CSE474/574 course webpage Documentation

Release 2020

Varun Chandola

May 08, 2021

Contents:

CHAPTER 1

Syllabus

Computer Science and Engineering, University at Buffalo

Spring Semester 2021

1.1 Instructors

- Varun Chandola (lead instructor; chandola[at]buffalo.edu)
- Deen Dayal Mohan (TA; dmohan[at]buffalo.edu)
- Seokmin Choi (TA; seokminc[at]buffalo.edu)
- Jie Zhang (TA; zhang326[at]buffalo.edu)
- Enshu Wang (TA; enshuwan[at]buffalo.edu)

Note: Students are strongly encouraged to use the Piazza's private messaging option to contact the intructors to ensure that the messages are dealt with promptly.

1.2 Class Website

https://cse.buffalo.edu/~chandola/machinelearning.html

1.3 Meeting times and locations

Every Monday, Wednesday and Friday - 1.50 to 2.40 PM, virtually on Zoom

Note: All lecture videos will be made available after the class. The links will be posted on Piazza and also available here. Please ensure that your video is turned off and the microphone is on mute. Use the zoom reactions and chat to interact with the instructor.

1.4 Office Hours

Who?	When?	Where?
Chandola	Fridays 3.00 PM - 5.00 PM	Virtually on Zoom
Deen	Mondays 6.30 PM - 7.30 PM	Virtually on Zoom
Seokmin	Mondays 9.00 AM - 10.00 AM	Virtually on Zoom
Jie	Thursdays 2.30 PM - 3.30 PM	Virtually on Zoom
Enshu	Tuesdays 1:00 PM - 2:00 PM	Virtually on Zoom

1.5 Prerequisites

CSE 250 and (EAS 305 or MTH 411 or STA 301 or MTH 309).

Note: This course requires a strong background in linear algebra, advanced calculus and statistics. Please refer to the FAQs for more.

1.6 Topic Schedule

Week	Topic	Pre-requisites
1	Introduction and Basics	
Supervised Learning::Linear Models		
1	Linear Regression	Linear Algebra,Gradient Descent Optimization, Matrix Calculus
2	Logistic Regression/Perceptrons	Newton's Method
2-3	Support Vector Machines	Constrained Optimization, Lagrangian Methods
Supervised Learning::Non-linear Models		
4	Non-linear Regression	
4	Regularization	
5	Kernel Regression	
5	Kernel Support Vector Machines	
6-7	Neural Networks	
Statistical Learning		
8	Generative Models	Laws of Probability, Statistical Distributions, Moments
9	Bayesian Learning Methods	Bayes Rule
10	Bayesian Classification	
11	Bayesian Linear Regression	
Fairness and Transparency Issues		
12	Fairness in Machine Learning (PA3 Review)	
12	Interpretable Models (Decision Trees)	
Unsupervised Learning		
13	Clustering (k-Means/Spectral)	Linear Algebra (Eigenvalue Decomposition)
14	Dimensionality Reduction Methods (Principal Component Analysis)	

1.7 Course Deliverables

Deliverable	Release Date	Due Date
Gradiance 0	Feb 1	*Feb 9*
Gradiance 1	Feb 10	Feb 16
PA 1	Feb 8	Mar 5
Gradiance 2	Feb 17	Feb 23
Gradiance 3	Feb 24	Mar 2
PA 2	Mar 8	Apr 9
Gradiance 4	Mar 3	Mar 9
Gradiance 5	Mar 10	Mar 16
Gradiance 6	Mar 17	Mar 23
Gradiance 7	Mar 24	Mar 31
PA 3	Apr 12	May 7
Gradiance 8	Apr 1	Apr 7
Gradiance 9	Apr 8	Apr 14
Gradiance 10	Apr 15	Apr 21
Gradiance 11	Apr 22	Apr 28
Gradiance 12	Apr 29	May 5

Note:

- Gradiance quizzes
  - Will be released every Wednesday at 9.00 AM EST
  - Due next Tuesday at 11.59 PM EST
  - Gradiance 0 will not be evaluated (warm up)
- All assignments are electronically due on Fridays by 11.59 PM EST through UBLearns.

