MACHINE LEARNING

CS6140

Predrag Radivojac
KHOURY COLLEGE OF COMPUTER SCIENCES
NORTHEASTERN UNIVERSITY

Fall 2024

Basic Information

Class meets:

Time: TF 3:25pm – 5:05pm
Place: Behrakis Health Sciences Center 310

Instructor:

Predrag Radivojac
Office: 908 @ 177 Huntington Ave.
Email: predrag@northeastern.edu
Web: https://www.ccs.neu.edu/home/radivojac/

Office Hours:

Time: TF 5:15pm-6:45pm, or by appointment
Place: Behrakis Health Sciences Center 310

Class Web Site:

https://www.ccs.neu.edu/home/radivojac/classes/2024fallcs6140/

About Myself, Briefly

Picture from the Internet.

Teaching Assistants

Ritwik Anand

Email: anand.r
Office hours: Wednesdays 3-4:30pm and Thursdays 10-11:30am, online.

Alfonso Barajas Cervantes

Email: barajascervantes.a
Office hours: Mondays and Wednesdays 9-10:30am, online.

Rashika Ramola

Email: ramola.r
Office hours: Thursdays and Fridays 10-11:30am, online.

Daniel Zeiberg

Email: zeiberg.d
Office hours: Tuesdays 8-9:30am and Thursdays 2:30-4pm, online.

JUNE 22, 2015

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 co[illegible]er yourself to be an ordinary person? Are you shy? Do you prefer problems t[illegible]require a lot of thought? Do you enjoy giving parties? When yo[illegible]frequently rebellious? Do you prefer to work with [illegible]o you believe people get stressed when they try to[illegible]mething in your day that makes you feel happy? F[illegible]u uncomfortable accepting help from ot[illegible]ever stressed at work? Do you think sometime[illegible]e? Do you like to have someone around at wo[illegible]uld you like to change a lot of things about yo[illegible]ss? Do you make new friends all the time? D[illegible]h you? Have you ever pretended to know m[illegible] form friendships at work? Would your colleagu[illegible]ou are ver[illegible]confident? How much does

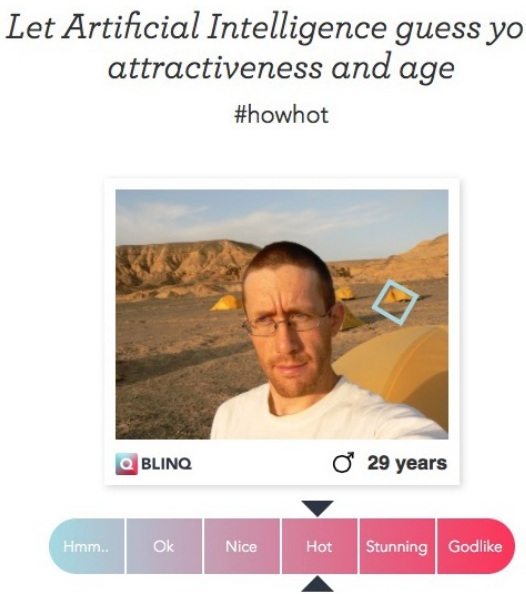
Do you often fantasize about being famous?

Do you find yourself getting angry easily?

Would you like to be an art collector?

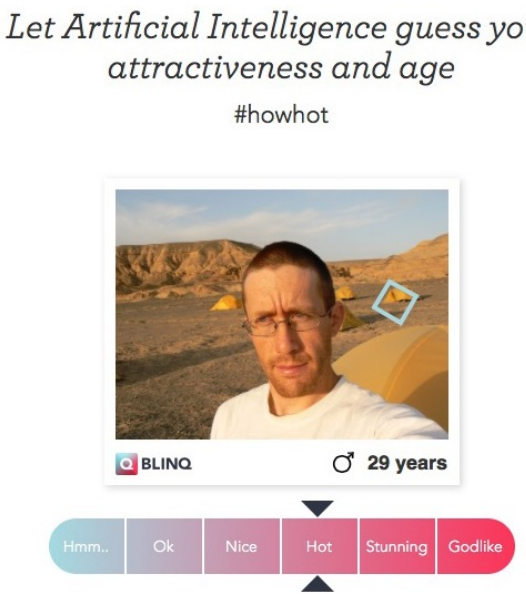
Do people say you are eccentric?

