

# Course Overview

CSCI 601 471/671  
NLP: Self-Supervised Models

<https://self-supervised.cs.jhu.edu/sp2023/>



# Why This Course?

## Why This Course?



- Let's start with an example: **ChatGPT**
- A computational model trained on massive piles of data and tuned to respond to dialogue commands.

<https://chat.openai.com/chat>



- Let's start with an example: **ChatGPT**



Can you write me a 4 paragraph academic essay comparing and contrasting the theories of nationalism of Benedict Anderson and Ernest Gellner?



Benedict Anderson and Ernest Gellner both have theories about the origins and nature of nationalism, but their perspectives differ in key ways. Anderson believes that nationalism is a cultural construct, arising from the development of print capitalism and the creation of imagined communities. Gellner, on the other hand, sees nationalism as a product of modernization and industrialization, with nations being formed as a means of organizing society and providing a shared identity.

<https://chat.openai.com/chat>



- Let's start with an example: **ChatGPT**
- A computational model trained on massive piles of data and tuned to respond to dialogue commands.
- Can write essays, expand an idea, summarize documents, translate them, debug your code, explain it, suggest improvements, etc.

Meanwhile ...



What will be the gender of the first female US President?



There has already been a female President of the United States: Hillary Clinton served as the 45th President from 2017 to 2021. Is there something else you would like to know about this



Was Lincoln's assassin on the same continent as Lincoln when the assassination occurred



Abraham Lincoln was assassinated by John Wilkes Booth while he was attending a play at Ford's Theatre in Washington, D.C. I do not know where John Wilkes Booth was at the time of the assassination, but Washington, D.C. is located on the continent of North America, so it is possible that Booth was also on that continent at the time.

## Why This Course?



- Let's start with an example: **ChatGPT**
- A computational model trained on massive piles of data and tuned to respond to dialogue commands.
- The model is **extremely good** at important tasks that we can about: why? how?
- It is **easily fails** at seemingly trivial tasks: why?



# Course Learning Objectives

- Learn about **architectures** that serve as foundation for tools like ChatGPT
- Understand a wide range of self-supervised **algorithms**
- Fearlessly build and train self-supervised **models** 🤖
- Gain intuitions about the **capabilities** and **limitations** of our models
- Explore their **applications**, e.g., question answering, text generations
- Reflect on **societal impacts**: their misuse, biases, etc.

## Focus on Natural/Human Language

- **Most of the class revolves around natural language.**

### Why natural language?

- It is a **convenient medium of communication**.
- Natural language is our species' best attempt to encode **everything about the world** as **efficiently** as possible.
- A huge archive of natural language is **freely available** (e.g., on web)

## Course Prerequisites

- Comfortable with programming, particularly Python
- Calculus, linear algebra, probability
- Familiarity with Natural Language Processing
- HW1 should give a sense how prepared you are!

# Course Logistics Brief

- **Instructor:** Daniel Khashabi (You can call me “Daniel”, as long as we act mutually respectfully.)
- **TA:** Adam Byler
- **CAs:** Lingfeng Shen, Xiao Ye, Wufe Ma



Daniel Khashabi  
Instructor



Adam Bylerly  
Teaching Assistant



Lingfeng Shen  
Course Assistant



Xiao Ye  
Course Assistant



Wufe Ma  
Course Assistant

# Self-Supervised Models: Course Website

CSCI 601.471/671	Content	Schedule	Project	Suggest an edit
------------------	---------	----------	---------	-----------------

## CSCI 601.471/671 NLP: Self-supervised Models

Johns Hopkins University - Spring 2023

The rise of massive self-supervised (pre-trained) models have transformed various data-driven fields such as natural language processing (NLP). In this course, students will gain a thorough introduction to self-supervised learning techniques for NLP applications. Through lectures, assignments, and a final project, students will learn the necessary skills to design, implement, and understand their own self-supervised neural network models, using the Pytorch framework.

**Note:** The course is different from 601.771 (offered in the fall semesters) which involves reading recent papers and is geared toward grad students that want to specialize in the latest developments in self-supervised models.

