



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



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.

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 something in your day that makes you feel happy? Do you feel uncomfortable accepting help from others? Do you think sometimes someone around at work a lot of things about you? Do you make new friends all the time? Do you pretend to know more than you do? Would your colleagues be confident? How much does

The image shows a man from the chest up, holding a large sheet of paper that serves as a personality test form. The form is divided into several columns, each with a header: 'NEVER', 'SOMETIMES', 'OFTEN', and 'ALWAYS'. Below these headers are rows of questions, each followed by a checkbox. Some checkboxes are filled in with black ink. The man's face is partially visible behind the top of the form. The form is held in front of a background of text from the article.

Do you often fantasize about being famous?

Do you find yourself getting angry easily?

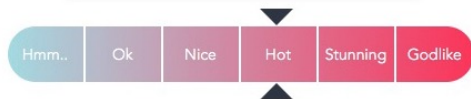
Would you like to be an art collector?

Do people say you are eccentric?

AN EXAMPLE FROM REDDIT

Let Artificial Intelligence guess your attractiveness and age

#howhot



AN EXAMPLE FROM REDDIT

Let Artificial Intelligence guess your attractiveness and age

#howhot



Let Artificial Intelligence guess your attractiveness and age

#howhot



AN EXAMPLE FROM LINKEDIN



Top job picks for you



Assistant Professor

Northeastern University · Boston, Massachusetts, United States



16 connections



Lecturer

Massachusetts Institute of... · Cambridge, Massachusetts, United States



2 connections



Programmer - NEW

Harvard University · Boston, Massachusetts, United States



14 school alumni

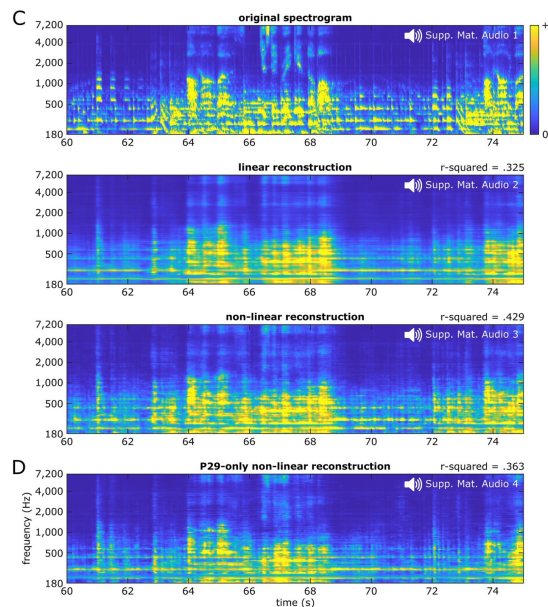
Music can be reconstructed from human auditory cortex activity using nonlinear decoding models

Ludovic Bellier , Anaïs Llorens, Déborah Marciano, Aysegül 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

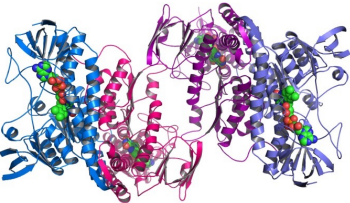
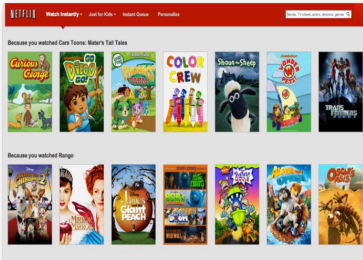
2. Mathematical and Statistical Foundations

- 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

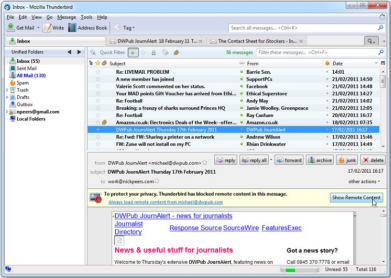
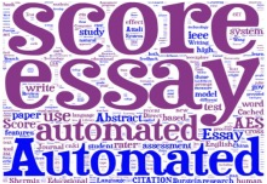
3. Supervised Learning Techniques

- Regression:

WHERE ELSE DO WE SEE IT?



