

Course Overview

CSCI 601.771: Self-supervised Statistical Models

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



JOHNS HOPKINS
UNIVERSITY

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



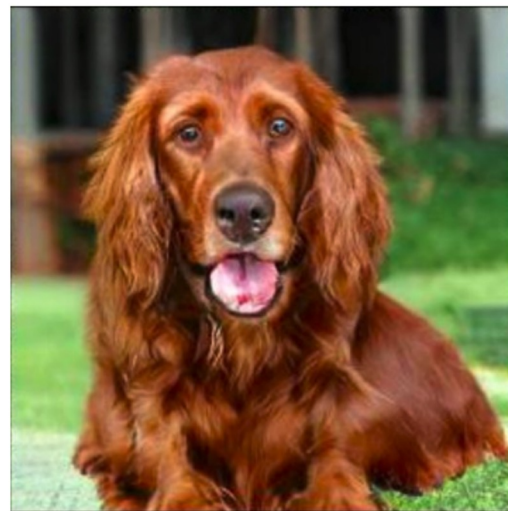
Dataset of natural images

[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



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

== tr
the t
trea
febr
brita
in ag
victo
seve
the s
the s
fren
american theatre,

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 distribute the weight of its load between the wheel ...

[Slide from Colin Raffel]

Self-Supervision



== lemon

WIKIPEDIA
The Free Encyclopedia

== w

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. The juice of the lemon is about 5% to 6% citric acid, with a pH

== tr

the t
treat
febr
brita
in ag
victo
seve
the s
the s
fren
american theatre,

load between the wheel ...

Self-Supervision



WIKIPEDIA
The Free Encyclopedia

== lemon

== w

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. The juice of the lemon is about 5% to 6% citric acid, with a pH

load between the wheel ...

== tr

the t

trear

febr

brita

in ag

victo

seve

the s

the s

fren

american

theatre,

== wings over kansas

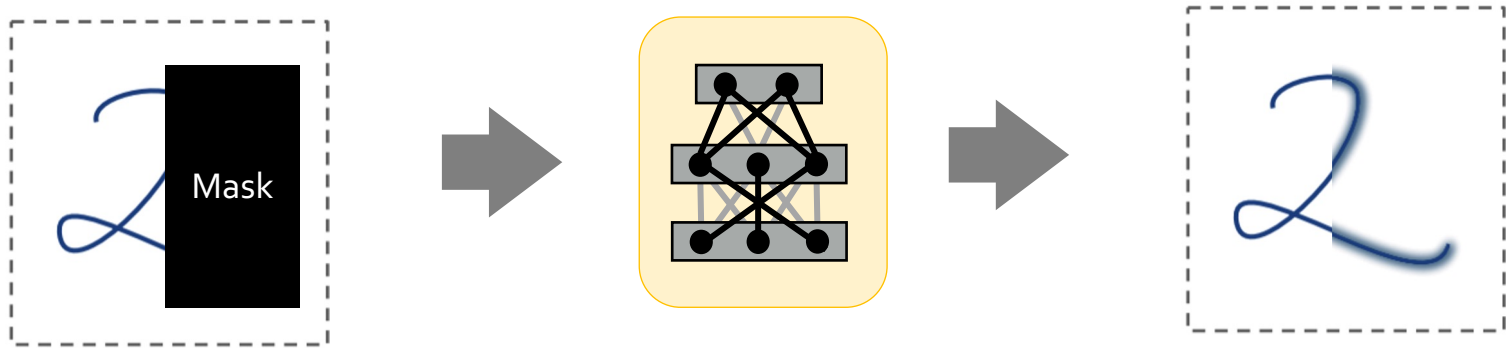
wings over kansas is the second studio album by jason ammons, john bolster and mo rosato. the album debuted at number one on the billboard 200, selling 35,000 copies in it first week at the time. it was the second highest selling album to debut at the billboard top 50 and the third highest selling album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

Self-Supervised Models

are *predictive models* of the world!

Self-Supervised Models

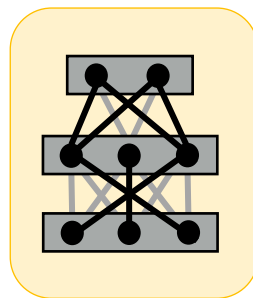
- Are trained to complete partial samples from the world.



