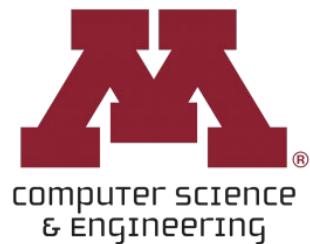


CSCI 5541: Natural Language Processing

Lecture 1: Class Overview

Dongyeop Kang (DK), University of Minnesota

dongyeop@umn.edu | twitter.com/dongyeopkang | dykang.github.io



Overview

- ❑ Instructors
- ❑ Course information
- ❑ Grading and Late Day Policy
 - Class Participation
 - Homework
 - Project
 - Reading Assignment
- ❑ Outline of Topics
- ❑ Computing resources
- ❑ Pre-class Survey

Instructors



Dongyeop Kang (DK)
Instructor



Shirley Anugrah Hayati
Graduate TA



James Mooney
Graduate TA



Zheng Robert Jia
Undergraduate TA

- DK: Friday 2:30pm - 3pm in Shepherd 259
- Shirley: Monday 3:30pm - 4pm via [Zoom](#)
- James: Wednesday 3pm - 4pm via [Zoom](#)
- Robert: Tuesday 10-11am in Lind 4-th floor at Lind/ME Lab

Course Information

- Class meets:
 - Tuesday and Thursday, 11:15AM to 12:30PM
 - Keller Hall 3-125
 - UNITE / Remote (recording will be available for everyone right after the class)
- Class homepage
 - <https://dykang.github.io/classes/csci5541/F24>
- HW/Project discussion on Slack
 - <https://csci5541f24.slack.com/>
- HW/Project submission and grading on Canvas
 - canvas.umn.edu/courses/460609

Outline

- The purpose of this course is to provide an overview of the *computational techniques* developed to enable computers to **understand** and **respond** appropriately to ideas expressed using natural languages.
- This course will cover a broad range of *computational models* in NLP,
 - from language representations, to development and evaluation of NLP applications, to practical coding/machine-learning skills to implement them.
 - text classification, distributional representation methods of language, large language models, and advanced techniques used in chatGPT

IMPORTANT NOTE for background

- This course covers NLP techniques based on recent *deep learning* models, assuming students' knowledge of basic machine learning
- This course focuses ***practical learning*** of NLP knowledge
 - Programming & reading assignments and a semester-long project
 - Assignments and project require programming with PyTorch
 - We use Python for programming and LaTex for report writing
- If students are not familiar with either one of the fields, they **should consider dropping the course** as it would be very challenging and the course is not designed to cover all aspects.

Prerequisite

□ Required:

- CSCI 2041 Advanced Programming Principles
- **CSCI 5521 Machine Learning Fundamentals,**

□ Recommended but not mandatory:

- CSCI 5527 Deep Learning, or
- Any other relevant course that covers fundamental machine learning algorithms.

□ Furthermore, this course assumes:

- Background in basic **probability**, **linear algebra**, and **calculus**.
- Good coding ability, corresponding to senior undergraduate CS major.

✓ **Assignments will be in Python with PyTorch.**

What we don't cover in this class (1)

❑ Classical ML techniques and NLP algorithms

- E.g., SVM, HMM, CRF, Viterbi
- Take CSCI 5521 Machine Learning

❑ Advanced techniques for NLP models

- E.g., GNNs, Recursive Nets, Diffusion, latent variable models,
- Take CSCI 5980 NLP with Deep Learning

❑ Human aspects of AI and NLP systems

- E.g., interaction design, formative/field study, Qual analysis, theory grounding
- Take CSCI 5115 User Interface Design or other HCI classes

What we don't cover in this class (2)

❑ Linguistic aspects or computational linguistics:

- Take classes in Institute of Linguistics
<https://cla.umn.edu/linguistics/graduate/courses>

❑ Emerging and latest NLP techniques and research skills

- Take the CSCI 8980 Intro to NLP Research

❑ Basic Python programming skills

- Basic tutorials on PyTorch and HuggingFace programming will be given by TAs

8980 vs 5980 vs 5541

- Some lectures will be shared across the three classes but they have different focuses:
 - 5980 (NLP with Deep Learning)
 - ✓ focuses on more "processing" parts of NLP, particularly with deep learning methods.
 - 8980 (Intro to NLP Research)
 - ✓ covers broad aspects of NLP research, including theory grounding, data annotation, error analysis, emerging topics, and applications to different fields.
 - 5541 (NLP, *this course*)
 - ✓ is an *introductory* class to cover some basic computational methods with applications and representations, and learn practical skills to apply them to your work
 - 55XX (Large Language models, *TBD*)
 - ✓ is an advanced course that cover a broad ranges of LLMs from technical knowledge, to latest learning methods, to applications, to societal impacts

Class Slack

- ❑ <https://csci5541f24.slack.com/>
- ❑ Slack will be the primary method of communication.
- ❑ A separate channel for UNITE/Remote students
- ❑ “prj-[A/B]-MENTOR-TEAMNAME”
 - With instructors
- ❑ “prj-[A/B]-TEAMNAME”
 - Without instructors

```
# announcement  
# hw0  
# hw1  
# hw2  
# hw3  
# hw4  
# project  
# random  
# unite-and-remote-students
```

prj-b-zae-semantics



Group A/B Mentor Project
Mentor name

- ▼ projects
 - ❑ prj-a-dk-golish-project
 - ❑ prj-a-dk-lm_bros
 - ❑ prj-a-dk-the-adversaries
 - ❑ prj-a-dk-transformative-attentors
 - ❑ prj-a-shirley-clickbait-analysis
 - ❑ prj-a-shirley-pilot
 - ❑ prj-a-shirley-verbavores
 - ❑ prj-a-zae-word-nerds
 - ❑ prj-b-dk-language-model-alignment
 - ❑ prj-b-dk-tbd
 - ❑ prj-b-shirley-title-fight
 - ❑ prj-b-shirley-word-wizards
 - ❑ prj-b-zae-flashcard-generator
 - ❑ prj-b-zae-nlp-vision
 - ❑ prj-b-zae-semantics
 - ❑ prj-b-zae-vjk

Slack communication etiquette

- ❑ Don't ask about homework via *emails* or *Direct Messages* to TAs;
 - Questions MUST be posted in the homework channel (e.g., hw#1, hw#2) or during TA's office hours.
- ❑ Technical / coding questions should follow the stack overflow format
 - <Error message> + context <what you tried so far>
- ❑ Responses outside office hours (9-6pm) will be delayed
- ❑ Be **respectful** and **supportive**
 - Anyone can respond to each other (Slack participation point). Feel free to answer others' questions if you have already addressed similar issues.

Canvas

<https://canvas.umn.edu/courses/460609>

2024 Fall (08/12/2024-01...

- Home
- Announcements
- UNITE Videos
- Assignments
- Discussions
- Grades
- People
- Pages
- Files
- Syllabus
- Outcomes
- Rubrics
- Quizzes
- Modules
- Collaborations
- Attendance
- Chat
- LockDown Browser
- Send or Export Grades to PeopleSoft
- Item Banks
- Course Admin Tools
- Check Accessibility

Course Summary:

Date	Details	Due
Fri Sep 13, 2024	HW1: Building MLP-based text classifier with pytorch	due by 11:59pm
Thu Sep 19, 2024	Team formation	due by 11:59pm
Fri Sep 20, 2024	HW2: Finetuning text classifier using HuggingFace	due by 11:59pm
Tue Oct 1, 2024	Project brainstorming	due by 11:59pm
Fri Oct 4, 2024	HW3: Authorship attribution using language models (LMs)	due by 11:59pm
Tue Oct 8, 2024	Proposal pitch	due by 11:59pm
Sun Oct 13, 2024	Proposal report	due by 11:59pm
Fri Oct 18, 2024	HW4: Generating and evaluating text generated from pretrained LMs	due by 11:59pm
Thu Oct 31, 2024	Midterm office hour participation	due by 11:59pm
Fri Nov 8, 2024	HW5: Prompting with large language models (LLMs)	due by 11:59pm
Fri Nov 22, 2024	HW6: Essay writing with ChatGPT	due by 11:59pm
Thu Dec 5, 2024	Poster presentation	due by 11:59pm
Thu Dec 12, 2024	Final report	due by 11:59pm
	Class Participation	
	Office Hour Participation	

- Homework
 - HW1: Building MLP-based text classifier with pytorch
Due Sep 13 at 11:59pm | 5 pts
 - HW2: Finetuning text classifier using HuggingFace
Due Sep 20 at 11:59pm | 10 pts
 - HW3: Authorship attribution using language models (LMs)
Due Oct 4 at 11:59pm | 10 pts
 - HW4: Generating and evaluating text generated from pretrained LMs
Due Oct 18 at 11:59pm | 15 pts
 - HW5: Prompting with large language models (LLMs)
Due Nov 8 at 11:59pm | 15 pts
 - HW6: Essay writing with ChatGPT
Due Nov 22 at 11:59pm | 5 pts
- Project
 - Team formation
Due Sep 19 at 11:59pm | 1 pts
 - Project brainstorming
Due Oct 1 at 11:59pm | 1 pts
 - Proposal pitch
Due Oct 8 at 11:59pm | 3 pts
 - Proposal report
Due Oct 13 at 11:59pm | 5 pts
 - Midterm office hour participation
Due Oct 31 at 11:59pm | 5 pts
 - Poster presentation
Due Dec 5 at 11:59pm | 5 pts
 - Final report
Due Dec 12 at 11:59pm | 10 pts
- Participation
 - Slack Participation
4 pts
 - Office Hour Participation
4 pts
 - Class Participation
3 pts



Grading and Late Day Policy

- 60% Homework (hw1/2/3/6 for individual, hw4/5 for team)
- 30% Project (team)
- 10% Class Participation (individual)

- Each student will be granted **5 late days** to use for homework and reading assignment over the duration of the semester. After all free late days are used up, penalty is 1 point for each additional late day.
- The late days and penalty will be applied to **all team members** for group homework and project.
- DRC and other medical urgencies can only be accommodated with a doctor's note or a letter from the DRC



Class participation (10%)

Template:						
Weekly Participation on Slack (5): Weekly Participation in office hours (5): Weekly Participation in class (5): Scaled on 10 (with minimum 6/10):						
Weekly Participation Count on Slack	Weekly Participation on Slack (5 scaled)	Weekly Participation Count in office hours	Weekly Participation in office hours (5 scaled)	Weekly Participation in class (5 scaled)	Total (15)	Repeatedly Participating?
6	3	1		5	5	Very active
						10

- Your class participation is **thoroughly** evaluated.
- The following metrics will be used to grade your participation:
 - Participation and discussion in class (no count for UNITE/Remote students)
 - **Discussion on Slack and during Office Hours for both instructor and TAs
 - Discussion and QA during the presentation of the project proposal and poster
- We explicitly **count the number of your offline and online participation**, and (min/max) normalize them at the end of the class.
- Your participation score will be zero if you haven't participated in class, Slack or other discussions.