1.8 Assignments (Tentative Schedule)

- Programming Assignment 1 - This assignment will focus on building linear models for supervised learning. This will include implementing a linear regression model for regression, and three classification models, viz., logistic regression, perceptron, and support vector machine (SVM).
- Programming Assignment 2 - In this assignment, your task will be to explore non-linear machine learning models to learn from text and image data.
- Programming Assignment 3 - This programming assignment has two parts. In the first part, you will implement a Naive Bayes Classifier and test it on a publicly available data set. In the second part, you will manipulate the data characteristics to understand how classifiers get impacted by the underlying bias in the training data. Focus will be on developing a COMPAS style risk assessment system.

1.9 Course Texts

- Kevin Murphy, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.
- Tom Mitchell, Machine Learning. McGraw-Hill, 1997.

- Chris Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- David Mackay, Information Theory, Inference, and Learning Algorithms, Cambridge Press, 2003.
- Trevor Hastie, Robert Tibshirani and Jerome Friedman, The Elements of Statistical Learning. Springer, 2009.
- Richard S. Sutton and Andrew G. Bart, Reinforcement Learning: An Introduction. MIT Press, 2015.

1.10 Grading

- Short weekly quizzes using Gradiance (12) – 20%
- Programming Assignments (3) – 45%
- Mid-term Exam (virtual-UBLearns, open book/notes) – 15%
- Final Exam (virtual-UBLearns, open book/notes) – 20%
- Final grade (*Tentative*)

- A [92.5,100]
- A- [87.5,92.5)
- B+ [82.5,87.5)
- B [77.5,82.5)
- B- [72.5,77.5)
- C+ [67.5,72.5)
- C [62.5,67.5)
- C- [57.5,62.5)

1.11 Exams

- **Mid-term Exam** March 19, 1.50 PM - 2.40 PM, virtually using UBLearns
- **Final Exam** May 14, 11.45 AM - 2.45 PM, virtually using UBLearns

Note: The mid-term will held during the regular Friday lecture.

1.12 Expectations

- Students are expected to act in a professional manner during the virtual classes and office hours.
- Programming assignments will be graded and returned to students.
- Late submission of assignments will receive a grade of zero.
- No late submission of Gradiance quizzes are allowed. The quizzes will automatically become unavailable immediately after the due date and no accomodations will be made for missed quizzes.
- Students are encouraged to discuss assignments and share ideas, but each student must independently write and submit their own solution.
- Makeup exams will be given in the following circumstances only:
  1. You contact the instructor prior to the exam
  2. You have a valid and documented reason to miss the exam

1.13 Accessibility Services and Special Needs

If you have a disability and may require some type of instructional and/or examination accommodation, please inform me early in the semester so that we can coordinate the accommodations you may need. If you have not already done so, please contact the Office of Accessibility Services (formerly the Office of Disability Services) University at Buffalo, 25 Capen Hall, Buffalo, NY 14260-1632; email: stu-accessibility@buffalo.edu Phone: 716-645-2608 (voice); 716-645-2616 (TTY); Fax: 716-645-3116; and on the web at http://www.buffalo.edu/accessibility/. All information and documentation is confidential. The University at Buffalo and the School of Engineering and Applied Sciences are committed to ensuring equal opportunity for persons with special needs to participate in and benefit from all of its programs, services and activities.

1.14 Academic Integrity

This course will operate with a zero-tolerance policy regarding cheating and other forms of academic dishonesty. Any act of academic dishonesty will subject the student to penalty, including the high probability of failure of the course (i.e., assignment of a grade of "F"). It is expected that you will behave in an honorable and respectful way as you learn and share ideas. Therefore, recycled papers, work submitted to other courses, and major assistance in preparation of assignments without identifying and acknowledging such assistance are not acceptable. All work for this course must be original for this course. Additionally, you are not allowed to post course homeworks, exams, solutions, etc., on a public forum. Please be familiar with the University and the School policies regarding plagiarism. Read the Academic Integrity Policy and Procedure for more information: http://undergrad-catalog.buffalo.edu/policies/course/integrity.shtml. Visit the Senior Vice Provost for Academic Affairs web page for the latest information at http://vpue.buffalo.edu/policies/

Machine Learning Honor Code

Against the ML honor code to:

1. Collaborate on Gradiance quizzes
2. Collaborate or cheat during exams
3. Submit someone else's work, including from the internet, as one's own for any submission
4. Misuse Piazza forum

You are allowed to:

1. Have discussions about homeworks. Every student should submit own homework with names of students in the discussion group explicitly mentioned.