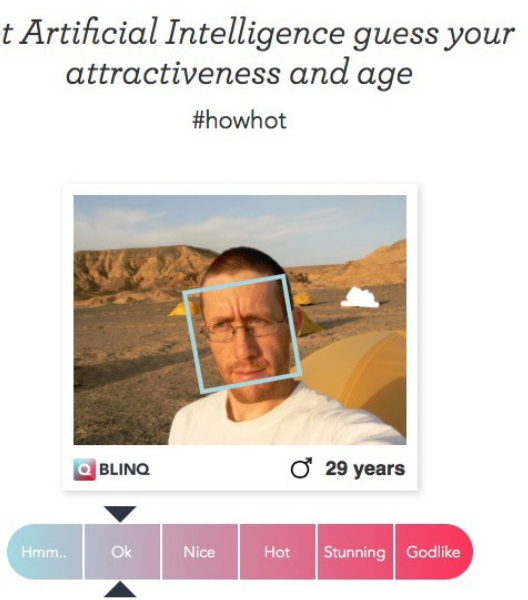
An Example from Reddit



Pictures from the Internet.

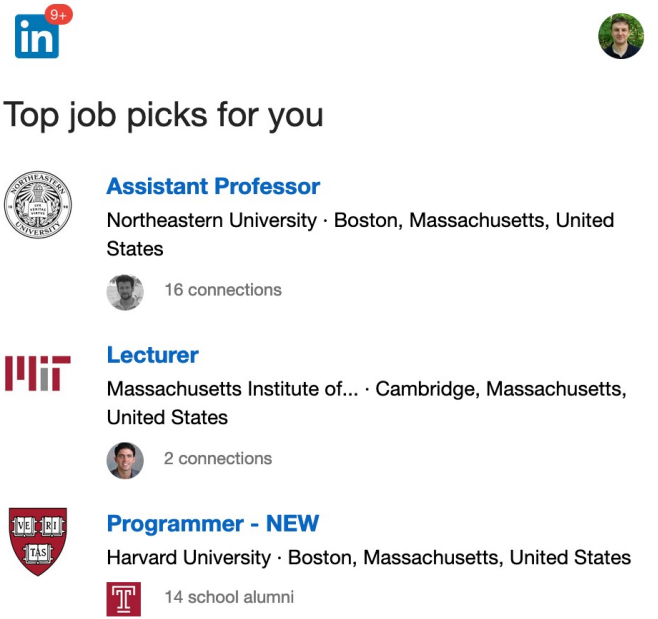
An Example from Reddit





Pictures from the Internet.

An Example from LinkedIn



Pictures from the Internet.

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

AN EXAMPLE FROM A PAPER

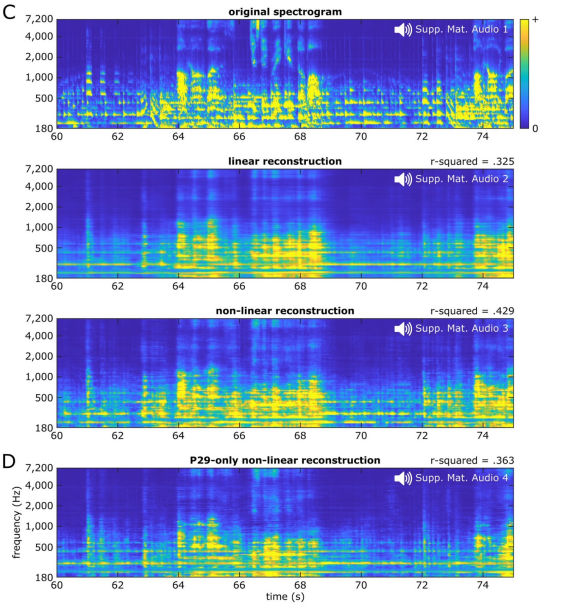
Music can be reconstructed from human auditory cortex activity using nonlinear decoding models

Ludovic Bellier, Anaïs Llorens, Déborah Marciano, Aysegul Gunduz, Gerwin Schalk, Peter Brunner, Robert T. Knight

Published: August 15, 2023 • https://doi.org/10.1371/journal.pbio.3002176

Abstract

Music is core to human experience, yet the precise neural dynamics underlying music perception remain unknown. We analyzed a unique intracranial electroencephalography (iEEG) dataset of 29 patients who listened to a *Pink Floyd* song and applied a stimulus reconstruction approach previously used in the speech domain. We successfully reconstructed a recognizable song from direct neural recordings and quantified the impact of different factors on decoding accuracy. Combining encoding and decoding analyses, we found a right-hemisphere dominance for music perception with a primary role of the superior temporal gyrus (STG), evidenced a new STG subregion tuned to musical rhythm, and defined an anterior–posterior STG organization exhibiting sustained and onset responses to musical elements. Our findings show the feasibility of applying predictive modeling on short datasets acquired in single patients, paving the way for adding musical elements to brain–computer interface (BCI) applications.