[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

- *Learn from unlabeled data – a special case of **unsupervised** learning*

Why not then call them “unsupervised”?

Self-Supervised Models

Why not then call them "unsupervised"?

Predictive models of the world
are tightly connected to
many **tasks we care about***!

* Thus far

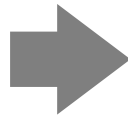
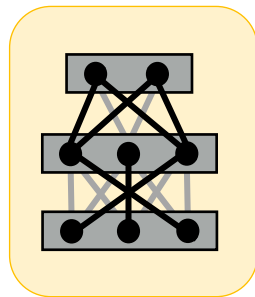
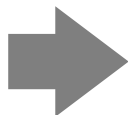
Self-Supervised Models and End Tasks

- **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 and End Tasks

- **Goal:** Sentiment classification

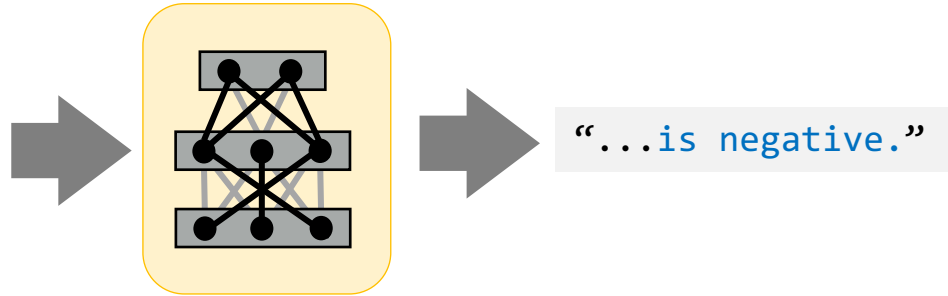
Review: “While this restaurant is popular on Google, I absolutely disliked it.”



“We want to decide whether the sentiment of the review is “positive” or “negative”.

Review: “While this restaurant is popular on Google, I absolutely disliked it”.

The sentiment of this review is” [MASK]



Self-Supervised Models

*Are representations of the world learned from **unlabeled** data in a way that the conveniently applicable **end tasks**.*

Self-Supervised Models: Course Website

CSCI 601.771 Course Schedule Project Edit this page!


CSCI 601.771: Self-supervised Statistical Models

Johns Hopkins University - Fall 2022

The rise of massive self-supervised (pre-trained) models has transformed various data-driven fields such as natural language processing, computer vision, robotics, and medical imaging. This advanced graduate course aims to provide a holistic view of the issues related to these models: We will start with the history of how we got here, and then delve into the latest success stories. We will then focus on the implications of these technologies: social harms, security risks, legal issues, and environmental impacts. The class ends with reflections on the future implications of this trajectory.

Prerequisites: Students must have extensive experience with deep learning, machine learning, artificial intelligence, and natural language processing. Familiarity with linear algebra, statistics and probability are necessary, as well as with the design and implementation of learning models (via one of the learning libraries, such as PyTorch, Tensorflow, Keras, JAX). Students must be comfortable with reading papers and extracting key concepts and ideas from papers.

Instructor



Daniel Khashabi

Logistics

- **Classes:** on Tuesday/Thursday 1:30 - 2:45 pm EST (Malone 228)
- **Office hours:** by appointment
- **Contact:** If you have any questions about the course email [the instructor](#).
- **Class Structure:** The class will be in-person. Each session will involve the presentation and/or discussion of recent important papers on the self-supervised statistical models. The course also involves a project.
- **Coursework:** Your grade is based on three activities: paper presentation by undertaking your role and present it in a clear and compelling way (40%), in-class participation in discussions for any class that you're not presenting (20%) and the class project (40%).
- **Changes:** The professor reserves the right to make changes to the syllabus or project due dates. These changes will be announced as early as possible.
- **Recordings:** Recorded versions of the sessions will be available online after each class on Canvas.