**Prerequisites:** (1) Data Structures (601.226), (2) All the class assignments will be in Python/PyTorch. If you don't know Python or PyTorch but have experience in other programming languages (Java, C++, etc.) you can probably pick Python/PyTorch pretty quickly. (3) Calculus and linear algebra: you should be comfortable with matrix operations (matrix multiplication, transpose, inverse, dot product, gradients). (4) Probability: basic probability properties (conditionals, marginals, mean, standard deviation), distributions (gaussian, categorical). (5) Background in Natural Language Processing & Machine Learning or having finished one of the relevant courses such as Machine Learning (475.675), Artificial Intelligence (464.664), Natural Language Processing (600.465), Machine Translation (600.468), or Introduction to HLT (601.467/667).

**Relevant Courses at Hopkins:** This course has some overlap with "Natural Language Processing" (EN.601/665), "Introduction to Human Language Technology" (601.467/667) and "Artificial Agents" (EN.601.470/670), though the courses have different focuses.

<https://self-supervised.cs.jhu.edu/sp2023>

## Course Logistics Brief (2)

- Lots of important information on the **website**:
  - <https://self-supervised.cs.jhu.edu/sp2023>
  - Syllabus, lectures, videos, tutorials, etc.
- **Recorded videos**: will be available on the course website and/or Piazza.
- **Office hours**: 3 office hours each week; check the course website.
- **Questions or discussions**: Piazza

# Course Work and Grading Policy

- **Homework:** 12 x 1-week assignments: 60%
- **Midterms :** 20%
- **Final course project (1-2 people):** 20%
  - Project proposal: 5%, final poster or web summary: 5%, report 10%

# Homework

- **Timeline:** Released **every Tuesday** and due in a week.
- Each have both **conceptual (theory)** and **empirical (coding)** parts.
- All assignments will be submitted to Gradescope.
- **Late day policy**
  - 10 free late days; afterwards, 5% off course grade per day late
  - Assignments not accepted more than 7 days late (unless given permission in advance)
- **Collaboration policy:** Allowed within certain limits:
  - Needs to be well-documented; acknowledge working with other students.
  - Avoid: copy code off the web; a write your own assignment solutions.
- **HW1 is released today! Due next week. Submitted via Gradescope.**



# Final Project

- Must be **exploring a topic related to the focus of the class.**
- This is your **chance to gain research experience** on a topic of interest.
- Topic choice will be (relatively) free. We will help you develop your ideas!

## **Deliverables:**

1. Submit project proposal outline (for our formal review and suggestions)
  - To make sure that the project is scoped reasonably and doable in your limited time.
2. Get excited 🥳 and work on the project
3. Submit a structured final report
4. Final project presentation

## Attendance Policy

- Not mandatory!

# Attendance Policy

- **Not mandatory!**
- No additional grade if you attend.
- However, it is highly encouraged!!
  - Participation in class is our chance to learn more effectively.
  - Come to the class and participate in the discussions!

## Pop Quiz (1)

- I am a .....: CS major, CS minor, MS student, PhD student, studying something else

respond at **PollEv.com/danielkhashabi994**

## Pop Quiz (2)

respond at **[PollEv.com/danielkhashabi994](https://pollev.com/danielkhashabi994)**

- Select the most appropriate option:
  1. NLP and DL are rather new to me. Looking forward to learning more.
  2. I am a DL, NLP or self-supervised learning enthusiast – I read casually on these topics.
  3. I am practitioner of NLP or ML and have already worked with a variety of self-supervised models.
  4. I am a pro. I could be part of the teaching staff of this class.

## Pop Quiz (3)

- I have read (and understood) the syllabus! Yes, No

respond at **[PollEv.com/danielkhashabi994](https://pollev.com/danielkhashabi994)**

Back to  
Self-Supervised Models ...

## Taco Tuesday



Jacqueline Bruzek ×

## Taco Tuesday

Hey Jacqueline,

Haven't seen you in a while and I hope you're doing well.



The best course in Johns Hopkins  
computr science ...

• SPELLING

~~computr~~ → computer

If you don't want **computr** to be marked as misspelled in the future, you can add it to your personal dictionary.