Warning:

- Violation of ML honor code and departmental policy will result in an automatic F for the concerned submission
- Two violations fail grade in the course

CHAPTER 2

Lecture Videos

Links will be added here as the course progresses.

Warning: You will need to authenticate using your UBIT account.

Week	Monday	Wednesday	Friday
Week 1	Feb 01	Feb 03	Feb 05
Week 2	Feb 08	Feb 10	Feb 12
Week 3	Feb 15	Feb 17	Feb 19
Week 4	Feb 22	Feb 24	Feb 26
Week 5	Mar 01	Mar 03	Mar 05
Week 6	Mar 08	Mar 10	Mar 12
Week 7	Mar 15	Mar 17	Mar 19
Week 8	Mar 22	Mar 24	Mar 26
Week 9	Mar 29	Mar 31	Apr 02
Week 10	Apr 05	Apr 07	Apr 09
Week 11	Apr 12	Apr 14	Apr 16
Week 12	Apr 19	Apr 21	Apr 23
Week 13	Apr 26	Apr 28	Apr 30
Week 14	May 03	May 05	May 07

CHAPTER 3

Documents

Note: The *handouts* have all the content that the *slides* have, along with some additional discussion which is not on the slides. If you want to save these for future use or for printing, please use the handouts and **not** the slides.

Topic	Documents
ML Basics	`slides` `handouts` `scans`
Supervised Learning::Linear Models	
Linear Regression	`slides` `handouts` `scans`
Logistic Regression/Percepton	`slides` `handouts` `scans`
Support Vector Machines	`slides` `handouts` `scans`
Kernel Methods	
Kernel Regression	`slides` `handouts` `scans`
Kernel Support Vector Machines	`slides` `handouts` `scans`
Supervised Learning::Non-linear Models	
Non-linear Regression and Regularization	`slides` `handouts` `scans`
Neural Networks	`slides` `handouts` `scans`
Statistical Learning	
Generative Models	`slides` `handouts` `scans`
Bayesian Learning	`slides` `handouts` `scans`
Bayesian Classification	`slides` `handouts` `scans`
Bayesian Linear Regression	`slides` `handouts` `scans`
Fairness in Machine Learning	
Fairness aspects in Machine Learning	`slides` `handouts` `scans`
Fairness primer	`fairness primer`
Decision Trees	`slides` `handouts` `scans`
Unsupervised Learning	
Clustering (k-Means/Spectral Methods)	`slides` `handouts` `scans`
Principal Component Analysis	`slides` `handouts` `scans`

CHAPTER 4

Frequently Asked Questions

4.1 Logistics

4.1.1 Where is the course website?

The course website is here. The website has links to the resources for the course.

4.1.2 Where is the course syllabus?

Syllabus is here.

4.1.3 When does the course begin?

First class is on February 1st, 2021 (Monday), virtually on Zoom.

4.1.4 What are the other resources I should be aware of?

The course website has links to all the resources that we will be using during the course. We will be using **Piazza** for all communications and announcements and **Gradiance** for the weekly quizzes. The UBLearns system will be used to manage assignment submissions and grades. The slides and handouts (for printing) are available here. We have also prepared a glossary of terms that you can refer to.

We will be using **Python** as the primary programming language. We will use Jupyter Notebooks for in-class demonstrations. The notebooks are available here. You can also check out the notebooks to your personal computers from Github.

4.1.5 What do I need to do before class starts?

- Sign-up for Piazza (if you do not have an account aready) and enroll into the CSE 474/574 class.

- Confirm that the class shows up in your UBLearns account.
- Create an account on Gradiance.