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

Self-Supervised Models: Topics

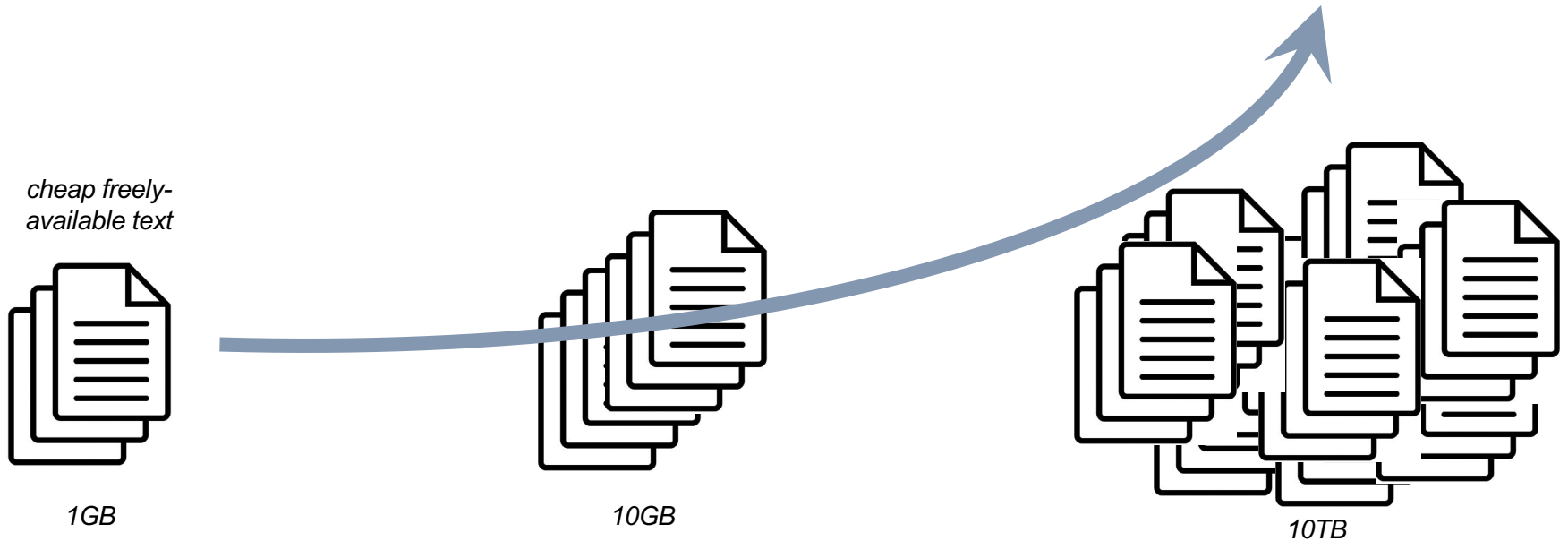
- Self-supervised models for language, vision, audio/speech, etc.
 - What **models** are available?
 - What **architectures** are these based on?
 - How are these models **trained**?
 - What are the **capabilities**?
 - What are their **weaknesses**?

Self-Supervised Models: Topics

- Self-supervised models for language, vision, audio/speech, etc.
- Behavior change as a function of scale