 Add to dictionary



6:56



Google Translate



English



Spanish



ENGLISH



I would like a beer please.



Camera



Conversation



Transcribe

SPANISH



Me gustaría una cerveza por favor.



Hi, how can I help?

what time will my Hermes parcel arrive



Your parcel from Hermes  
will arrive today



# DALL-E Generation

“Photo of hip hop cow in a denim jacket recording a hit single in the studio”



# Popular Media: AI is Solved!!

FINANCIAL TIMES

**Is AI finally closing in on human intelligence?**

INSIDER

**Google's DeepMind artificial intelligence has figured out how to talk**

*The Atlantic*

An Artificial Intelligence Developed Its Own Non-Human Language

## Self-Supervised Models



# Self-Supervised Models

# Self-Supervision



[Slide from Colin Raffel]

# Self-Supervision



[Slide from Colin Raffel]



# Self-Supervision



[Slide from Colin Raffel]

# Self-Supervision



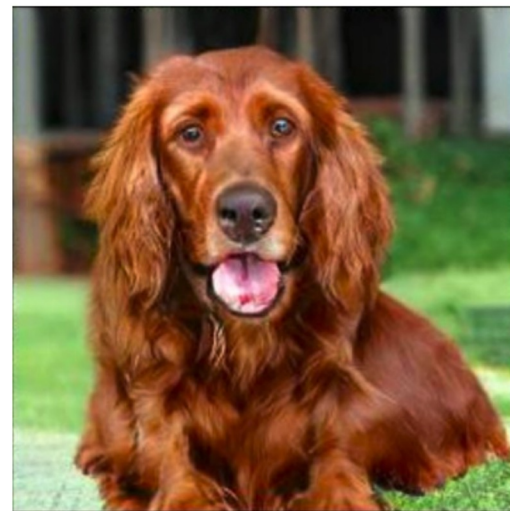
*Dataset of natural images*

[Slide from Colin Raffel]

# Self-Supervision



Dataset of natural images



Generated image, from "Large Scale GAN Training for High Fidelity Natural Image Synthesis", Brock et al.

[Slide from Colin Raffel]

# Self-Supervision



*Dataset of natural images*



*Generated image, from "Large Scale GAN Training for High Fidelity Natural Image Synthesis", Brock et al.*

[Slide from Colin Raffel]





# Self-Supervision



Dataset of natural images



Generated image, from "Large Scale GAN Training for High Fidelity Natural Image Synthesis", Brock et al.

[Slide from Colin Raffel]

# Self-Supervision

== treaty of paris (1763)

the treaty of paris, also known as the treaty of 1763, was signed on 10 february 1763 by the kingdoms of great britain, france and spain, with portugal in agreement, after great britain's victory over france and spain during the seven years' war.

the signing of the treaty formally ended the seven years' war, known as the french and indian war in the north american theatre, ....

# Self-Supervision

== wheelbarrow

== t

A wheelbarrow is a small hand-propelled vehicle, usually with just one wheel, designed to be pushed and guided by a single person using two handles at the rear, or by a sail to push the ancient wheelbarrow by wind. The term "wheelbarrow" is made of two words: "wheel" and "barrow." "Barrow" is a derivation of the Old English "barew" which was a device used for carrying loads. The wheelbarrow is designed to .....  
north american theatre, ....



# Self-Supervision



== lemon

WIKIPEDIA  
The Free Encyclopedia

== w

== t

A wh  
the prop  
trea one  
febr and  
grea two  
port to p  
brita win  
spai mad  
"bar  
the the  
end devi  
as th  
north american theatre, ....

The lemon (*Citrus limon*) is a species of small evergreen trees in the flowering plant family Rutaceae, native to Asia, primarily Northeast India (Assam), Northern Myanmar or China.[2] The tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses.[2] The pulp and rind are also used in cooking and baking. ....

# Self-Supervision



WIKIPEDIA  
The Free Encyclopedia

== lemon

The lemon (*Citrus limon*) is a species of small evergreen trees in the flowering plant family Rutaceae, native to Asia, primarily Northeast India (Assam), Northern Myanmar or China.[2] The tree's ellipsoidal yellow fruit is used for culinary and non-culinary purposes throughout the world, primarily for its juice, which has both culinary and cleaning uses.[2] The pulp and rind are also used in cooking and baking. ....

