CS6140

MACHINE LEARNING

Predrag Radivojac
Khoury College of Computer Sciences
Northeastern University

Fall 2020
Class meets:
  Time: TF 3:25pm – 5:05pm
  Place: Richards Hall 236, online.

Instructor:
  Predrag Radivojac
  Office: WVH 310D
  Email: predrag@northeastern.edu
  Web: https://www.ccs.neu.edu/home/radivojac/

Office Hours:
  Time: T 5:30pm-7:00pm
  F 5:30pm-7:00pm or by appointment
  Place: online.

Class Web Site:
  https://www.ccs.neu.edu/home/radivojac/classes/2020fallcs6140/
TEACHING ASSISTANTS

Clara De Paolis Kaluza
Email: depaoliskaluza.m
Office hours: Mondays 3-4pm, Fridays 10-12pm, online.

Piyush Goel
Email: goel.pi
Office hours: Wednesdays 1-4pm, online.

Vikram Shenoy
Email: shenoy.vi
Office hours: Tuesdays 1:30-3 pm and Thursdays 10:30am-12pm, online.

INSTRUCTIONAL ASSISTANTS

Miles Wilson
Email: wilson.mil@northeastern.edu
Available in classroom for NUflex.

** Available for in-person meetings during class.
A LITTLE ABOUT ML

• Machine Learning is concerned with 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
Electing American Presidents (18th, 19th, 20th Century)

• if the Washington win their last home game before the election, the incumbent’s party will be re-elected

• no Republican has ever won a presidential election without carrying Ohio

• no incumbent with a four-letter last name has ever been re-elected (Polk, Taft, Ford, Bush Sr.)

• Americans won’t unseat a wartime President
ELECTING AMERICAN PRESIDENTS (18\textsuperscript{TH}, 19\textsuperscript{TH}, 20\textsuperscript{TH} CENTURY)

• if the Washington win their last home game before the election, the incumbent’s party will be re-elected (2012: Washington vs. Carolina: 13-21)

• no Republican has ever won a presidential election without carrying Ohio

• no incumbent with a four-letter last name has ever been re-elected (Polk, Taft, Ford, Bush Sr.) (2004: GWB)

• Americans won’t unseat a wartime President
A LITTLE ABOUT ML

• Machine Learning is concerned with 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
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 get stressed, do you believe people get stressed when they try to do something in your day that makes you feel happy? How do you feel when you uncomfortable accepting help from others? How? Do you ever stressed at work? Do you think someone around at work would like to change a lot of things about your work? Do you make new friends all the time? Do you wish you? Have you ever pretended to know someone very, very well? 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…

\[ \Gamma(t|k, \theta) = \frac{t^{k-1}e^{-\frac{t}{\theta}}}{\theta^k \Gamma(k)} \]
WHERE ELSE DO WE SEE IT…

What Can Be Automated?
What Cannot Be Automated?
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
OVERVIEW OF THE CS 6140 COURSE

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)
Main books:
• Pattern Recognition and Machine Learning - by C. M. Bishop, Springer 2006.


Recommended readings:
• The Elements of Statistical Learning - by T. Hastie et al., Springer, 2009


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 POLICY

• 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)
More on 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.)
## Roadmap

<table>
<thead>
<tr>
<th>Month</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>September</td>
<td>7, 14, 21, 28</td>
</tr>
<tr>
<td>October</td>
<td>5, 12, 19, 26</td>
</tr>
<tr>
<td>November</td>
<td>2, 9, 16, 23, 30</td>
</tr>
<tr>
<td>December</td>
<td>7, 14</td>
</tr>
</tbody>
</table>
**ROADMAP**

| September    | 7
|             | 14 h1
|             | 21
|             | 28 H1, h2
| November    | 2
|             | 9 h4
|             | 16
|             | 23 H4
|             | 30
| October     | 5
|             | 12 H2, pp (h3)
|             | 19
|             | 26 M, PP (H3)
| December    | 7 F
|             | 14 P
**Late Assignment Policy**

- Homework assignments are due on the specified due date through Blackboard.

- Late assignments will be accepted* according to the following rules:
  - points (on time)
  - 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)

* 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!
  – Reporting every incident to the university
**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