



CS6140

# MACHINE LEARNING

Predrag Radivojac  
KHOURY COLLEGE OF COMPUTER SCIENCES  
NORTHEASTERN UNIVERSITY

Spring 2021

# ZOOM CALLS

- We have 34 students in class, as of January 18
- I will keep my video on
- Please keep your video on if you can
- Unmute yourself if you'd like to speak (and turn the video on)
  - or Raise hand on Zoom
  - or Use Zoom's chat option
- Classes are recorded
  - Password will be emailed to you after class
  - Office hours are not recorded

# BASIC INFORMATION

## Class meets:

Time: TF 3:25pm – 5:05pm

Place: Online

## Instructor:

Predrag Radivojac

Office: WVH 310D

Email: [predrag@northeastern.edu](mailto:predrag@northeastern.edu)

Web: <https://www.ccs.neu.edu/home/radivojac/>

## Office Hours:

Time: TF 5:30pm-7:00pm, or by appointment

Place: Online

## Class Web Site:

<https://www.ccs.neu.edu/home/radivojac/classes/2021springcs6140/>



# ABOUT MYSELF, BRIEFLY



# TEACHING ASSISTANTS

## **Clara De Paolis Kaluza**

Email: depaoliskaluza.m

Office hours: Thursdays 2-3:30pm, Fridays 10:30am-12pm, online.

## **Vikram Shenoy**

Email: shenoy.vi

Office hours: Mondays and Wednesdays 11am-12:30pm, online.

# A LITTLE ABOUT MACHINE LEARNING

- Machine learning is concerned with establishing theories as well as developing, analyzing, and applying algorithms that make useful inferences in the real world
- “Learn” functions and rules from data
- Specific problems always in mind, but frameworks are very important
- Balance between theory and application, slanted towards theory
- Probability theory, statistics, computer science
  - artificial intelligence
  - engineering
  - optimization
  - psychology
  - biology

# TIME

## How High Is Your XQ?

Your next job might depend on it

BY ELIZA GRAY

Is it true to say you have never hated anyone? Do you understand why stars twinkle? Have you used a display of emotion to get what you want? Would you rather read or watch TV? Do you usually notice when you are boring people? Do you hate opera singing? Would you consider yourself to be an ordinary person? Are you shy? Do you prefer problems that require a lot of thought? Do you enjoy giving parties? When you frequently rebellious? Do you believe people get stressed when they try to make you feel happy? Frequently stressed at work? Do you think sometimes accepting help from others makes you uncomfortable? Do you like to have someone around at work? Do you like to change your job? Do you make new friends all the time? Do you pretend to know more than you do? Do you need to form friendships at work? Would your colleagues say you are very confident? How much does

Do you often fantasize about being famous?

Do you find yourself getting angry easily?

Would you like to be an art collector?

Do people say you are eccentric?

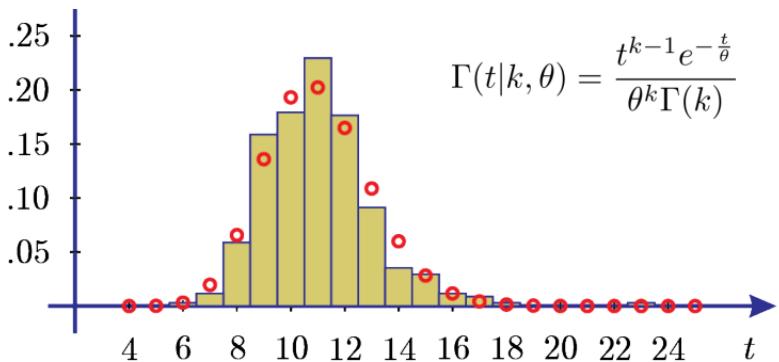
# AN EXAMPLE FROM REDDIT

*Let Artificial Intelligence guess your  
attractiveness and age*

#howhot



## A LITTLE MORE DETAIL...



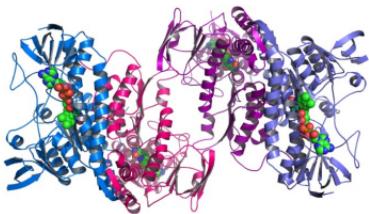
LILAC PETAL					12.16
SHRUB-4	9	-	-10	10.8	
S	10	14	11.9	12.10	
B	11	12.9	11.5	14.11	
S	9	11	13	12.11	
SHRUB-3	11	13	12.10	10.10	
H	13	10.11	9?	12.?	
	9	12.10	11	11.12	
	10	12.23	12.12	10.9	
	11	12	12.10	12.8	
	8	14	8.9	11.1	
	-	11	-9	5.8	
	-	12.10	12.15	8.6	