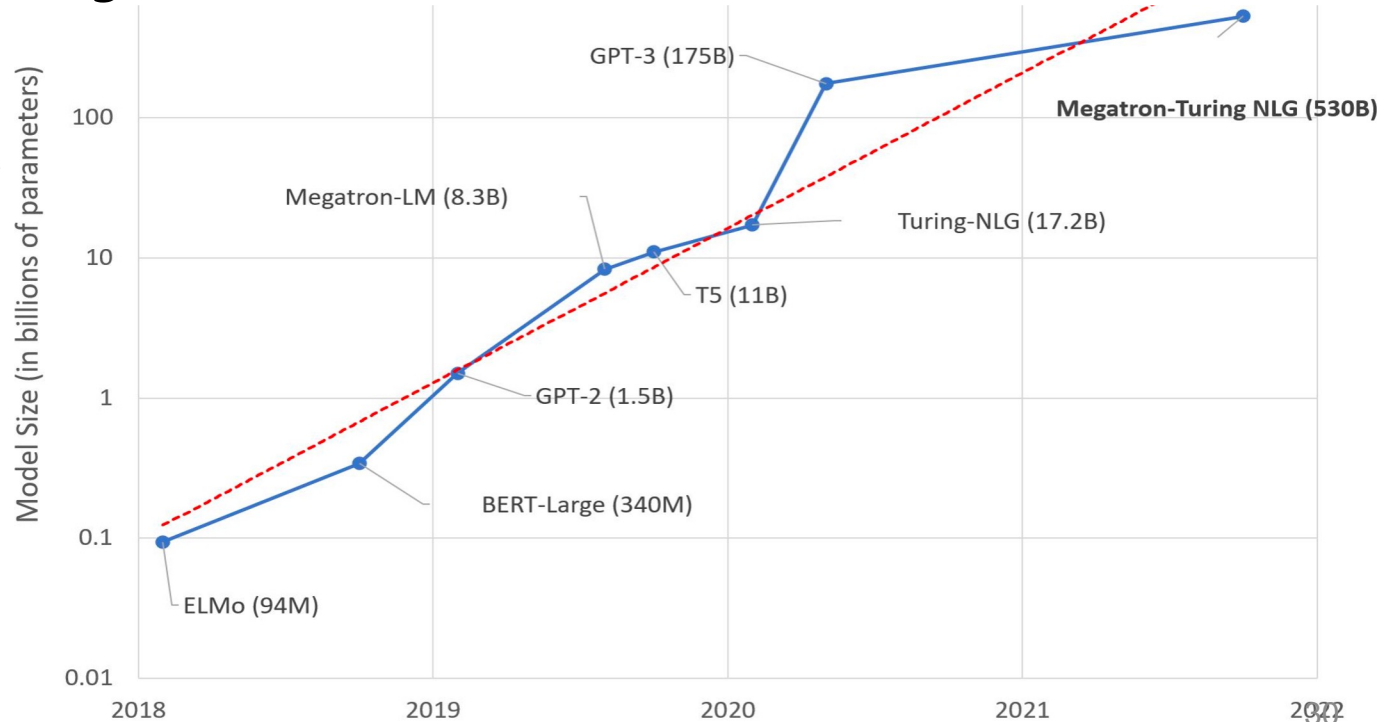
Scaling

- Larger pre-training datasets



Scaling

- Larger pre-training datasets
- Larger models



Self-Supervised Models: **Topics**

- Self-supervised models for language, vision, audio/speech
- Behavior change as a function of **scale**
 - What is really the nature of “**scale**”?
 - How does the **quality change** with scale? (end tasks, biases, etc.)
 - How **efficiently** can we scale?
 - In an ideal **limit**, how far will scale lead to benefits?
 - Is scale **the right “hill”** to climb? (Should we really scale?!)

Self-Supervised Models: Topics

- Self-supervised models for language, vision, audio/speech
- Behavior change as a function of scale
- The problematic content: bias, toxic language, etc.

Outlook

He Is A Doctor, She A Nurse: How Language Carries Gender Bias Into Algorithms, Perpetuates Status Quo

May 7



Self-Supervised Models: Topics

- Self-supervised models for language, vision, audio/speech
- Behavior change as a function of scale
- The problematic content: bias, toxic language, etc.
- Environmental impacts

Self-Supervised Models: Topics

- Self-supervised models for language, vision, audio/speech
- Behavior change as a function of scale
- The problematic content: bias, toxic language, etc.
- Environmental impacts
- The future of these models and their impacts on our lives

A Bias ...

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 quantity of writing is **freely available** on the internet.

What We Will **Not** Cover

- Multilinguality & Cross-linguality → see Kenton Murray's class
- Multi-step decision-making, dialogue → see Ben Van Durme's class
- Actively learning
- Human-in-loop
- Data augmentation
-

What I Expect You

- Comfortable with machine learning.
 - **Modeling:** linear models, classification, neural networks
 - **Training:** gradient descent, backpropagation, train/test/dev splits
 - **Measuring quality:** generalization: overfitting vs underfitting
- Familiarity with NLP is helpful, though not necessary.
- Being open to reading [somewhat esoteric] papers and presenting their gist to the class.

Questions so far?

Class Structure

- The class will be **in-person**.
- Each session will involve **the presentation/discussion** of a recent important paper on self-supervised models.
- The course also involves **a project**.

Class Presentations

- Role-based presentation

Role-Playing Paper-Reading Seminars

Alec Jacobson and Colin Raffel

March 17th, 2021

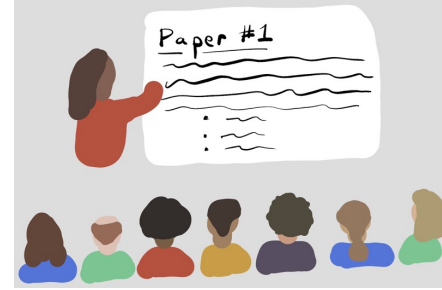
colinraffel.com/blog

<https://colinraffel.com/blog/role-playing-seminar.html>



• Role-based presentation vs.