Homework Details (60%)

- HW1, 2, 3 and 6 are individual assignments
- HW4 and 5 are team assignments (the same team for projects)
- Team assignments
 - Maximum of 4 people
 - Submit your team information by **Sep 19**
- All questions regarding homework **MUST be communicated with the lead TA** over Slack homework channels (e.g., #hw1) or during their office hours.

Homework Details (60%)

- ❑ The use of outside resources (books, research papers, websites), collaboration (students, professors) or AI tools (ChatGPT) are encouraged but **MUST be explicitly acknowledged**. Check out the Academic Integrity policy (<https://dykang.github.io/classes/csci5541/F24/index.html#note>).
- ❑ The deadline for all homework is by **Friday midnight (11:59PM)** of the due date. Since our schedule is quite tight, there will be no deadline extension, but you can still use your late days. For the delayed **team homework and project**, late days for every team member will be used.



Use of AI Tools (e.g., chatGPT)

- ❑ Encouraged to use AI tools like ChatGPT, CoPilot, etc.
- ❑ You can use it to:
 - Clarify concepts learned in class
 - Get instant answers to your questions before emailing TAs and instructor
 - Test your knowledge by asking it questions
- ❑ Warning:
 - Please properly mention what kinds of help you get from the tools in your homework and project (missing acknowledgement will be treated as **cheating**)
 - Use the tools with responsibility:
 - ✓ Generative AI tools often **hallucinate** and only provide **general** answers

Homework Details (60%)

- ❑ HW1: Building MLP-based text classifier with pytorch
 - 5 points, Individual, due: **Sep 13**
- ❑ HW2: Finetuning text classifier using HuggingFace
 - 10 points, Individual, due: **Sep 20**
- ❑ HW3: Authorship attribution using language models (LMs)
 - 10 points, Individual, due: **Oct 4**
- ❑ HW4: Generating and evaluating text from pretrained LMs
 - 15 points, Team, due: **Oct 18**
- ❑ HW5: Prompting with large language models (LLMs)
 - 15 points, Team, due: **Nov 8**
- ❑ HW6: Essay writing with ChatGPT
 - 5 points, Individual, due: **Nov 22**



Homework Details (60%)

<https://dykang.github.io/classes/csci5541/F24/#homework>

Homework Details (60%)

All questions regarding homework **MUST be communicated with the lead TA over Slack homework channels** (e.g., #hw1, #hw2) or during their office hours. Homework 1, 2, 3, and 6 should be done individually, while homework 4 and 5 are team-based (maximum of 4 people). Your team for homework 4 and 5 should be the same for the project team. The use of outside resources (books, research papers, websites, etc.) or collaboration (students, professors, chatGPT, etc.) **must be explicitly acknowledged in your report**. Check out the [notes](#) for academic integrity.

The deadline for all homework is by **Friday midnight (11:59PM)** of the due date. Due to a tight schedule, there will be no deadline extension, but you can still use your late days. For the delayed team homework, late days for every team member will be counted. Check out the homework description  and link to canvas  for submission:

Here are homework assignments with dues:

- HW1: Building MLP-based text classifier with pytorch (5 points, [Individual](#), due: **Sep 13 Friday**)  
- HW2: Finetuning text classifier using HuggingFace (10 points, [Individual](#), due: **Sep 20 Friday**)  
- HW3: Authorship attribution using language models (LMs) (10 points, [Individual](#), due: **Oct 4 Friday**)  
- HW4: Generating and evaluating text generated from pretrained LMs (15 points, [Team](#), due: **Oct 18 Friday**)  
- HW5: Prompting with large language models (LLMs) (15 points, [Team](#), due: **Nov 8 Friday**)  
- HW6: Essay writing with ChatGPT (5 points, [Individual](#), due: **Nov 22 Friday**)  

HW1: Building text classifier with pytorch

```
from torch import nn

class MLP(nn.Module):
    def __init__(self, vocab_size, embed_dim, num_class):
        super().__init__()
        self.embedding = nn.EmbeddingBag(vocab_size, embed_dim, sparse=True)
        num_layer = 100
        # self.fc = nn.Linear(embed_dim, num_class)
        self.fc1 = nn.Linear(embed_dim, num_layer)
        self.fc2 = nn.Linear(num_layer, num_class)

        # initialize the weights
        self.init_weights()

    def init_weights(self):
        initrange = 0.5
        self.embedding.weight.data.uniform_(-initrange, initrange)
        self.fc1.weight.data.uniform_(-initrange, initrange)
        self.fc1.bias.data.zero_()
        self.fc2.weight.data.uniform_(-initrange, initrange)
        self.fc2.bias.data.zero_()

    def forward(self, text, offsets):
        embedded = self.embedding(text, offsets)
        return self.fc2(self.fc1(embedded))
```

```
| end of epoch 1 | time: 6.61s | valid accuracy 0.714
-----
| end of epoch 2 | time: 6.69s | valid accuracy 0.672
-----
| end of epoch 3 | time: 6.86s | valid accuracy 0.806
-----
| end of epoch 4 | time: 6.69s | valid accuracy 0.829
-----
| end of epoch 5 | time: 6.60s | valid accuracy 0.834
-----
| end of epoch 6 | time: 6.68s | valid accuracy 0.809
-----
| end of epoch 7 | time: 6.79s | valid accuracy 0.843
-----
| end of epoch 8 | time: 6.81s | valid accuracy 0.850
-----
| end of epoch 9 | time: 6.61s | valid accuracy 0.844
-----
| end of epoch 10 | time: 6.61s | valid accuracy 0.844
```

Does my classifier learn
something?

HW2: Finetuning text classifier using HuggingFace

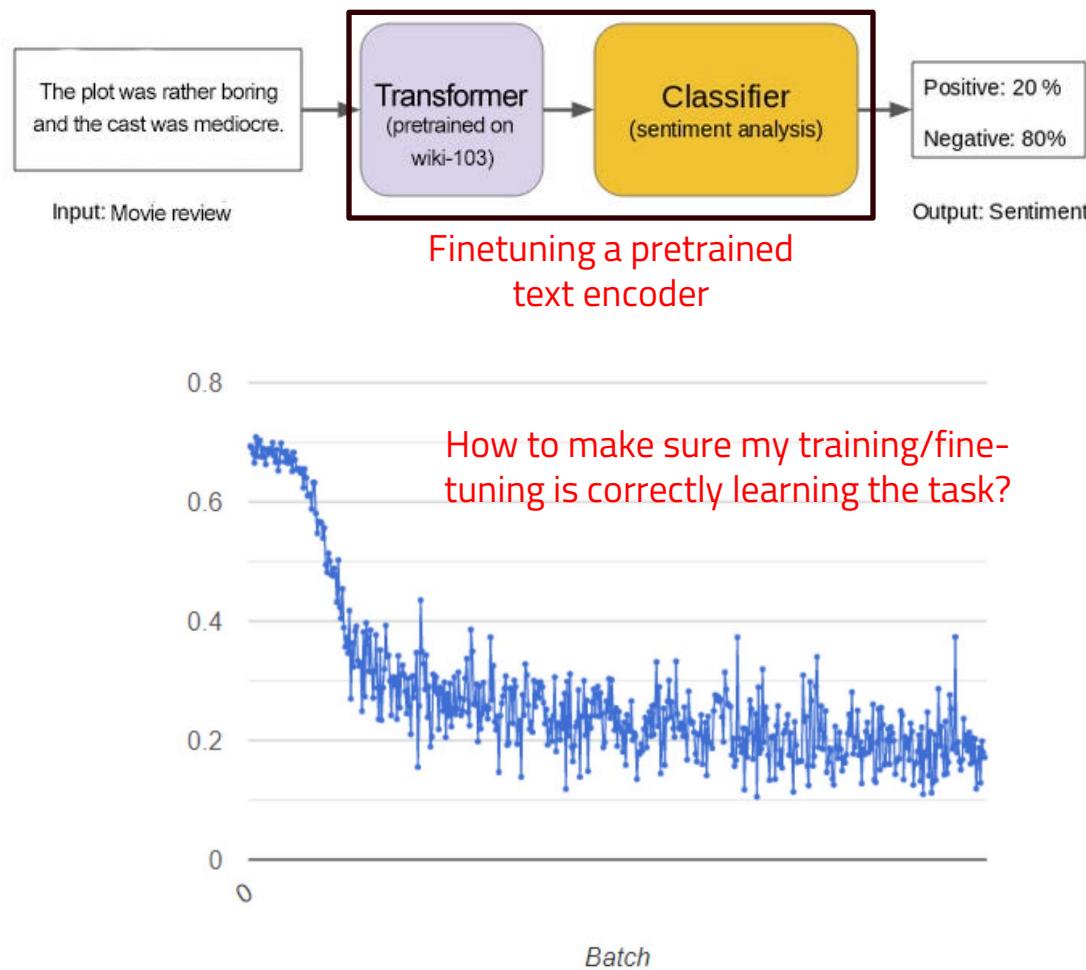
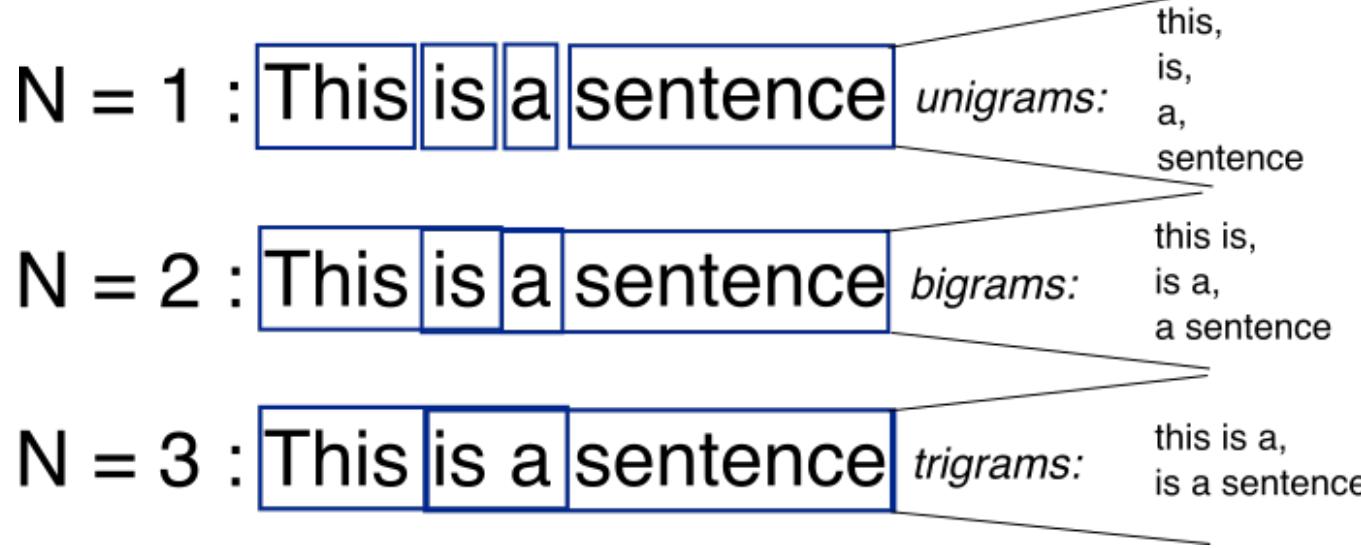