12	15	11	17	12	15	13
13	10	10	12	11	12	
11	14	14	14	13	18	16
12	9	11	—	8	9	
11	14	13	10	11	9	13
?	15	12	10	11	12	13(1)
3	9	13	10	11	—	
11	13	12	11	11	12	29
9	10	10	12	10	11	11(1)
13	13	11	8	—	12	17
13	13	11	8	15	—	15(1)
10	—	10	9	15	13	15(1)
?	13	13	—	12	10	11(1) 19

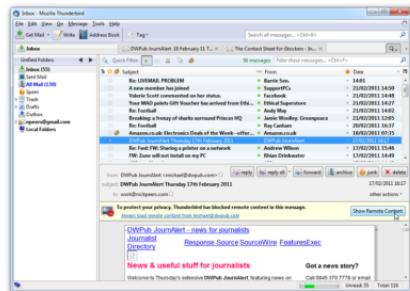
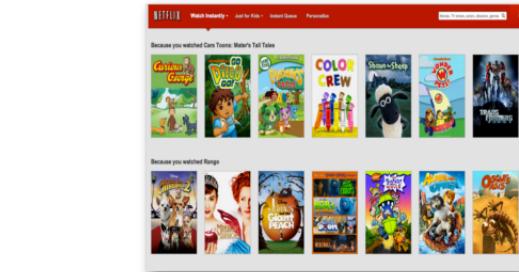
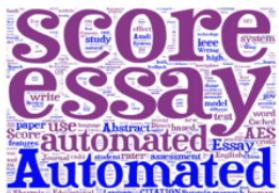
Barcode: CL138  
Title: SUMMER NIGHT  
Author: DEEBS, COLOR  
Call Number: 139-149  
Notes:  
1. Due date: 12/12  
2. Date issued: 11/13  
3. Date returned: 11/13  
4. Status: 37  
5. Shelf location: 1010  
6. Date due: 12/12  
7. Date issued: 11/11  
8. Date returned: 11/14  
9. Status: 12/12  
10. Shelf location: 1010  
11. Date due: 12/12  
12. Date issued: 11/13  
13. Date returned: 11/13  
14. Status: 11/9  
15. Shelf location: 1010  
16. Date due: 12/13  
17. Date issued: 11/13  
18. Date returned: 11/13  
19. Status: 11/13

Picture of the car from the Internet.

# WHERE ELSE DO WE SEE IT?



What Can Be Automated?  
What Cannot Be Automated?



Pictures from the Internet.

# So...

- In real situations we have uncertainty
  - We have incomplete knowledge of the environment
  - Actions of other actors are not provided
- Applied everywhere to learn from data and make predictions, some of which facilitate decisions
- Utility theory: incorporates an agent's preferences towards certain scenarios
- Decision theory: probability theory + utility theory
- Rational decision: decision that maximizes expected utility

# BRIEF OVERVIEW OF CS6140

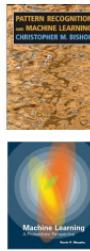
See online syllabus...

- mathematical foundations of machine learning
- overview of machine learning
- foundations of parameter estimation
- basic unsupervised learning
- classification (prediction of discrete outputs)
- regression (prediction of continuous outputs)
- kernel methods (within classification/regression)
- ensemble methods
- practical aspects in machine learning
- special topics (if time permits)

# TEXTBOOK INFORMATION

## Main books:

- Pattern Recognition and Machine Learning - by C. M. Bishop, Springer 2006.
- Machine Learning: A Probabilistic Perspective - by K. P. Murphy, The MIT Press, 2012



## Recommended readings:

- The Elements of Statistical Learning - by T. Hastie et al., Springer, 2009
- Machine Learning - by Tom M. Mitchell, McGraw-Hill, 1997

Supplementary material will be provided in class!



# WHAT DO I EXPECT AND ASSUME?

- Basic mathematical skills
  - calculus
  - probabilities
  - linear algebra
- You are patient and hardworking
- Your integrity is impeccable
- You are motivated to learn (machine learning)
- You are motivated to succeed in class

# GRADING

- Midterm exam: 20%
  - Final exam: 20%
  - Homework assignments (4): 30%
  - Mini project: 25%
  - Class participation: 5%
- 

- I decide on the final grade (I don't necessarily enjoy this)

## GRADING

- Top performers in the class will get As
- Distributions of scores will be shown (I hope regularly)
- If you don't know where you stand in class, ask me
- All assignments count, must be typed to show formulas properly! Plan ahead!
- All assignments are individual!
- All the sources used for problem solution must be acknowledged (people, web sites, books, etc.)