What Can Be Automated?
What Cannot Be Automated?



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

“Probability, [...] 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.”

BRIEF OVERVIEW OF CS6140

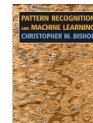
See online syllabus...

- 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.)

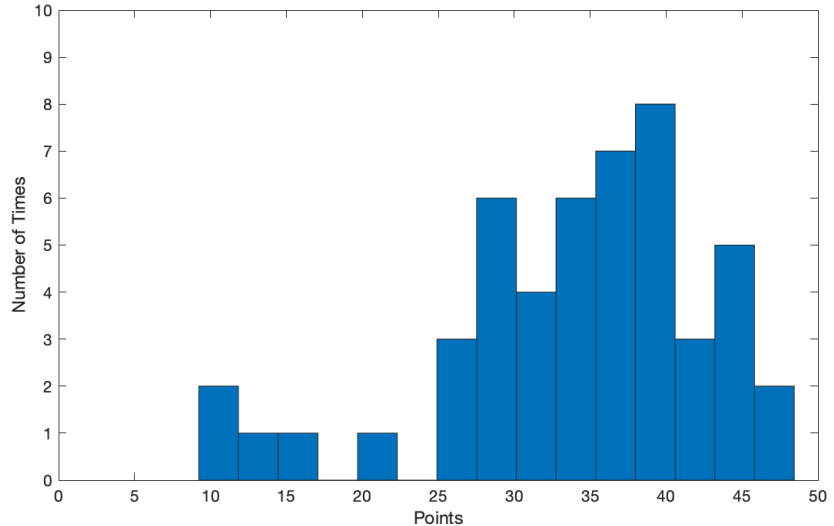
ONE OF PREVIOUS YEARS AFTER MIDTERM

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

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



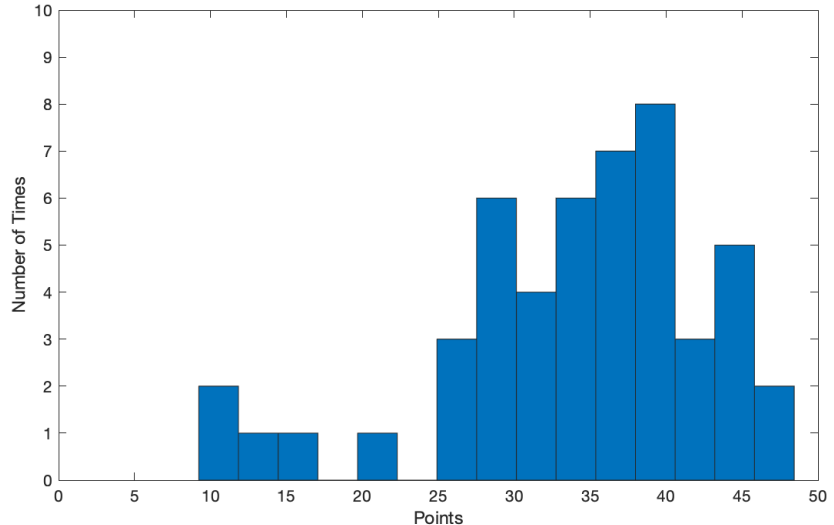
ONE OF PREVIOUS YEARS AFTER MIDTERM

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

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

Percentiles:

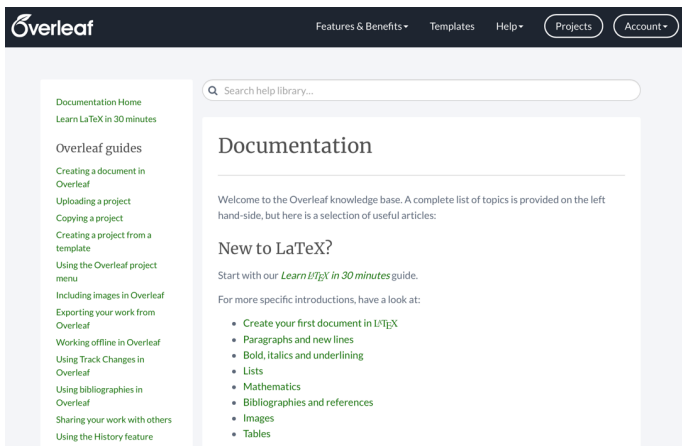
10%: 22.1 points
25%: 29.6 points
50%: 35.6 points
75%: 39.3 points
90%: 44.9 points