Table 1: Samples the model struggles with.

Sample	Ground-Truth Label	Predicted Label	Confidence
"This shit is fire"	Positive	Negative	0.870
"I love when my thoughts are definitely valued"	Negative	Positive	0.995
"Abso-fucking-lutely"	Positive	Negative	0.931
"Poggers"	Positive	Negative	0.895
"You are my favorite person to never see"	Negative	Positive	0.682

In which cases, does the fine-tuned classifier make mistakes?

HW3: Authorship attribution using language models



<s> i thought recalling the entrenchment and in the flank
had never thought nothing

Written by Tolstoy? Or Written by Hamlet?

Tolstoy Samples

<s> i thought recalling the entrenchment and in the flank
had never thought nothing
perplexity : 499.5855375522907

<s> i burn my regiment and me five oclock then accept
for the innumerable
perplexity : 1072.2662726616034

How plausible is text as Tolstoy language?

HW4: Generating and evaluating text from pretrained LMs

Los Angeles Times

WORLD & NATION

Palestinian Authority joins International Criminal Court as 123rd member



Palestinian Authority President Mahmoud Abbas speaks during a meeting of the Fatah revolutionary council in the West Bank city of Ramallah on Oct. 18, 2014. (Majdi Mohammed / Associated Press)

Ground Truth Summarization

Membership gives the ICC jurisdiction over alleged crimes committed in Palestinian territories since last June . Israel and the United States opposed the move, which could open the door to war crimes investigations against Israelis .

How to evaluate the generated text against the human's summary?

Algorithm	Meteor AVG
Greedy	37.12
Beam	38.33
Top-K	38.64
Top-P	38.14

Greedy Search

The Palestinian Authority becomes the 123rd member of the International Criminal Court. The move gives the court jurisdiction over alleged crimes in Palestinian territories. Israel and the United States opposed the Palestinians' efforts to join the body. But Palestinian Foreign Minister Riad al-Malki says it is a move toward greater justice.

Beam Search

Palestinian Authority becomes 123rd member of the International Criminal Court. The move gives the court jurisdiction over alleged crimes in Palestinian territories. Israel and the United States opposed the Palestinians' efforts to join the body. But Palestinian Foreign Minister Riad al-Malki said it was a move toward greater justice.

Top-K Sampling

The Palestinian Authority becomes the 123rd member of the International Criminal Court. The move gives the court jurisdiction over alleged crimes in Palestinian territories. Israel and the United States opposed the Palestinians' efforts to join the body. But Palestinian Foreign Minister Riad al-Malki said it was a move toward greater justice.

Different versions of generated text/summaries using various decoding algorithms

Other aspects humans evaluate text summaries?

Average Coherence (out of 5)	Average Factuality (out of 5)
4.52	4.20
4.70	4.28
4.57	4.25
4.33	4.20

HW5: Prompting with large LMs



Q: Minneapolis and Saint Paul are known as the Twin Cities in Minnesota. What's the capital?

A: Saint Paul.

Q: New York is a large city in New York state. What's the capital?

A: Albany.

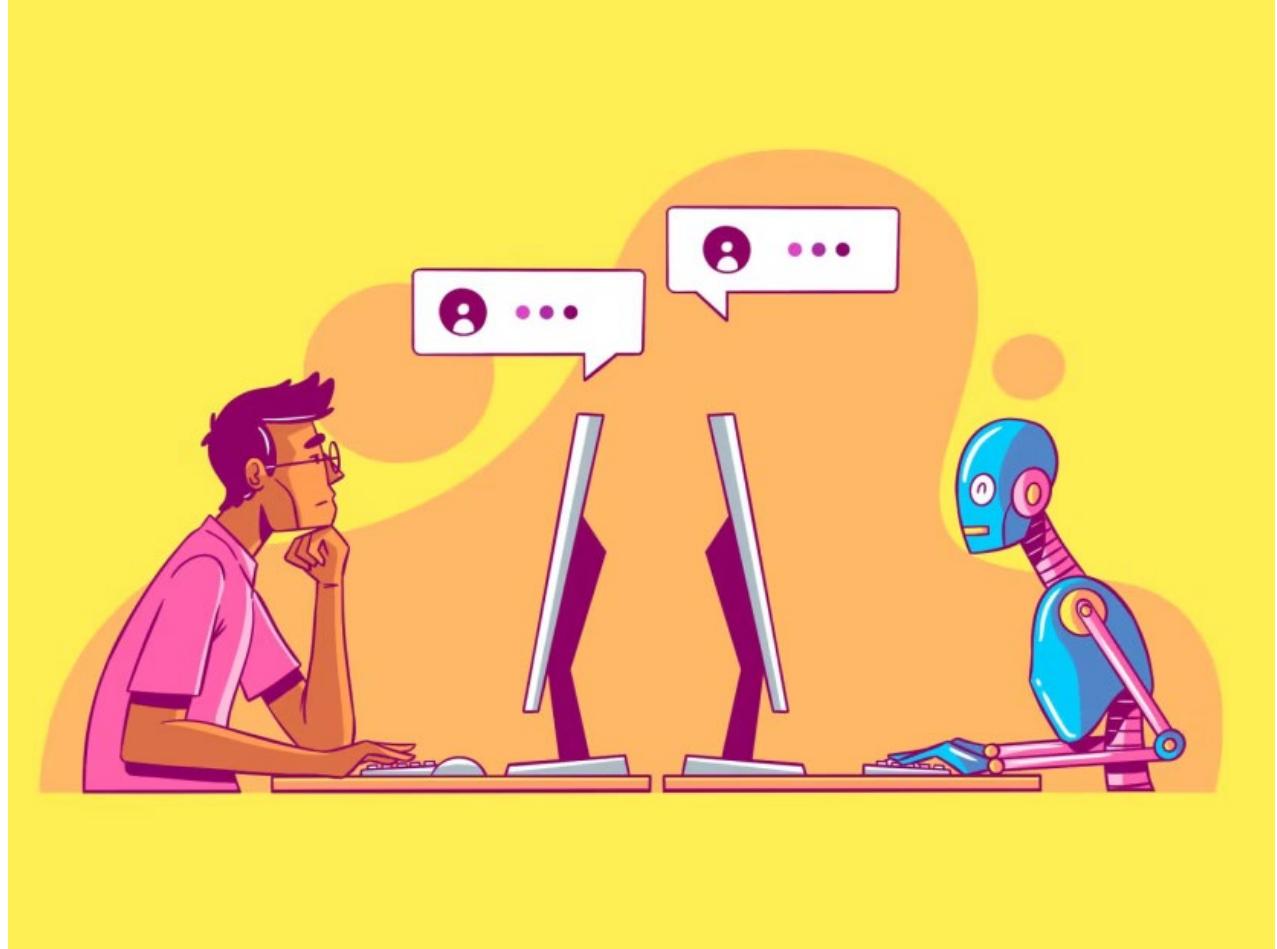
Q: Kansas City is another large city. Kansas City is the capital of which state?

A: Kansas ✗

Task Description	Instructions	Number of examples	Examples	Input Task	Expected Answer	Predicted Answer by GPT3
Convert movie titles into emoji.	N/A	0	N/A	Convert movies with emojis: The Hunger Games	🦁🔥💀	🔥🍁🔥😊
Convert movie titles into emoji.	N/A	1	This is an example of converting a movie title into an emoji. Harry Potter and the Sorcerer's Stone: 🧙‍♂️🧙‍♀️🔮	Convert movies with emojis based on the example: Harry Potter and the Goblet of Fire	🧙‍♂️🔥🏆	🧙‍♂️🔥🏆💡☀️
Convert movie titles into emoji.	N/A	0	N/A	Convert movie titles into emoji: Full River Red	🏞️🩸	🏔️💧🏔️🩸
Convert movie titles into emoji.	N/A	1	Full River Red 🏞️🩸	Convert movie titles into emoji: From Vegas to Macau	**) !	**) !

HW6: Essay writing with ChatGPT

Q2 [Education]. How would the classroom look in the future? Do you believe AI will impact and reshape the classroom and education? If you can learn most of what you need from chatGPT or other LLMs, what is the role of teachers? What do you expect students to learn in the classroom?