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.
- The “role” defines **the lens** through which you read/present a paper.



• One-to-Many presentations

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage

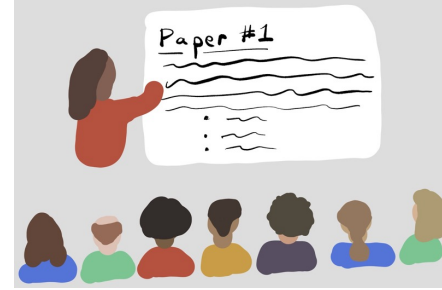


- Role-based presentation

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.

Role: Stakeholder 🖋️

Act as if you're the author of this paper. Try to sell it!



- One-to-Many presentations

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage

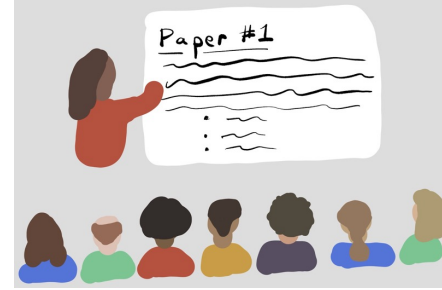


- **Role-based presentation**

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.

Role: Scientific Reviewer 

Do a complete conference-style critical peer-review of the paper.



- **One-to-Many presentations**

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage

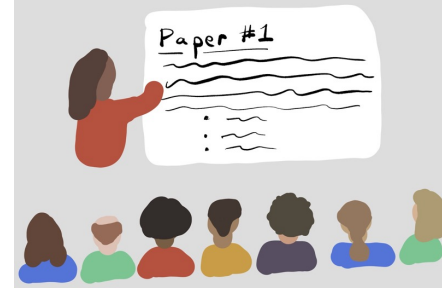


- **Role-based presentation**

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.

Role: Empiricist 🧑🏫

Implement something related to the paper and your finding and code with the class.



- **One-to-Many presentations**

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage

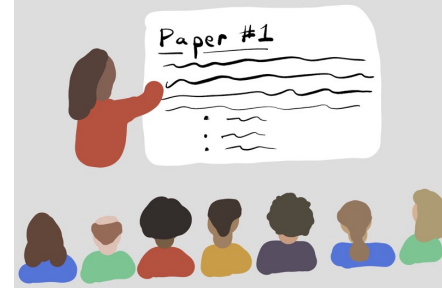


- **Role-based presentation**

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.

Role: Archaeologist 🏺

Determine the [prior and recent] literature that inspired and was inspired by this work.



- **One-to-Many presentations**

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage

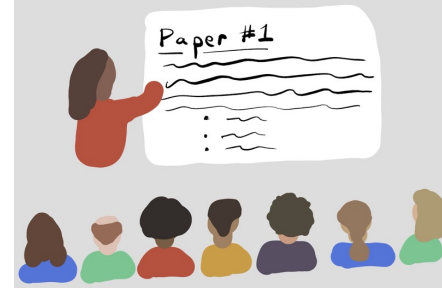


- Role-based presentation

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.

Role: Visionary 

Propose an imaginary follow-up -- research project or a new application.



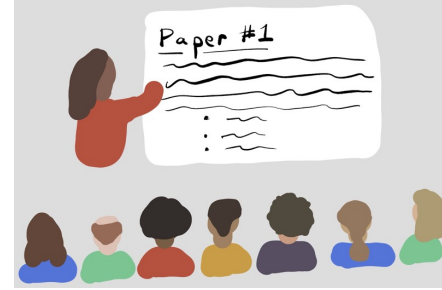
- One-to-Many presentations

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage



- Role-based presentation

- Many students **cooperatively** present a paper.
- Each subgroup of students takes a specific **“role”**.
- Students **rotate** “roles” each week.



- One-to-Many presentations

- A single (subgroup of) student(s) presenting a paper to the class.
- **Pro:**
 - Easy division of labor
- **Cons:**
 - Too much work for one person
 - Audience easy to disengage



- **Role-based presentation**

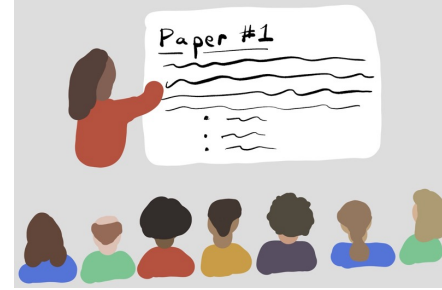
- Many students **cooperatively** present a paper.