Note: Make sure that you use your UBIT name as the user id. Gradiance requires the user name to be at least 6 characters. If your UBIT name has less than 6 characters, pad it with the required number of x's to make it 6 character long. For example, if your UBIT name is **jliu**, your Gradiance username should be **jliuxx**. Please use lower-case username only

Improper usernames could result in incorrect assignment of grades.

- Add the CSE 474/574 class on Gradiance using this token **14C9F9A9**. Will be available on Feb 1st at 9.00 AM.

4.1.6 How do I contact the instructors?

Best way is through Piazza. You may send a private or a public message, depending on the nature of the query. Direct email to instructors might not get handled in time.

4.2 Content

4.2.1 What will I learn in this course?

This course is aimed at providing an introduction to the field of Machine Learning. We will focus on understanding the breadth of topics within this field. ML is a broad field and due to the limited time, we will not be able to explore the topics to great depths. However, we will also go deep into some core concepts that are relevant to a many sub-topics within ML. We will explore three ML tasks, viz., supervised, unsupervised and reinforcement (*new*) learning. At the same time, we will also explore cross-cutting issues such as ethics and fairness associated with ML algorithms.

4.2.2 What will I not learn in this course?

Due to the limited time, we will not be able to cover many ML algorithms that you might have heard of, e.g., decision trees, convolutional neural networks (though we will look at simpler neural networks). At the same time, this course is not about teaching how to use a certain library or a tool. You will learn how to build such tools from scratch.

4.2.3 How important is Math in this course?

Very important. In fact, this course is one of the most math intensive course in the undergraduate CSE program. By the second class (Wednesday), we will be deriving a solution for the minimization of an objective function of a vector variable. Similarily, I will assume a strong background in Probability and Statistics.

4.2.4 What linear algebra topics do I need to know before the class starts?

Some of these topics will be covered in-class and during the recitation sections, but it would be good to brush up on these before the class starts. Also, refer to glossary for the notation that we will follow in the class:

- Matrix-vector product, matrix-matrix product, vector-vector dot product, vector-vector cross product
- Special matrices: Identity, diagonal, symmetric, matrix transpose
- Matrix operations: trace, norms (of vectors and matrices), linear independence and rank of a matrix, orthogonal matrices, matrix determinant, quadratic form, eigendecomposition of a matrix
- Matrix calculus: Gradient of a function, Hessian, Gradients and Hessians of a quadratic function

4.3 Attendance

4.3.1 Is in-class attendance mandatory?

While the class videos will be posted (almost) immediately after the class, I would strongly recommend against relying on the videos to understand the material.

4.3.2 Will class videos be posted?

Yes, they will be available within the class UBLearns website. Look out for a **Panopto Recordings** link on that page. Videos are typically available within 2 hours of the lecture.

4.4 Gradiance

4.4.1 What is Gradiance?

Gradiance is an online quiz taking system. We will be using that to administer weekly quizzes. Check the syllabus for the exact schedule.

4.4.2 How do I enroll?

- Sign-up here.

> **Warning:** Make sure you use your UBIT name as your username. Any other accounts will be deleted!

- After creating your account, add the class using token **14C9F9A9**.

4.4.3 How does it work?

- Every week one quiz will be available on Wednesday morning at 9.00 AM and will be due the next Tuesday at 11.59 PM.
- Each quiz will contain 4-5 problems on topics covered the previous week.
- Each problem will have multiple choices, with only one correct answer.
- At the end of a submission, the system will give you hints for problems that you answer incorrectly.
- There will be a 5 minute between successive tries.
- Maximum 3 tries are allowed.

- Every wrong answer will result in one negative point.
- A practice quiz will be posted in the first week.

4.5 Piazza

4.5.1 What do I need Piazza for?

We will use Piazza as our primary medium of communication. Students with questions can post on Piazza (either private or public). Additionally, it will be used as a discussion forum to have discussions among students and instructors regarding various course aspects.

4.5.2 Why should I be active on Piazza?

It is well-documented that student led discussions on Piazza result in much better learning outcomes compared to a single-direction discourse.

Note: The top contributor (questions or answers) on Piazza will get "recognized".

4.6 Assignments

4.6.1 What will the assignments entail

The course consists of three programming assignments. Assignments will be done in groups of 3.