A Large Language Model Example

What topics should we cover in CS6140 at Northeastern University?

To design a curriculum for CS6140 at Northeastern University, particularly if it is a Machine Learning course, consider covering the following key topics:

1. Introduction to Machine Learning

- Overview of Machine Learning: Definitions, Examples, and Types (Supervised, Unsupervised, and Reinforcement Learning)
- Key Concepts: Datasets, Features, Labels, Models, Training, and Testing

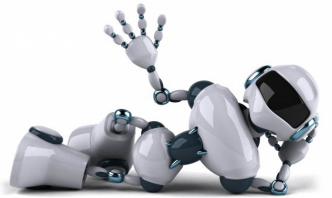
2. Mathematical and Statistical Foundations

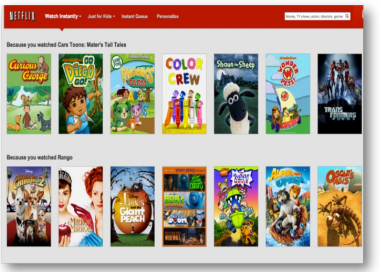
- Linear Algebra: Vectors, Matrices, Eigenvalues, Singular Value Decomposition
- Probability and Statistics: Probability Distributions, Bayesian Inference, Hypothesis Testing
- Optimization Techniques: Gradient Descent, Stochastic Gradient Descent, Convex Optimization

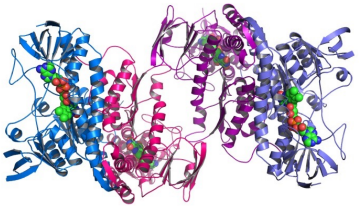
3. Supervised Learning Techniques

- **Regression:**

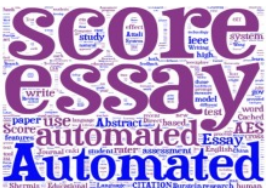
Where Else Do We See It?

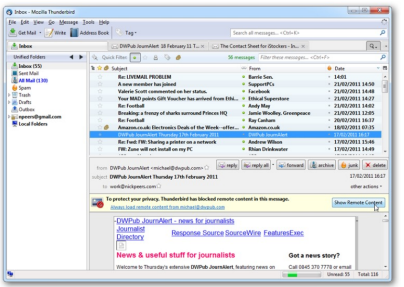






What Can Be Automated?
What Cannot Be Automated?





Pictures from the Internet.

The Field of 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 observations (data)
- Specific problems always in mind, but frameworks are very important
- Balance between theory and application, slanted towards theory
- Deals with the uncertainty
  - incomplete knowledge of the world
  - actions of actors are unknown
  - interest in making rational decisions (those that maximize utility)
- Probability theory, statistics, computer science
  - artificial intelligence
  - engineering
  - optimization
  - psychology
  - biology

THE ART OF CONJECTURING

Jacob Bernoulli

Probability

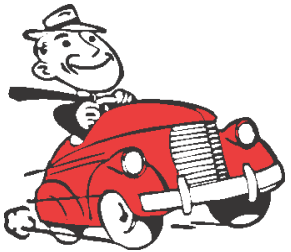
“Pobability, [...] is the degree of certainty, and it differs from the latter as a part differs from the whole”