- **Pro:**

- More engagement
- Distributed and less workload
- Present more frequently (once a week) 😭

- **Cons:**

- Need to manage role assignment



- **One-to-Many presentations**

- A single (subgroup of) student(s) presenting a paper to the class.

- **Pro:**

- Easy division of labor

- **Cons:**

- Too much work for one person
- Audience easy to disengage
- Present a 1-2 times only.


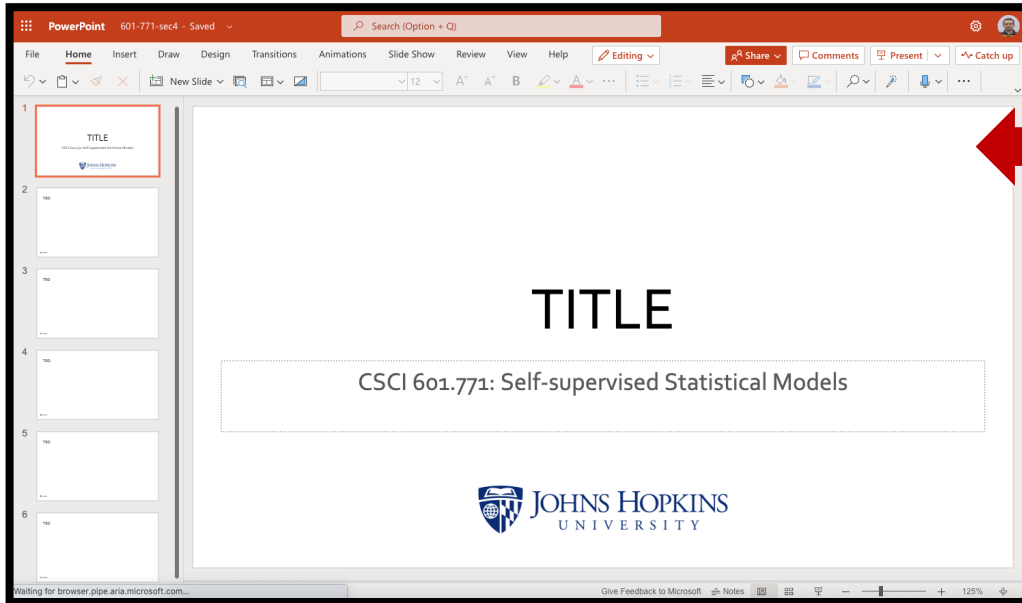
Role-Based Presentation

- You will be split into two halves:
 - One half presenting on **Tuesdays** and the other on **Thursdays**.
- Each member of the presenting half will be given a random role every week.
- Each role has a time budget:
 - 18 mins for Stakeholder 🖋️
 - 12 mins for the rest of the roles

Questions so far?

Slides?

- Using slides is encouraged though optional.
 - A shared pool of slides (links accessible on the class website).