TYPING ASSIGNMENTS

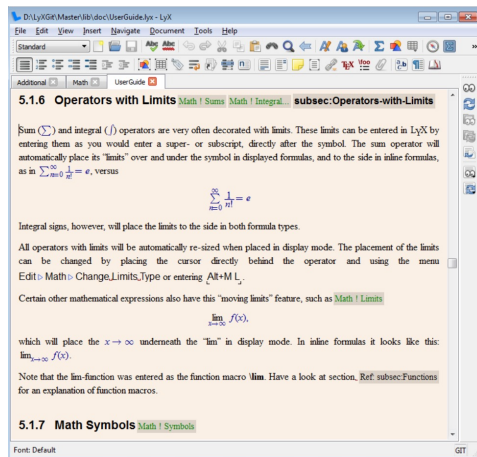
- Latex (TeXShop + MacTeX or TeXnicCenter + MiKTeX)
- Word

Overleaf



The screenshot shows the Overleaf website homepage. At the top, there is a navigation bar with the Overleaf logo on the left and links for 'Features & Benefits', 'Templates', 'Help', 'Projects', and 'Account'. Below the navigation bar is a search bar with the placeholder text '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:'. Below this, there is a section titled 'New to LaTeX?' with the text 'Start with our [Learn L^AT_EX in 30 minutes](#) guide.' and a list of links for more specific introductions: 'Create your first document in L^AT_EX', 'Paragraphs and new lines', 'Bold, italics and underlining', 'Lists', 'Mathematics', 'Bibliographies and references', 'Images', and 'Tables'. On the left side of the page, there is a sidebar with a 'Documentation Home' section and a list of 'Overleaf guides' including 'Learn LaTeX in 30 minutes', '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 software interface. The window title is 'D:\LyX\Git\Master\M\doc\UserGuide.lyx - LyX'. The menu bar includes 'File', 'Edit', 'View', 'Insert', 'Navigate', 'Document', 'Tools', and 'Help'. The toolbar contains various icons for editing and navigation. The main text area displays the following content:

5.1.6 Operators with Limits [Math](#) | [Sums](#) | [Math](#) | [Integral](#) | [subsec:Operators-with-Limits](#)

Sum (\sum) and integral (\int) operators are very often decorated with limits. These limits can be entered in L^AT_EX 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 $\sum_{x=-\infty}^{\infty} \frac{1}{2^x} = e$, versus

$$\sum_{x=0}^{\infty} \frac{1}{2^x} = e$$

Integral signs, however, will place the limits to the side in both formula types.

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**.

Certain other mathematical expressions also have this "moving limits" feature, such as [Math](#) | [Limits](#)

$$\lim_{x \rightarrow \infty} f(x),$$

which will place the $x \rightarrow \infty$ underneath the "lim" in display mode. In inline formulas it looks like this: $\lim_{x \rightarrow \infty} f(x)$.

Note that the `lim`-function was entered as the function macro `\lim`. Have a look at section [Ref. subsec:Functions](#) for an explanation of function macros.

5.1.7 Math Symbols [Math](#) | [Symbols](#)

Font: Default

PLAN

September: 2
9
16
23
30

November: 4
11
18
25

October: 7
14
21
28

December: 2
9

PLAN

September: 2
9 h1
16
23 H1, h2
30

November: 4 h4
11
18 H4
25

October: 7 H2, pp (h3)
14
21 M, PP (H3)
28

December: 2 F
9 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)	}	not recommended!
– 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!
 - 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