== wings over kansas

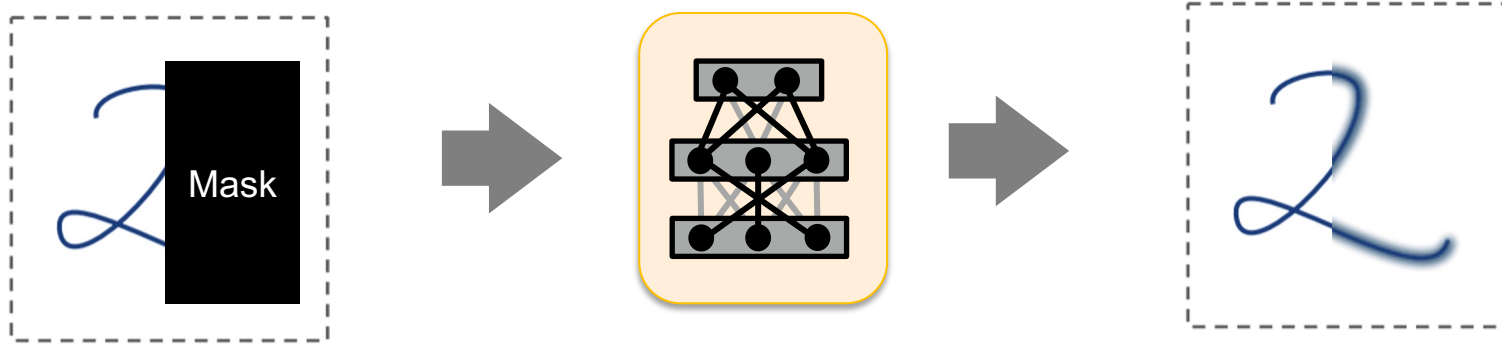
wings over kansas is the second studio album by jason ammons, john bolster and mo rosato. the album debuted at number one on the billboard 200, selling 35,000 copies in it first week at the time. it was the second highest selling album to debut at the billboard top 50 and the third highest selling album to debut at the top heatseekers, with 26,000 copies sold. this is the supremes album earning the nickname ...

# Self-Supervised Models

*are **predictive models** of the world!*

# Self-Supervised Models

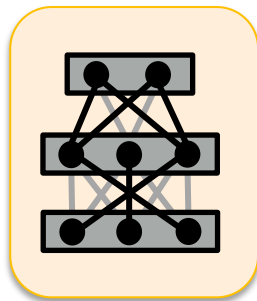
- Are trained to complete partial samples from the world.



[Bengio et al. 2004, Hinton et al. 2006, Peters et al. 2018, ...]

# Self-Supervised Models

“Wings over Kansas is [MASK]”



“Wings over Kansas is  
an aviation website  
founded in 1998 by Carl  
Chance owned by Chance  
Communications, Inc.”

[Bengio et al. 2004, Hinton et al. 2006, Peters et al. 2018, ...]

# Self-Supervised Models

*Learned from **cheaply available** unlabeled data*

# Self-Supervised Models

*Learned from **cheaply available** unlabeled **data***

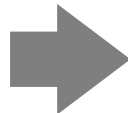
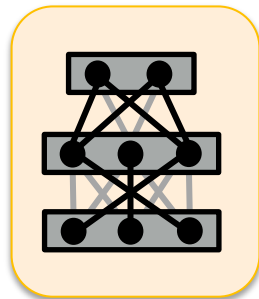
# Self-Supervised Models and End Tasks

- **Goal:** Answering questions

**Question:** “Where is the birthplace of the American national anthem?”



“The birthplace of the American national anthem” [MASK]



“The birthplace of the American national anthem, “The Star-Spangled Banner,” lies in Baltimore, Maryland.”



## Self-Supervised Models

- Are *predictive models* of the world.
- Are learned from *unlabeled* data.
- Tightly connected to *tasks* we care about.

Questions? 