Slides

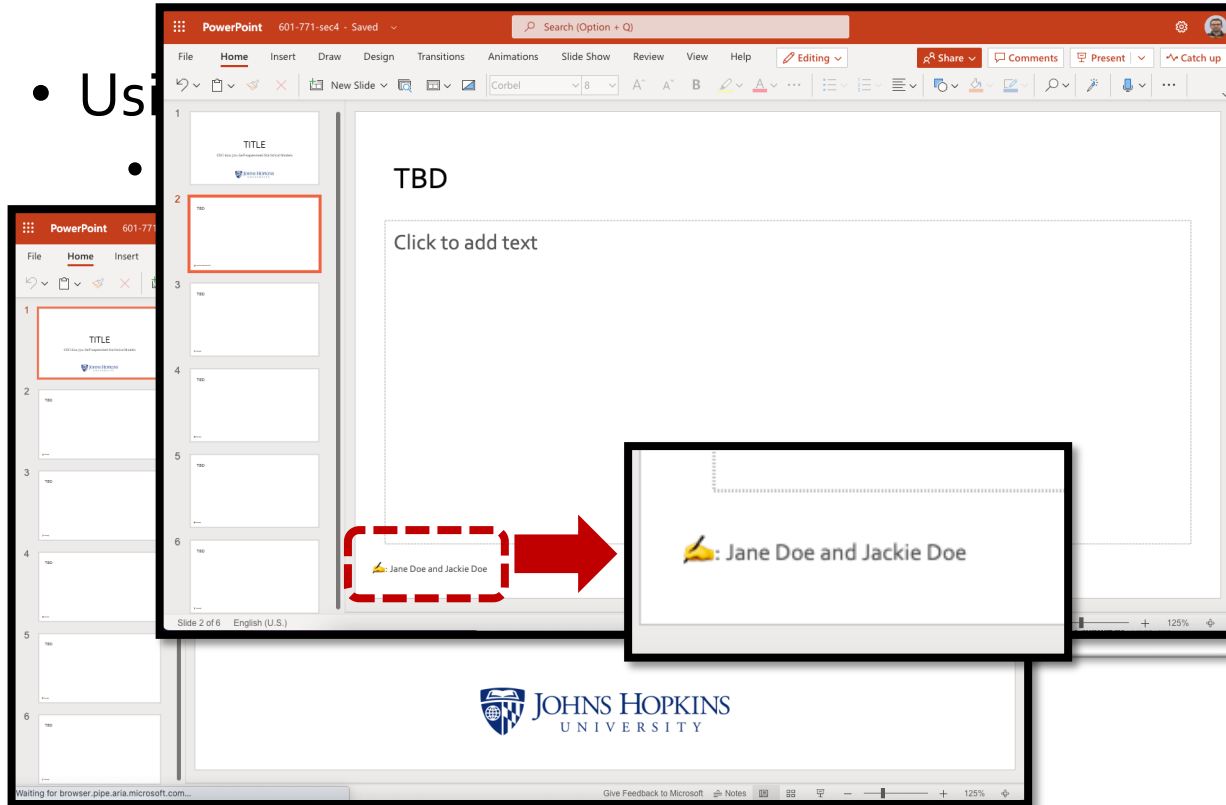
Main Reading: BERT: Pre-training of Deep Bidirectional

Additional Reading(s):

1. RoBERTa: A Robustly Optimized BERT Pretraining
2. Exploring the limits of transfer learning with a uni
3. BART: Denoising Sequence-to-Sequence Pre-train

Slides?

- Use



... website).

...g: BERT: Pre-training of Deep Bidirectional

...ading(s):

...RTa: A Robustly Optimized BERT Pretraining

...ring the limits of transfer learning with a uni

...: Denoising Sequence-to-Sequence Pre-trai

Non-presenter Activity

- **Before the class:** a short answer to a prompt. For example,
 - “An alternative title for this paper is _____”
 - “My least favorite thing about this paper is _____”
- **Beginning of each class:** one question about the paper
 - Something you're confused about or something you'd like to hear discussed more.
- **During the class:** come to class ready to participate in the discussions.

Class Engagement

- You have the freedom (and encouraged) to share your **concerns, questions** or **suggestions**, while being respectful.
- I hope this class will be an **open** and **safe** space for class discussions.
- Remember to *Step Up & Step Back*.

Attendance

- You can miss 3 sessions. Drop me a note before the class.
 - If you're "non-presenting", that's **easy!**
 - If you're "presenting", that's a bit **complicated**:
 - Find someone willing to swap presentations with.
 - Create the presentation for that role and find someone else to present.
- If you have any **COVID symptoms**, skip the class.
 - Does not count toward your 3 sessions.
 - Drop me a note before the class.

Class Project

- **One** project for the whole class.
- **Groups** of size 2 or 3. You can start looking now! 🥹
- The topic is **open-ended**.
 - Examples,
 - demonstrating systemic limitations of prior work,
 - suggesting improvements on methods or benchmarks discussed in the class, and so on.
 - If you want to chat/brainstorm about it, shoot me an email.
- Check the course page for the project deadlines.

Your Grade

- It's a function of three items:
 - Paper presentation (33%)
 - In-class participation (33%)
 - Class project (33%)
 - Bonus point (1%)
- If you're engaged in class presentations and on top of your project, you should not be worried about the grade.

Last call for questions! 

Question for You

- What is the best medium of communication for us? (Slack? Email? Canvas?)
 - Announcements, role assignments, cancellations, broad discussions, etc.
- How can I/we make this a better learning environment for each other?
- How many people have [used/read the paper for] X?