Project Goal

- A course project would be one of the following types:
 - New research results judged suitable for acceptance to a NLP or ML workshop,
 - Critical analysis of existing NLP models or datasets,
 - Collection of your own dataset on new problems or adversarial datasets that can fool the existing systems ,
 - An in-depth literature survey on emerging topics,
 - Interactive demonstration (e.g., Chrome Extension, Flask) or visualization of existing systems,
 - New open-source repository or dataset with a high impact on the community
- Please carefully read the project guideline first (will be posted on Sep 17)

Project Evaluation Rubric

- ❑ HWs are generously graded but the **projects are not!** Therefore, students should consider the potential contribution of the projects rather than **trying to play it safe**. Playing it safe won't give them full marks.
- ❑ Three important rubrics:
 - **Novelty:** Compared to the state-of-the-art methods/systems/datasets, how novel is your approach?
 - **Significance:** How strong is your result? Is your finding still holding if different setups or prompting tricks?
 - **Clarity:** How clear and easy-to-follow is your report? Do you have well organized presentation of your results and problem definition?
 - <https://dykang.github.io/classes/csci5541/F24/rubrick.html>

Project Deliveries and Dues (30%)

- ❑ Team formation (1 point, due: **Sep 19**)
- ❑ Project brainstorming (1 point, due: **Oct 1**)
- ❑ Proposal pitch (3 points, due: **Oct 8 and 10**)
- ❑ Proposal report (5 points, due: **Oct 13**)
- ❑ Midterm office hour participation (5 points, due: **Oct 31**)
- ❑ Poster presentation (5 points, due: **Dec 3 and 5**)
- ❑ Final report (10 points, due: **Dec 12**)



Project Details (30%)

First, **carefully read the project description**  as most project information, dues, rubric, and answers to your questions are in the description document. It is your responsibility to miss any information regarding the project. Your team (maximum of 4 people) should submit their report, link to code (or a zipped code), and presentation slides/poster to Canvas before the deadline. Use official ACL style templates ([Overleaf](#) or [links](#)). Here are some dues you have to submit for project (note that some dues are during week days):

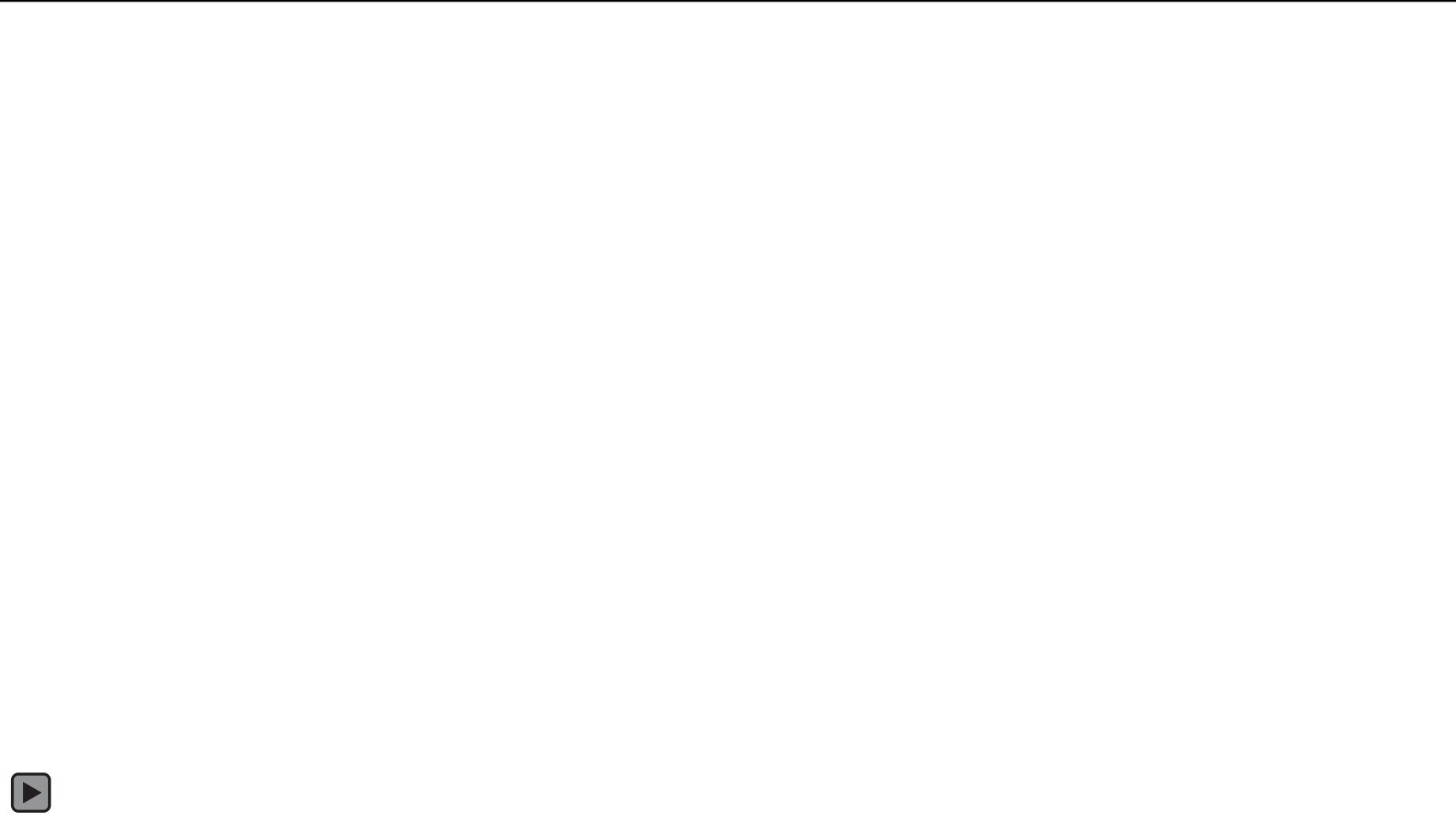
- **Team formation** (1 point, due: Sep 19) 
- **Project brainstorming** (1 point, due: Oct 1) 
- **Proposal pitch** (3 points, due: Oct 8 and 10) (Slides decks for Group A and Group B)
- **Proposal report** (5 points, due: Oct 13) 
- **Midterm office hour participation** (5 points, due: Oct 31) 
- **Poster presentation** (5 points, due: Dec 3 and 5) 
- **Final report** (10 points, due: Dec 12)  (evaluation rubric)

You can find some selected project reports and posters from the previous years' NLP classes below. Some projects are extended and published top-tier workshop and conferences:

- [CSCI 5541 S23] *Simulating Everyone's Voice: Exploring ChatGPTs Ability to Simulate Human Annotators*  
- [CSCI 5541 S23] *Vision & Language-guided Generalized Object Grasping*  
- [CSCI 5541 S23] *Generalizability of FLAN-T5 Model Using Composite Task Prompting*  
- [CSCI 5541 S23] *Comparing the Effectiveness of Fine-tuning vs. One-Shot Learning on the Kidz Bopification Task*  
- [CSCI 5980 F22] *Generating Controllable Long-dialogue with Coherence*  → Published in AAAI 2024 
- [CSCI 8980 S22] *Understanding Narrative Transportation in Fantasy Fanfiction*  → Published in Workshop on Narrative Understanding (WNU) @ACL 2023 

VLanGOGh: Vision and Language guided Generalized Object Grasping

CSCI 5541 Spring 2023
Nikhilanj Pelluri



Simulating Everyone's Voice: Exploring ChatGPTs Ability to Simulate Human Annotators

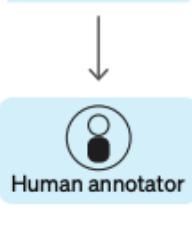
CSCI 5541 Spring 2023

Abdirizak Yussuf, Claire Chen, Dinesh Challa, Venkata Sai Krishna

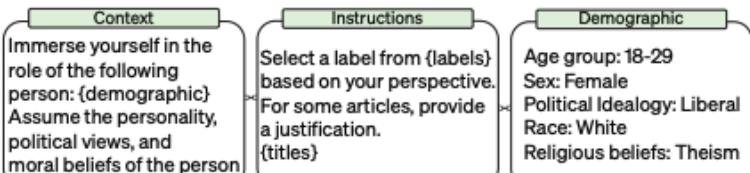
Step 1
Scraping and filtering data.



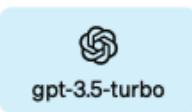
Step 2
Human annotation.
Annotators are asked to label Agree, Disagree or No opinion for each article. For 10 articles, they also provide a justification.



Step 3
ChatGPT annotation.
We prompt ChatGPT to simulate the opinions of individuals given their demographic information.



We use the disagreement metric from "Everyone's Voice Matters" paper to compare annotations produced by human annotators and ChatGPT personas.



Topic	Human Annotators	ChatGPT Personas
Abortion	0.22	0.32
Immigration	0.15	0.40
Social Issues	0.11	0.40
Political Issues	0.017	0.50
Racial Justice	0.19	0.40
Religion	0.18	0.36
All Topics Combined	0.15	0.42

- **Human annotators:** 0.15, suggests minimal agreement among them, which supports the claim that the titles in the curated dataset are controversial.
- **ChatGPT personas:** 0.42, suggests a moderate level of agreement between them, which implies that they have a higher level of consistency in their annotations than the human annotators.

Who is speaking? Distinguishing Artificial Intelligence Generated and Human Written Text

CSCI 5541 Spring 2023

Moyan Zhou, Mingsheng Sun, Yutong Sun

How Robotic is Your Content?

In today's business world, B2B sales are increasingly becoming a crucial aspect of any company's success. The business-to-business (B2B) market involves selling products or services to other businesses, rather than to individual consumers. Whether you are just starting out in B2B sales or are looking to improve your current approach, this blog post will provide you with strategies that can help you succeed in the B2B market.

Know Your Target Market

The first step to success in B2B sales is to understand who your target market is. Who are the businesses that you want to sell to? What are their needs and wants? What are the pain points that your products or services can solve? The better you understand your target market, the easier it will be for you to create a sales strategy that appeals to them.