4.6.2 What programming language will be used?

We will be using **Python** as the programming language for the assignments. We will be using the Python 3.x version. Note that if you are using Python 2.x, you might run into issues.

4.6.3 What if I do not know Python?

We will have a couple of recitation sessions to introduce you to Python. However, I will strongly advise you to checkout resources on the web to get started on learning Python, something that will help you in future too.

- Installing python, ipython
- Python IDE - Canopy
- More about ipython notebooks
- Python for Developers, a complete book on Python programming by Ricardo Duarte
- CodeAmerica - Python
- An introduction to machine learning with Python and scikit-learn (repo and overview) by Hannes Schulz and Andreas Mueller

4.6.4 Are there any computing resources available?

While the programming assignments can be completed on a reasonably modern laptop or desktop, you can utilize the CSE resources (See here for more information).

CHAPTER 5

Glossary

Term (s)	Description	Notation
Data object (point, observation, sample, example)	A unit of analysis. Typically, a data object is represented as a **vector** of **features**.	Typically denoted as lower case letter, often bold, e.g., $\mathbf{x}$ or $\mathbf{x}_i$ or $\mathbf{x}^{(i)}$, where the subscript or superscript i denotes membership in a **data set**.
Data set	A collection of data objects.	Typically denoted as upper case letter, often bold, e.g., $\mathbf{X}$.
Vector	A list (or array) of real values ($\in \{-\infty,\infty\}$).	Typically denoted as bold lower case letter, e.g., $\mathbf{x}$. $\mathbf{x} \in \mathbb{R}^{\mathbf{d}}$, means that the vector represents a point in a d-dimensional vector space. An individual element of the vector is denoted as x_i.
Matrix	A 2-way array of real values ($\in \{-\infty,\infty\}$).	Typically denoted as bold upper case letter, e.g., $\mathbf{X}$. • $\mathbf{X} \in \mathbb{R}^{\mathbf{m}\times\mathbf{n}}$, means that the matrix has m rows and n columns. A vector is represented as a $m \times 1$ matrix. • An individual element of the matrix is denoted as X_{ij}.
Transpose	$\mathbf{X}^\top$, $\mathbf{x}^\top$	A transpose of a matrix is an operator which flips a matrix over its diagonal, i.e., $X_{ij} = X^\top_{ji}$. Transpose of a vector is a $1 \times d$ matrix.
Matrix multiplication	$\mathbf{X} = \mathbf{Y}\mathbf{Z}$	Only valid if the number of columns in $\mathbf{Y}$ is equal to the number of rows in $\mathbf{Z}$. • The ij^{th} entry of the matrix $\mathbf{Z}$ is the dot product (see below) between the i^{th} row of $\mathbf{X}$ and j^{th} column of $\mathbf{Y}$, i.e., $Z_{ij} = X^\top_{i*}Y_{*j}$ Where X_{i*} denotes the i^{th} row of $\mathbf{X}$ and Y_{*j} is the j^{th} column of $\mathbf{Y}$
Vector dot (inner) product	$\mathbf{x}.\mathbf{x} = \sum_{i=1}^{d} x_i^2$	In matrix notation, the dot product is expressed as $\mathbf{x}^\top\mathbf{x}$
Data Matrix	A $n \times d$ matrix, $\mathbf{X}$	If each data object in a data set can be represented as a vector in $\mathbb{R}^d$, the data set of n such objects is typically arranged as a $n \times d$ matrix, $\mathbf{X}$, where the transpose of each row of the matrix corresponds to a data object, i.e., $\mathbf{x}_i = X^\top_{i*}$.
Random Variable	A variable whose possible values are outcomes of a random phenomena (distribution)	Typically denoted as an upper case letter, X (bold - $\mathbf{X}$, if multivariate)
Probability	A measure of the	$P(A)$ denotes the probability of an event A to occur.

CHAPTER 6

Other Links

- Piazza
- UBLearns
- iPython Notebooks
- Gradiance
- Videos from previous years:
  - Panopto Videos (2020) *For University at Buffalo Students only*
  - YouTube Videos (2017)