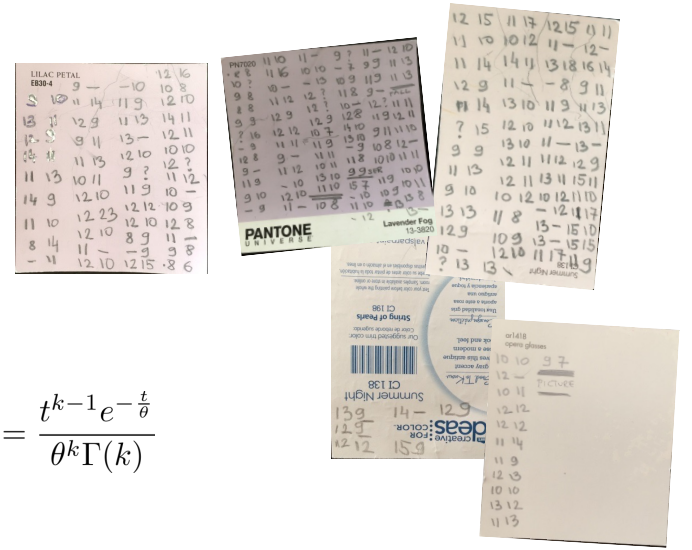
The Art of Conjecturing

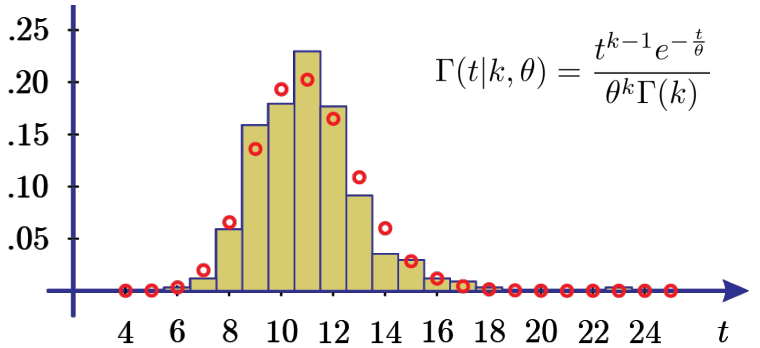
“To make a conjecture [prediction] about something is the same as to measure its probability. Therefore, we define the art of conjecturing [science of prediction] or stochastics, as the art of measuring probabilities of things as accurately as possible, to the end that, in judgements and actions, we may always choose or follow that which has been found to be better, more satisfactory, safer, or more carefully considered.”

Jacob Bernoulli, The Art of Conjecturing, 1713.

A Little More Detail…







$$\Gamma(t|k,\theta) = \frac{t^{k-1}e^{-\frac{t}{\theta}}}{\theta^k \Gamma(k)}$$

Picture of the car from the Internet.

BRIEF OVERVIEW OF CS6140

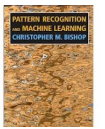
See online syllabus…

- mathematical foundations of machine learning
- overview of machine learning
- foundations of parameter estimation
- theory of supervised 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)

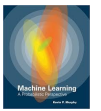
TEXTBOOK INFORMATION

Main books:

- Pattern Recognition and Machine Learning – by C. M. Bishop, Springer 2006.

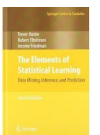


- Machine Learning: A Probabilistic Perspective – by K. P. Murphy, The MIT Press, 2012



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

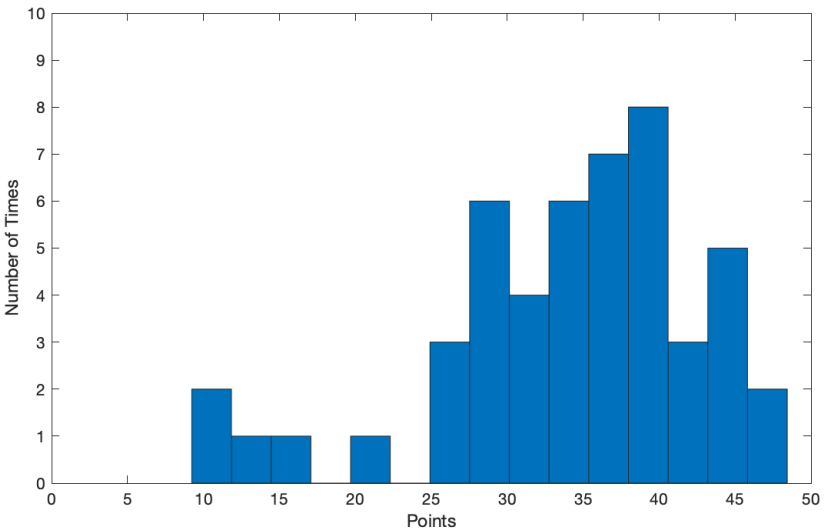
- Top performers in the class will get As
- Distributions of scores will be shown (I hope regularly)
- If you don't know where you stand in class, ask me
- All assignments count, must be typed to show formulas properly! Plan ahead!
- All assignments are individual!
- All the sources used for problem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