Build Relationships

Building strong relationships with your prospects and Up to 25,000 characters will be used. 1666 Characters

Want Undetectable AI Content?

Our proprietary content platform uses a mix of 3 AI engines, NLP and semantic analysis algorithms, crawls Google, and parses all the top ranking content to put research-backed, long-form, SEO driven blog posts together.

This isn't an AI writing assistant, this is a human level, long-form, blog post producing machine!

Request an Invite

Check For AI Content

RQ1: Do people **agree with each other** when distinguishing AI-generated and Human-written text?

Fleiss' Kappa
0.05 (p-value = 0.017)

RQ3: How does the existing tools work?

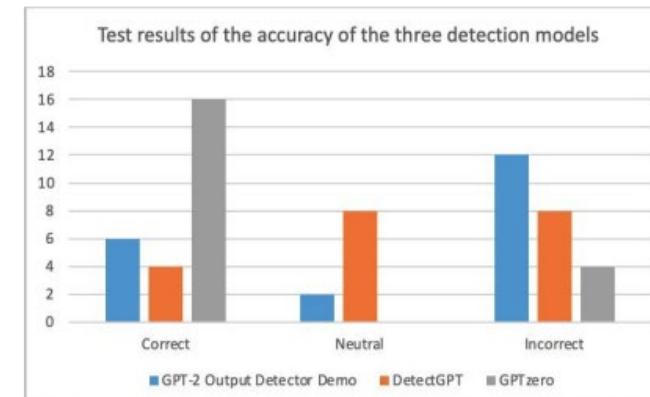


Figure 5. Test results of the accuracy of the three detection models

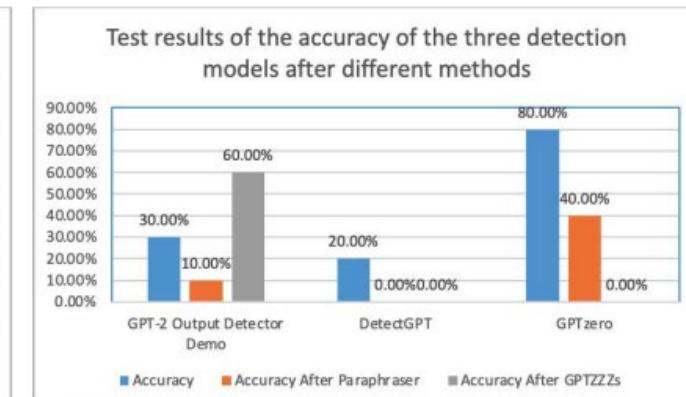


Figure 6. Test results of the accuracy after different methods

Comparing the Effectiveness of Fine-tuning vs. One-Shot Learning on the Kidz Bopification Task

CSCI 5541 Spring 2023

Jeonghoon Kim, Matthew Olson, Marco Berriodi

“I buy a new car for the b**** (For real)
I tear down the mall with the b**** (For real)”
from Bank Account by 21 Savage

“Lean with me, pop with me
Get high with me if you rock with me
Smoke with me, drink with me
F***** up liver with some bad kidneys”
from Lean Wit Me by Juice WRLD



You are a helpful assistant who is tasked with cleaning up songs. You will remove explicit material from a song's lyrics.

So we'll piss off the neighbors
In the place that feels the tears
The place to lose your fears
Yeah, reckless behavior
A place that is so pure, so dirty and raw
In the bed all day, bed all day, bed all day
Fucking and fighting on
It's our paradise and it's our war zone
It's our paradise and it's our war zone



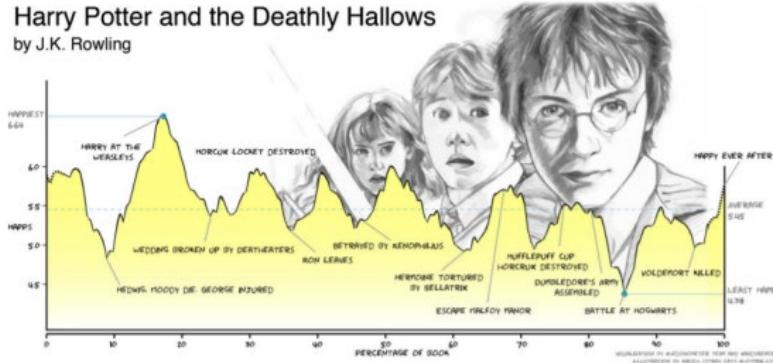
Here's a cleaned-up version of the lyrics:

We'll tick off the neighbors
In the place that feels the tears
The place to lose your fears
Yeah, wild behavior
A place that is so real, so gritty and raw
In the bed all day, bed all day, bed all day
Arguing and laughing on
It's our haven and it's our battleground
It's our haven and it's our battleground.



Understanding Narrative Transportation in Fantasy Fanfiction

CSCI 8980 Spring 2022
Kelsey Neis, Yu Fang



Select and highlight phrases that increase your sense of presence or connection to the story (blue) or decrease it (red).

He'd been dreaming of it since the defeat of Voldemort. The Veil, that is. The one that Sirius had fallen behind. The last time he'd dreamed about the Department of Mysteries, Sirius had died. The world had also finally woken up to the truth about Voldemort, but the price had been too high for him to be grateful. And once the truth was out, the attacks had gotten worse.

He wondered, briefly, why he was there. Ginny was at home, she was pregnant, she needed him. Yet he could not seem to stay away. Voldemort was gone, true. But then, so were so many other people, good people, who should not have died. He counted the steps until he was standing right in front of where Sirius had fallen through.

Cedric. Sirius. Dumbledore. Hedwig. Moody. Dobby. Tonks father. Remus. Colin Creevy. Tonks. Snape. Fred. Hell, even Crabbe didn't deserve to die then. There were more, many more, but none of them close to him.

Figure 1: Sample emotion arc for Harry Potter. (Reagan et al., 2016)

That was why he found himself, during his latest bout with insomnia, browsing a lesser known hero forum.	2.520
He'd found it several years earlier and quickly figured out that a lot of underground heroes used it to communicate with each other, since it offered encrypted chats and accounts were only known by random numbers, rather than usernames.	1.890
He'd spent about two weeks back then figuring out which accounts corresponded to which heroes, but he had never posted himself.	3.950
As Izuku drowsily scrolled through old posts, a crazy idea occurred to him.	4.580
If he couldn't be a hero himself, why couldn't he help the real heroes be better?	5.580
In the morning, he'd blame it on sleep deprivation and then promptly die of mortification, but that didn't change the fact that, at two o'clock in the morning, Izuku Midoria sent ten underground heroes in depth analyses of their quirks and fighting styles.	7.070

Published in Workshop on Narrative Understanding (WNU) @ACL 2023 <https://arxiv.org/abs/2306.04043>

Transidiomation: Optimizing translation of idioms embedded in text

CSCI 5541 Spring 2024

Dhondup Dolma, Jaeeun Lee, Yongtian Ou, Jiyoon Pyo Team

NLPitch

English sentences with PIE	GPT 3.5	Google Translate	SeamlessM4T
It was raining cats and dogs last weekend.	지난 주말에는 개구리가 비 내리는 것 같았어요.	지난 주말에는 고양이와 개에게 비가 내렸습니다.	지난 주말에는 고양이와 개가 비를 내고 있었다.
It was like frogs were raining last weekend. (hallucination)	Cats and dogs were rained on last weekend. (incorrect)	Cats and dogs were rained on last weekend. (incorrect)	Not translatable
Getting that done was a piece of cake.	그것을 해내는 것은 쉬운 일이었다.	그 일을 끝내는 건 아주 쉬운 일이었어	그 일을 처리하는 것은 쉬운 일이었습니다.
	It was easy to get that done.	It was easy to get that done.	It was easy to get that done.
Teaching is my bread and butter.	가르치는 것이 내 소득원이다.	가르치는 것이 나의 빵과 버터이다.	가르치는 게 내 과 버터야.
	Teaching is my source of income.	Teaching is my bread and butter. (literal translation)	Not translatable

Table 1: Examples of various NMT models translating English sentences with PIE to Korean. Text that have red Xs are incorrect translations.

“Translate the sentence into target language by following these steps:

Step1. Identify the idiom.

Step2. Find an idiom with the same meaning in the target language. If there is no equivalent idiom, give the figurative meaning of it.

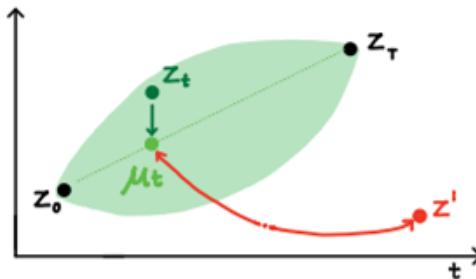
Step3. Include response from Step2 to translate the sentence.”

Generating Controllable Long-dialogue with Coherence

CSCI 5980 Fall 2022

Zhecheng Sheng, Chen Jiang and Tianhao Zhang

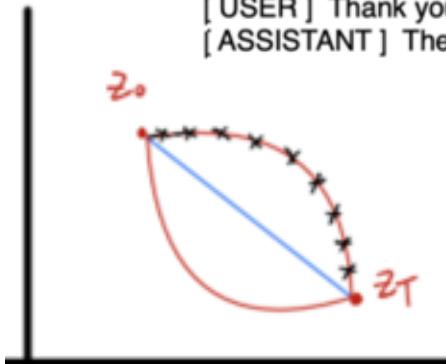
Time control in language model using Brownian bridge (Wang et al., ICLR 2022)



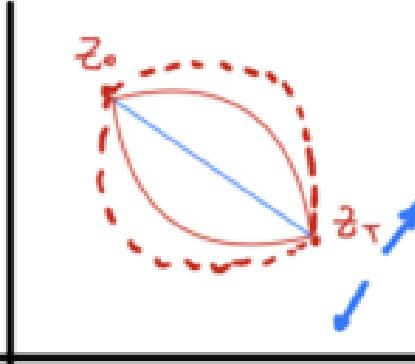
x_0 : [USER] Hello, I'd like to buy tickets for tomorrow.
 x_t : [ASSISTANT] What movie theater do you prefer?
 x_T : [USER] Could you confirm my tickets just in case?
 x' : [USER] Hi, I'm looking to purchase tickets for my family.

$$\mathcal{L} = -\log \frac{\exp(d(z_t, \mu_t))}{\exp(d(z_t, \mu_t)) + \exp(d(z', \mu_t))}$$

[USER] I am thinking about seeing a movie tonight, please.
[ASSISTANT] What movie do you have in mind?
[USER] The Elizabeth Theatres.
[ASSISTANT] They have 2 tickets available.
[USER] Thank you.
[ASSISTANT] They have been waiting in the line for you.