# ONE OF PREVIOUS YEARS AFTER MIDTERM

$$T = 30 \cdot \frac{1}{140} \cdot \sum_{i=1}^1 hw(i) + 20 \cdot \frac{m}{100}$$

No. Students = 49

Mean = 33.9 out of 50

Standard Deviation = 9.1

Highest: 48.4

Lowest: 9.1

## Percentiles:

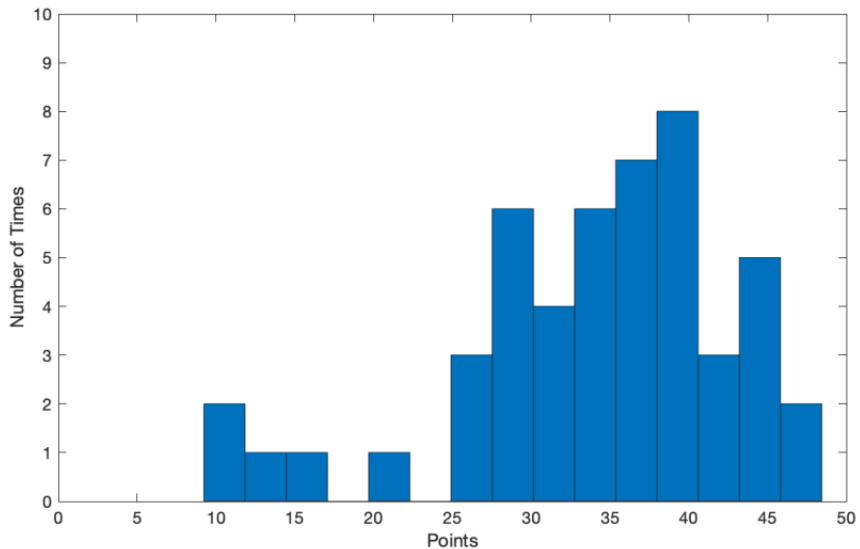
10%: 22.1 points

25%: 29.6 points

50%: 35.6 points

75%: 39.3 points

90%: 44.9 points



# ONE OF PREVIOUS YEARS AFTER MIDTERM

$$T = 30 \cdot \frac{1}{n} \cdot \sum_{i=1}^4 hw(i) + 20 \cdot \frac{m}{100} + 25 \cdot \frac{mp}{100} + 20 \cdot \frac{f}{100} + 5 \cdot \frac{p}{10}$$