One of Previous Years After Midterm

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

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



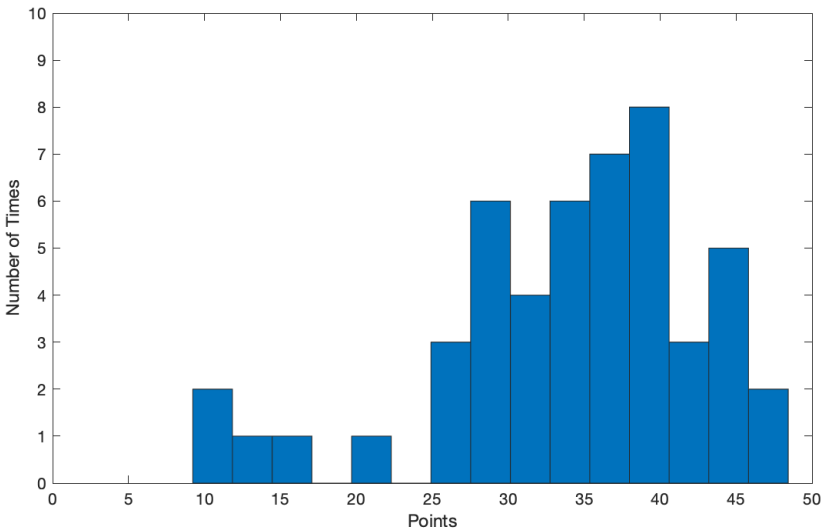
One of Previous Years After Midterm

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

← total potential score

No. Students = 49
Mean = 33.9 out of 50
Standard Deviation = 9.1
Highest: 48.4
Lowest: 9.1

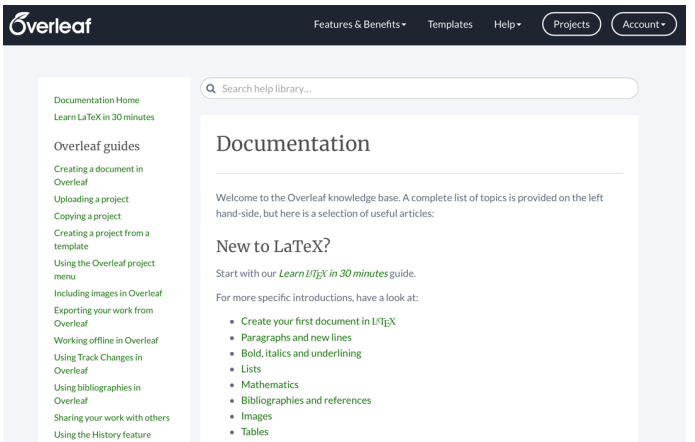
Percentiles:
10%: 22.1 points
25%: 29.6 points
50%: 35.6 points
75%: 39.3 points
90%: 44.9 points



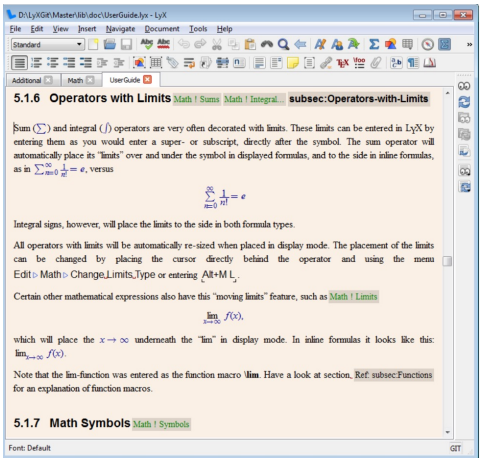
TYPING ASSIGNMENTS

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

Overleaf



Lyx



PLAN

September:	2	November:	4
	9		11
	16		18
	23		25
	30		
October:	7	December:	2
	14		9
	21		
	28		

PLAN

September:	2	November:	4 h4
	9 h1		11
	16		18 H4
	23 H1, h2		25
	30		
October:	7 H2, pp (h3)	December:	2 F
	14		9 P
	21 M, PP (H3)		
	28		

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.
  - If there is a problem with assignment 4, I will go back and check assignments 1-3

MISCELLANEA

- Do not record instructor(s) without explicit written 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

Pictures from the Internet.