[USER] I would love to get to the movies tonight.
[ASSISTANT] OK. And where will you be seeing the movie?
[USER] Creek's End, Oregon.
[ASSISTANT] Creek's End, Oregon. Got it. Is there a particular movie you have in mind?
[USER] No wait, the visuals are so darned.
[ASSISTANT] No problem.
[USER] No problem.
[ASSISTANT] No problem.
[USER] No problem.
[ASSISTANT] No problem.
[ASSISTANT] No problem.



Published in AAAI 2024, <https://arxiv.org/abs/2312.16893>

Topics

Schedule

We will cover basic NLP representations $g(x)$, to build text classifiers $P_{\theta}(y|g(x))$, language models $P_{\theta}(g(x))$, and large language models $P_{\{\theta \text{ is large}\}}(g(x))$. Based on knowledge you gain during the class, your team will develop your own NLP systems during the semester-long project. Pay attention to due dates and homework release. Lecture slides and homework/project description will be available in  PDF.

Date	Lectures and Dues	Readings
Sep 3	Class Overview 	
Sep 5	Intro to NLP  HW1 out 	



Topics

□ Text classification

- Programming Tutorials
- Finetuning Tutorial

□ Basic representations

- Distributional semantics
- Contextualized word embeddings
- Self-Attentions and Transformers

Sep 10	Text Classification PDF Tutorial on building a classifier with Scikit-Learn P and PyTorch P (Shirley)	<ul style="list-style-type: none">• Determining the sentiment of opinions• From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series• Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank• Text classifier with NLTK and Scikit-Learn
Sep 12	Text Classification with Finetuning PDF Tutorial on Finetuning PDF (James) HW1 due HW2 out PDF	<ul style="list-style-type: none">• Does BERT Learn as Humans Perceive? Understanding Linguistic Styles through Lexica• Style is NOT a single variable: Case Studies for Cross-Style Language Understanding• Beyond Accuracy: Behavioral Testing of NLP Models with CheckList• Blog post on Pre-training vs Fine-tuning in LLM: Examples• Tutorial on Text classification using HuggingFace's Transformers
Sep 17	Distributional Semantics and Word Vectors PDF Project description out PDF	<ul style="list-style-type: none">• From Frequency to Meaning: Vector Space Models of Semantics• Efficient Estimation of Word Representations in Vector Space• Linguistic Regularities in Continuous Space Word Representations• GloVe: Global Vectors for Word Representation• Retrofitting Word Vectors to Semantic Lexicons• Gensim's word2vec tutorial
Oct 15	Contextualized Word Embeddings PDF	<ul style="list-style-type: none">• Deep contextualized word representations• BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding• A Primer in BERTology: What we know about how BERT works
Oct 17	Transformers HW4 due	<ul style="list-style-type: none">• Attention is All you Need• Tutorial on Illustrated Transformer• Language Models are Unsupervised Multitask Learners• Language Models are Few-Shot Learners• Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer• Language Models are Few-Shot Learners• Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer

Topics (cont'd)

□ Language Models

- Ngram LM, Neural LM, RNN, LSTM, and seq2seq
- Search and decoding
- Search in Training
- Evaluation
- Applications

Sep 19	Language Models (1): Ngram LM, Neural LM PDF Project Team Formation due HW2 due HW3 out PDF	<ul style="list-style-type: none">• Chapter 3 of Jurafsky and Martin• A Neural Probabilistic Language Model
Sep 24	Project Guideline PDF P	
Sep 26	Language Models (2): RNNs, LSTMs and Sequence-to-Sequence PDF	<ul style="list-style-type: none">• Recurrent neural network based language model• Long Short-Term Memory• Multivariable chain rule, simple version• Long Short-Term Memory• Sequence to Sequence Learning with Neural Networks
Oct 1	Language Models (3): Search and Decoding PDF Project brainstorming due	<ul style="list-style-type: none">• The Curious Case of Neural Text Degeneration• Mutual Information and Diverse Decoding Improve Neural Machine Translation• Sequence Level Training with Recurrent Neural Networks• An Actor-Critic Algorithm for Sequence Prediction• Training language models to follow instructions with human feedback
Oct 3	Language Models (4): Evaluation and Applications PDF HW3 due HW4 out PDF	<ul style="list-style-type: none">• Perplexity of fixed-length models• BLEU: a Method for Automatic Evaluation of Machine Translation• ROUGE: A Package for Automatic Evaluation of Summaries

Topics (cont'd)

❑ Large language models:

- Pretraining and scaling laws
- Prompting
- Instructing and augmenting LLMs
- Multi-agents, agentic workflow
- Alignment and RLHF

❑ Efficient NLP

❑ Ethics and safety

❑ Societal alignment or Multimodal LM

❑ Data, annotation, and evaluation

❑ Concluding remarks

Oct 22	Pretraining and Scaling Laws PDF HW5 out PDF	<ul style="list-style-type: none">• Scaling Laws for Neural Language Models• On the Opportunities and Risks of Foundation Models• On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?
Oct 24	Prompting PDF	<ul style="list-style-type: none">• Chain-of-Thought Prompting Elicits Reasoning in Large Language Models• Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing• Prefix-Tuning: Optimizing Continuous Prompts for Generation
Oct 29	Instructing and augmenting LLMs PDF	<ul style="list-style-type: none">• Training language models to follow instructions with human feedback• Augmented Language Models: a Survey• Toolformer: Language Models Can Teach Themselves to Use Tools• Internet-augmented language models through few-shot prompting for open-domain question answering
Oct 31	Multi-agents, Agentic Workflow (Robert) PDF Project midterm office-hour due	TBD
Nov 5	Ethics and Safety (Shirley) PDF	<ul style="list-style-type: none">• The Ethics of Artificial Intelligence• On Calibration of Modern Neural Networks• Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings• Calibrate Before Use: Improving Few-Shot Performance of Language Models• "Why Should I Trust You?": Explaining the Predictions of Any Classifier• Differential Privacy• Case Studies in Princeton Dialogues on AI and Ethics
Nov 7	Alignment and RLHF PDF HW5 due HW6 out PDF	TBD
Nov 14	Efficiency NLP PDF (James)	<ul style="list-style-type: none">• Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding• LLM.int8(): 8-bit Matrix Multiplication for Transformers at Scale• Outrageously Large Neural Networks: The Sparsely-Gated Mixture-of-Experts Layer• The Power of Scale for Parameter-Efficient Prompt Tuning
Nov 19	Societal Alignment of LLMs PDF	TBD
Nov 21	All about Data and Annotation PDF HW6 due	<ul style="list-style-type: none">• Annotation Artifacts in Natural Language Inference Data• Dataset Cartography: Mapping and Diagnosing Datasets with Training Dynamics• Everyone's Voice Matters: Quantifying Annotation Disagreement Using Demographic Information• ChatGPT Outperforms Crowd-Workers for Text-Annotation Tasks



Topics (cont'd)

□ Project:

- Project guideline
- Project Proposal Pitch (Group A and B)
- Final Project Poster (Group B and A)

Sep 24	Project Guideline PDF P	
Oct 8	Project Proposal Pitch (1) Slides Deck for Group A	Group A:
Oct 10	Project Proposal Pitch (2) Slides Deck for Group B Project proposal due (Oct 13)	Group B:
Nov 28	No Class (Thanksgiving)	
Dec 3	Final Project Poster (1) Project poster due	Posters for Group B
Dec 5	Final Project Poster (2) Project final report due (Dec 12)	Posters for Group A



Textbook

- ❑ The lecture slides and reading list will be posted on the class site at least one day before the class.
- ❑ Textbook is **not required** but the following books are primarily referred
 - Jurafsky and Martin, Speech and Language Processing, 3rd edition [[online](#)]
 - These course materials are inspired by the slides from Stanford, UC Berkeley, and CMU.

Academic Integrity

- ❑ Everything you turn in must be your own work, and you must note **the names of anyone** you collaborated with on each problem and cite **any resources** (e.g., chatGPT prompts you used) that you used to learn about the problem.
- ❑ If you have any doubts about whether a particular action may be construed as cheating, ask the instructors for clarification before you do it. Cheating in this course will result in a grade of **F** for course and the [University policies](#) will be followed.