No. Students = 49

Mean = 33.9 out of 50

Standard Deviation = 9.1

Highest: 48.4

Lowest: 9.1

## Percentiles:

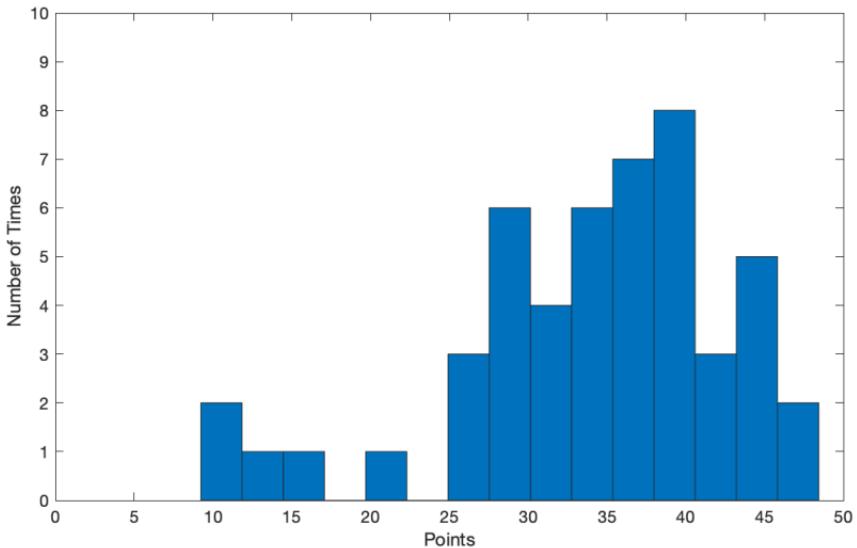
10%: 22.1 points

25%: 29.6 points

50%: 35.6 points

75%: 39.3 points

90%: 44.9 points



# TYPING ASSIGNMENTS

- Latex (TeXShop + MacTeX or TeXnicCenter + MiKTeX)
- Word

Overleaf

The screenshot shows the Overleaf documentation website. At the top, there's a navigation bar with links for 'Features & Benefits', 'Templates', 'Help', 'Projects', and 'Account'. Below the navigation bar, there's a search bar labeled 'Search help library...'. The main content area is titled 'Documentation' and contains a welcome message: 'Welcome to the Overleaf knowledge base. A complete list of topics is provided on the left hand-side, but here is a selection of useful articles:'. Under this, there's a section titled 'New to LaTeX?' with a sub-section 'Start with our [Learn LaTeX in 30 minutes](#) guide.' Further down, there's a list of topics: 'Create your first document in L<sup>A</sup>T<sub>E</sub>X', 'Paragraphs and new lines', 'Bold, italics and underlining', 'Lists', 'Mathematics', 'Bibliographies and references', 'Images', and 'Tables'. On the left side of the main content area, there's a sidebar with a list of 'Overleaf guides' including 'Creating a document in Overleaf', 'Uploading a project', 'Copying a project', 'Creating a project from a template', 'Using the Overleaf project menu', 'Including images in Overleaf', 'Exporting your work from Overleaf', 'Working offline in Overleaf', 'Using Track Changes in Overleaf', 'Using bibliographies in Overleaf', 'Sharing your work with others', and 'Using the History feature'.

Lyx

The screenshot shows the Lyx application interface. The title bar indicates the file is 'D:\Lyx\Get-Master\lib\doc\UserGuide.lyx'. The main window shows a LaTeX code snippet:  $\sum_{n=0}^{\infty} \frac{1}{n!} = e$ . To the right of the code, there's a detailed explanation: 'Sum ( $\sum$ ) and integral ( $\int$ ) operators are very often decorated with limits. These limits can be entered in LyX by entering them as you would enter a super- or subscript, directly after the symbol. The sum operator will automatically place its "limits" over and under the symbol in displayed formulas, and to the side in inline formulas, as in  $\sum_{n=0}^{\infty} \frac{1}{n!} = e$ , versus  $\sum_{n=0}^{\infty} \frac{1}{n!}$ '. Below this, there's another LaTeX code snippet:  $\lim_{x \rightarrow \infty} f(x)$ , with a note: 'Integral signs, however, will place the limits to the side in both formula types.' Further down, there's a note: 'All operators with limits will be automatically re-sized when placed in display mode. The placement of the limits can be changed by placing the cursor directly behind the operator and using the menu Edit > Math > Change Limits Type or entering Alt+M L.' At the bottom, there's a section titled '5.1.7 Math Symbols' with a link to 'Math | Symbols'.

# ROADMAP

January: 18  
25

March: 1  
8  
15  
22  
29

February: 1  
8  
15  
22

April: 5  
12  
19  
26

# ROADMAP

January: 18  
25 h1

March: 1  
8 M, PP (H3)  
15  
22 h4  
29

February: 1  
8 H1, h2  
15  
22 H2, pp (h3)

April: 5 H4  
12  
19 F  
26 P

# LATE ASSIGNMENT POLICY

- Homework assignments are due on the specified due date through Canvas
- Late assignments will be accepted\* according to the following rules

– points	(on time)	{}	recommended!
– points x 0.9	(1 day late)		
– points x 0.7	(2 days late)		
– points x 0.5	(3 days late)		
– points x 0.3	(4 days late)		
– points x 0.1	(5 days late)		
– 0	(after 5 days)		not recommended!

\* if there are legitimate circumstances to not apply this policy, please inform me early

# ACADEMIC HONESTY

- *The Code of Student Conduct*
  - <http://www.northeastern.edu/osccr/code-of-student-conduct/>
  - Interesting things there, including that...
  - “Students are expected to display proper respect for the rights and privileges of other members of the University community and their guests .”
  - “Furthermore, students must follow the reasonable directions of University personnel.”
  - “The Code of Student Conduct applies both on and off campus”
- Academic honesty taken seriously!
  - Rules I follow: problems with one assignment, 0 on that assignment; problems on another assignment, 0 for the course.

# MISCELLANEA

- Do not record instructor(s) without explicit permission
- Turn off cell phones and other similar devices during class
- Use laptops if you have to (unless it bothers someone)
- “will u be in ur office after class”; “I need a letter of recommendation.”
- BE NICE TO  
PEOPLE

