edu/~jadbabai/ESE504/linalg.pdf) , by Fernando Paganini

Probability resources

- David Blei's **review of probability** (http://www.cs.princeton.edu/courses/archive/spring07/cos424/scribe_notes/0208.pdf)

- **Review of Basic Concepts in Probability (https://www.ics.uci.edu/~smyth/courses/cs274/notes/notes1.pdf)** by Padhraic Smyth.
- **A review of probability theory (http://cs229.stanford.edu/section/cs229-prob.pdf)** from Stanford
- The **Probability and Information Theory (https://www.deeplearningbook.org/contents/prob.html)** chapter in the Deep Learning textbook.

Schedule

Schedule subject to adjustments.

Lecture	Week	Date	Topic	Material	HW
1	Week 1	1/18	Introduction, Course "Mechanics"	Slides Video	
2		1/20	Linear Algebra Recap	Slides Video	
3	Week 2	1/25	Linear Regression aka Linear Least Squares	Slides Video	A1
4		1/27	Linear Separability, Decision Boundaries, Perceptron	Slides Video Reading	
5	Week 3	2/1	Linear Regression, Overfitting, Gradient Descent	Slides Video	
6		2/3	Probabilistic View, Maximum Likelihood (MLE), Logistic Regression	Slides Video Reading	
7	Week 4	2/8	Bayesian Theorem, Bayesian Learning, Maximum A Posteriori (MAP)	Slides Video	A2
8		2/10	Naïve Bayes: training and inference	Slides Video	
9	Week 5	2/15	Generative vs Discriminative Models. K-Nearest Neighbors	Slides Slides Video	
10		2/17	Curse of Dimensionality, Decision Trees	Slides Slides Video	
11	Week 6	2/22	Decision Trees and Random Forests, Variance, Correlation	Slides Video	A3
12		2/24	Variance, PCA, Dimensionality Reduction, Clustering, K-Means	Slides Slides Video	
13	Week 7	3/1	Non-Linear Embeddings, Spectral Clustering	Slides Video	
14		3/3	Graph Laplacian, Spectral Clustering	Slides Video	
	Week 8	3/8	Midterm Review	Slides Video	
		3/10	**Midterm Exam**		
		3/15	Spring Recess		
		3/17	Spring Recess		
15	Week 9	3/22	Support Vector Machines	Slides Video	A4
16		3/24	Support Vector Machines and Stochastic Gradient Descent	Slides Video	
17	Week 10	3/29	Kernel Methods and the "Kernel Trick"	Slides Video Reading	
18		3/31	Neuronal Networks - Introduction	Slides Video	
19	Week 11	4/5	Neuronal Networks - Backpropagation	Slides Video	A5
20		4/7	Convolutional Neural Networks (CNN 1)	Slides Video	
21	Week 12	4/12	Autoencoder: Coding, Decoding, Compression	Slides Video	
22		4/14	Convolutional Neural Networks (CNN 2)	Slides Video	
23	Week 13	4/19	Variational Autoencoder (VAE)	Slides Video	A6
24		4/21	Generative Adversarial Networks (GAN)	Slides Video	
25	Week 14	4/26	Recurrent Neural Networks and Transformers (NLP)	Slides Video	

26		4/28	Closure: Advanced Topics, Summary, and Outlook	Slides Video	
		TBD	**Final Exam**		

Grading Policy

The final grade is computed as a weighted sum of the programming assignments (Homework), a midterm exam, and a final exam.

- 6 programming assignments (50%)
- Midterm exam (25%)
- Final exam (25%)
- Active class participation is a bonus

Assignments

Assignments will have several small tasks where selected code needs to be completed (usually only a few lines). Each Assignment has its own detailed instructions. In addition, own research on the details of the implementation needs to be conducted. Each Assignment needs to be completed in 7-14 days and submitted via Canvas. On several assignments, bonus points might be accumulated to come up with lost points in previous tasks.

Grading Scale

The final grade will be composed of 50% programming assignments and 50% exams. The grading scale normalized to 100 is as follows (might be subject to adjustments):

- A: 100-90,
- B+: 90-80,
- B: 80-70,
- C+: 70-60,
- C: 60-50,
- F: 50-0.

Grade Corrections

Check the grades in course work and report errors promptly. Please try and resolve any issue within one week of the grade notification.

Incomplete

A grade of I (incomplete) is given in rare cases where work cannot be completed during the semester due to documented long-term illness or unexpected absence for other serious reasons. A student needs to be in good standing (i.e., passing the course before the absence) and receives a provisional I if there is no time to make up for the documented lost time; a letter (or email) with a timeline of what is needed to be done will be sent to the student. Note that for most cases and I would be resolved within a few days, not months, and not the following semester! Not showing up in the final will probably get you an F rather than an I.

Course Policies

Absence

If you miss a class, it is up to you to make up for the lost time. Missing two exams leads to an automatic F in the course. If you miss one exam, you must contact the Dean of Students (DOS) within 2 working days from the day the reason for the absence is lifted with all necessary documentation. If DOS approves, your missing exam grade will be set equal to the average of the non-missing exam grades.

Collaboration and External Resources for Assignments

Some homework problems will be challenging. You are advised to first try and solve all the problems on your own. For problems that persist you are welcome to talk to the course assistant or the instructor. You are also allowed to collaborate with your classmates and search for solutions online. But you should use such solutions only if you understand them completely (admitting that you do not understand something is way better than copying things you do not understand). Also, make sure to give the appropriate credit and citation.

Honor Code

A set of ethical principles governing this course:

- It is okay to share information and knowledge with your colleagues, but
- **It is not okay** to share the code,
- **It is not okay** to post or give out your code to others (also in the future!),
- **It is not okay** to use code from others (also from the past) for this Assignment!

Any noticed disregard of these principles will be sanctioned as per the Academic Integrity Policy of NJIT (see below).

Late Policy

- There will be a 10% penalty of total regular points for every day an assignment is late.
- Max. late submission is 5 days late.

Academic Integrity

Academic Integrity is the cornerstone of higher education and is central to the ideals of this course and the university. Cheating is strictly prohibited and devalues the degree that you are working on. As a member of the NJIT community, it is your responsibility to protect your educational investment by knowing and following the academic code of integrity policy that is found at: **http://www5.njit.edu/policies/sites/policies/files/academic-integrity-code.pdf** (http://www5.njit.edu/policies/sites/policies/files/academic-integrity-code.pdf) .

Please note that it is the professional obligation and responsibility of the instructor to report any academic misconduct to the Dean of Students Office. Any student found in violation of the code by cheating, plagiarizing or using any online software inappropriately will result in disciplinary action. This may include a failing grade of F, and/or suspension or dismissal from the university. If you have any questions about the code of Academic Integrity, please contact the Dean of Students Office at **dos@njit.edu.** (mailto:dos@njit.edu)