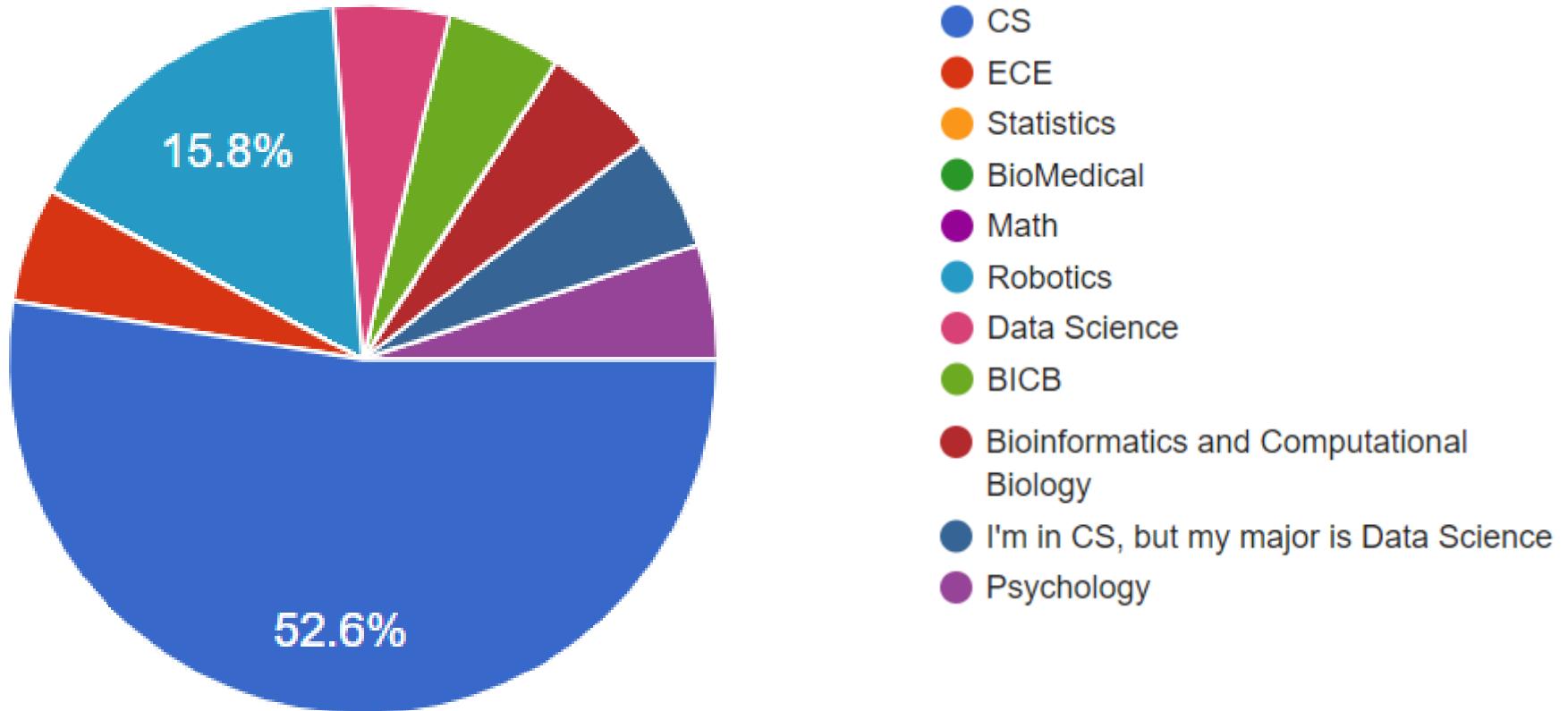


Computing Resources

- ❑ Reimbursement for Google Colab Pro by CS&E department
 - <https://docs.google.com/document/d/1Kdo6BN1h2yISYHTTSC6DVLWhNEaEcRxk/edit?usp=sharing&ouid=102305309222832534233&rtpof=true&sd=true>
- ❑ MSI Class Accounts:
 - <https://www.msi.umn.edu/content/applying-class-accounts>
 - Will be created soon

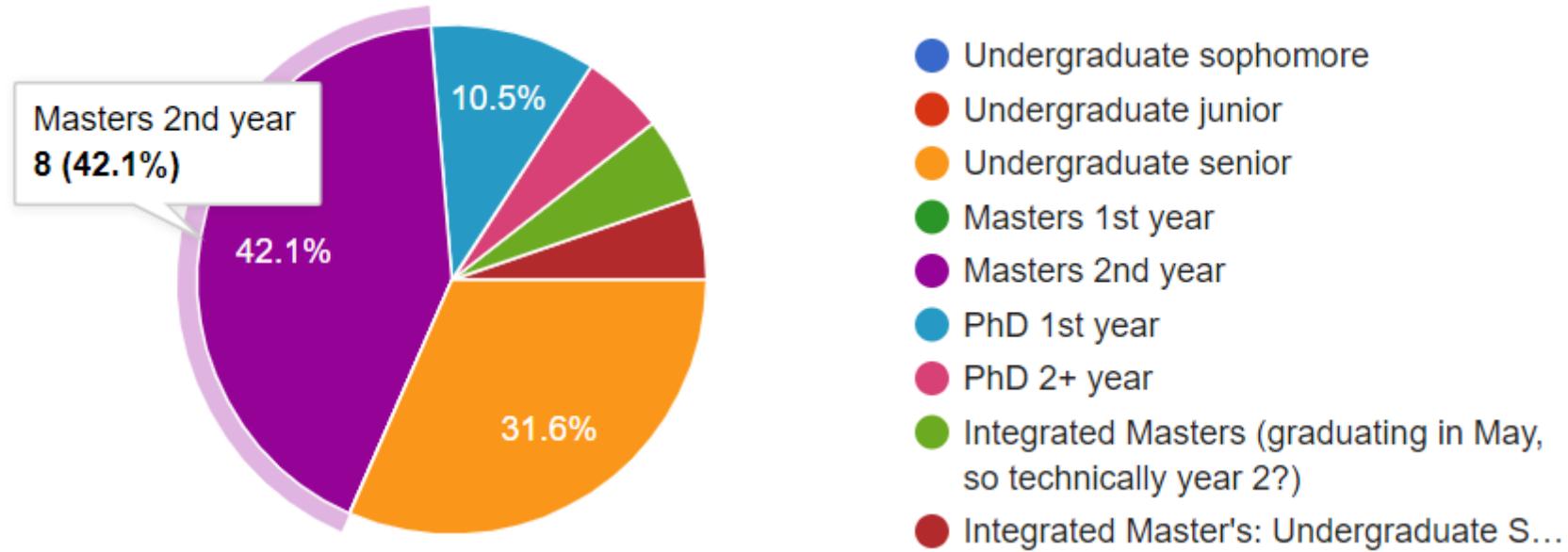
Pre-class Survey

Which department are you in? (19 responses)



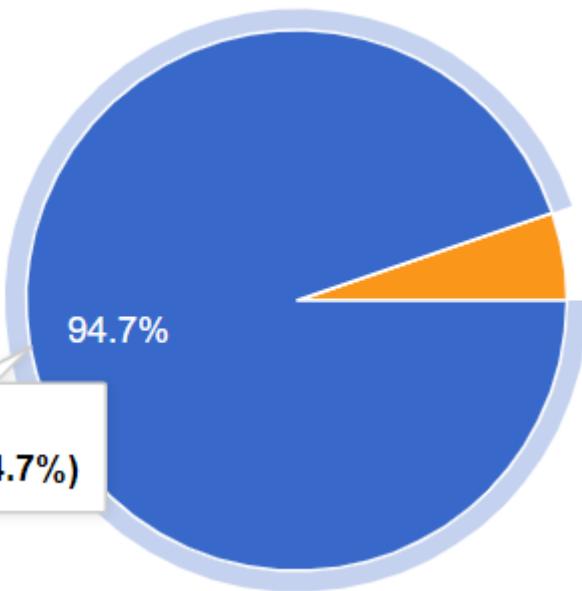
Pre-class Survey

Which program are you? (e.g., undergraduate senior, Master 2nd, Ph.D. 1st)

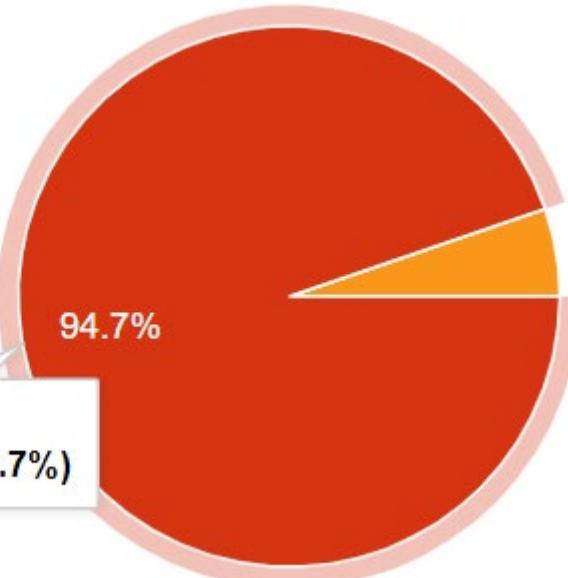


Pre-class Survey

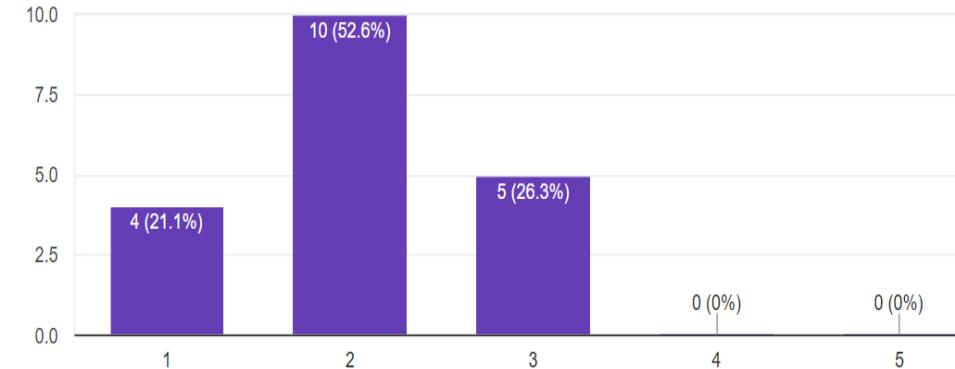
Have you ever taken any machine learning or deep learning courses before?



Have you ever taken any NLP courses before?



Can you rate your current knowledge about NLP from 1 to 5?



Pre-class Survey

What is the most interesting topic you like to learn from this class?

Have you had a chance to look at the current schedule in our class homepage?
(<https://dykang.github.io/classes/csci5541/F24/>) What is the most interesting topic you like to learn from this class?

19 responses

Transformers

LLMs and the math behind NLP

language models and transformers

Transformers

Yes, I'm excited for the week/lecture on prompting!

Machine Translation

Pre-class Survey

What is the most interesting topic you like to learn from this class?



You

Here are topics of interests from students. Can you summarize them by count?



ChatGPT

Here is a summary of the topics of interest from the students, along with the count of mentions:

1. Lexical Semantics: 3 mentions
2. Language Models: 8 mentions
3. BERT: 1 mention
4. COVID-Twitter-BERT: 1 mention
5. Deep Learning topics/Transformer/Attention: 2 mentions
6. Hugging Face Transformer: 2 mentions
7. PyTorch: 1 mention
8. Creating our own dataset: 1 mention
9. Mamba: 1 mention
10. Transformers in the context of NLP: 2 mentions
11. Introduction to LLMs: 2 mentions
12. AI-related concepts: 1 mention
13. Machine Translation: 1 mention
14. Probabilistic Language Model: 1 mention
15. Distributional Semantics: 1 mention
16. Computation efficiency: 1 mention
17. Text encoding: 1 mention
18. Encoder-Decoder: 1 mention

Note: Some topics may have variations in their naming, and I have consolidated similar topics under a common term (e.g., Language Models, LLMs).



Analyzed

```
python                                         Always s  
  
from collections import Counter  
  
# List of topics  
topics = [  
    "Transformers",  
    "LLMs and the math behind NLP",  
    "language models and transformers",  
    "Transformers",  
    "Yes, I'm excited for the week/lecture on prompting!",  
    "Machine Translation",  
    "I want to learn applications, like how can I apply NLP in o  
    "Large Language Models",  
    "I think prompt engineering will be the most interesting.",  
    "I glanced at it but not in depth.",  
    "Yes, I'm interested on the basics of language models and ho  
    "I'm really interested in learning more about Prompting. Con  
    "I don't know much about NLP, but learning about the applica  
    "Yes and Prompting",  
    "Language models",  
    "Language Models (2): RNNs, LSTMs and Sequence-to-Sequence",  
    "How to use NLP techniques for image captioning"  
]  
  
# Normalize topics by keywords
```

```
normalized_topics = []  
for topic in topics:  
    topic_lower = topic.lower()  
    if "transformers" in topic_lower:  
        normalized_topics.append("Transformers")  
    elif "prompt" in topic_lower:  
        normalized_topics.append("Prompting")  
    elif "language models" in topic_lower or "llms" in topic_lower:  
        normalized_topics.append("Language Models")  
    elif "machine translation" in topic_lower:  
        normalized_topics.append("Machine Translation")  
    elif "applications" in topic_lower:  
        normalized_topics.append("Applications")  
    elif "image captioning" in topic_lower:  
        normalized_topics.append("Image Captioning")  
    else:  
        normalized_topics.append(topic)  
  
# Count occurrences of each topic  
topic_counts = Counter(normalized_topics)  
  
# Display the counts  
topic_counts
```

```
Result
Counter({'Language Models': 6,
          'Prompting': 4,
          'Transformers': 3,
          'Machine Translation': 1,
          'Applications': 1,
          'I glanced at it but not in depth.': 1,
          'Image Captioning': 1})
```

Oct 24	Prompting 
Oct 29	Instructing and augmenting LLMs 
Oct 31	Multi-agents, Agentic Workflow (Robert)  Project midterm office-hour due

Here is the summary of topics by count:

- Language Models: 6
- Prompting: 4
- Transformers: 3
- Machine Translation: 1
- Applications: 1
- Image Captioning: 1
- I glanced at it but not in depth.: 1

Language Models and Prompting are the most popular topics among the students. 

Pre-class Survey

What do you want to get out of this course? What is important thing you hope to learn from this course? you most excited about in this course?



The students' responses can be summarized into the following key themes:

1. **Interest in AI/ML and NLP Applications:** Students express a strong desire to learn about AI, machine learning, and NLP to contribute to an evolving technological landscape. They are particularly interested in understanding how these technologies can be applied in real-world scenarios, such as improving daily tasks and enhancing job prospects in the tech industry.
2. **Focus on Language Models (LLMs):** Many students are keen on gaining a deep understanding of large language models, including their architecture, functionality, and potential applications. Some are specifically interested in using LLMs for specialized tasks like robotics and high-level planning.
3. **Skill Development for Future Careers:** Several students are motivated by the desire to acquire practical skills that will be valuable in their future careers, particularly in tech-related fields. This includes learning foundational concepts, practical applications, and engaging in projects that demonstrate their acquired knowledge.
4. **Interest in Prompt Engineering:** There is a notable interest in learning about how language models respond to prompts, with students wanting to understand the underlying mechanics to create more effective prompts.
5. **Research and Academic Growth:** A few students are interested in building on their existing knowledge, particularly in machine learning, to deepen their understanding of NLP for future research and academic purposes.
6. **Course Engagement:** Some students, including those auditing the course, are eager to learn specific applications and foundational knowledge, indicating a strong commitment to personal and professional development through this course.

Pre-class Survey

(optional) What are you most worried about?



The students' responses can be summarized into the following concerns and challenges:

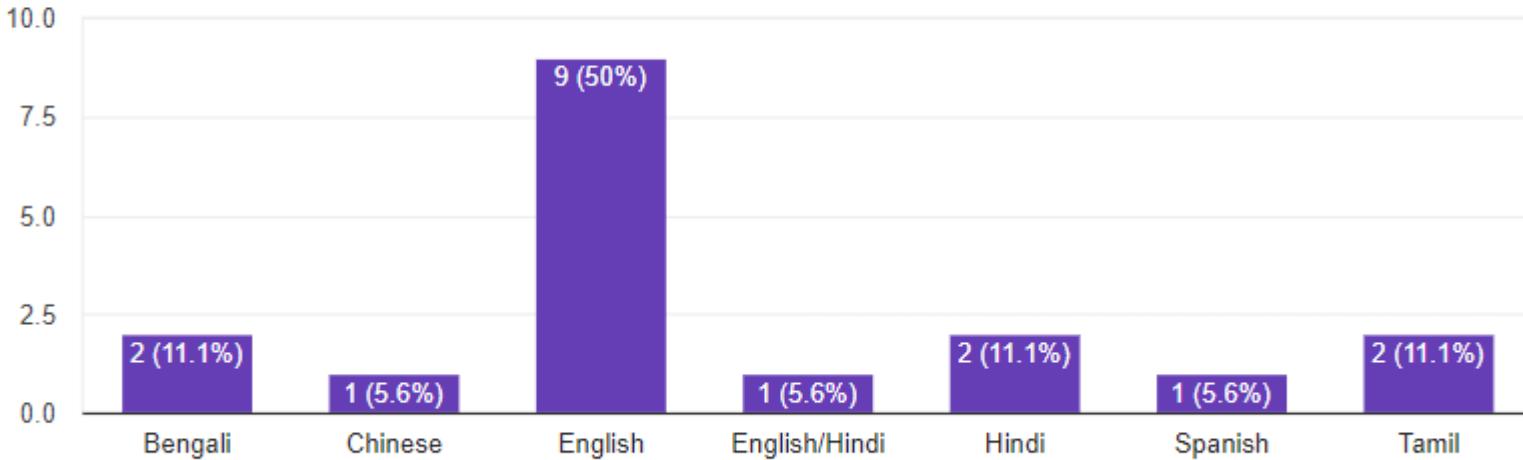
1. **Difficulty with Technical Content:** Some students are concerned about the complexity of NLP, particularly the theoretical aspects like the arithmetic and math behind language models, including transformers.
2. **Lack of Background Knowledge:** A few students are worried about not having sufficient coding experience or background knowledge in the subject matter, although they are optimistic about their ability to learn quickly.
3. **Balancing Commitments:** One student expressed concern about managing the demands of life, work, and study simultaneously.
4. **Group Project Concerns:** The large group project is a source of anxiety for at least one student, especially due to not knowing any classmates.
5. **Uncertainty:** A couple of students are unsure about what specific challenges they might face as the course progresses.

What's your mother tongue? (e.g., English, Chinese, Spanish)

18 responses

Copy

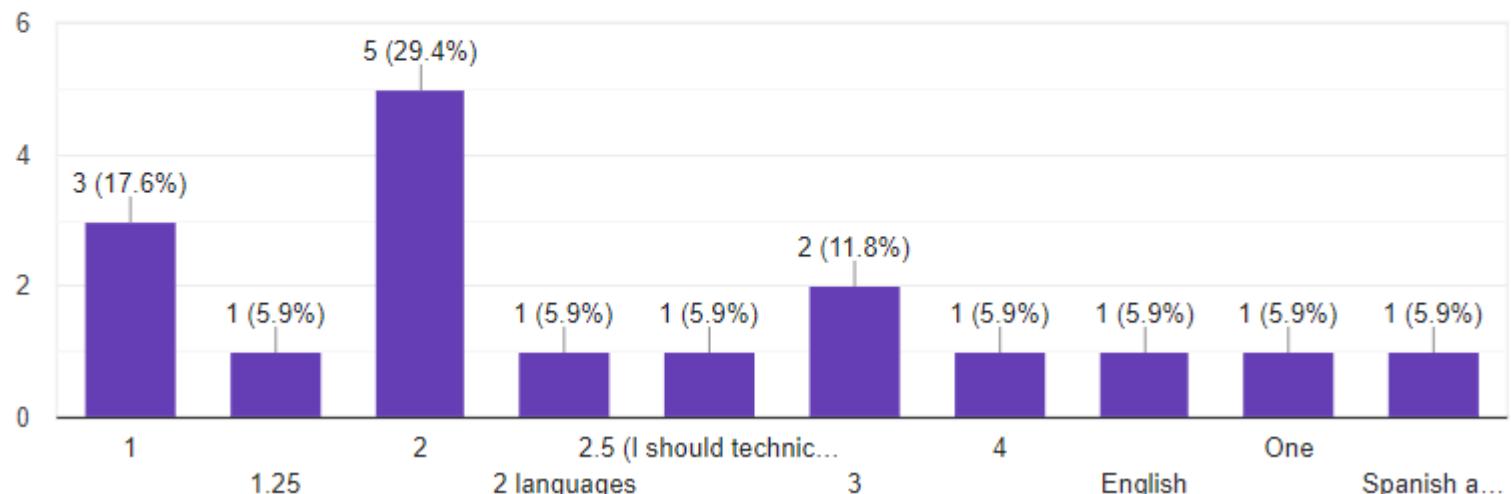
Copy chart



(optional) How many language can you speak?

17 responses

Copy



Waitlist & Recitation on computing basics

- ❑ If you plan to drop the course, please do it **by Today** so I can let some students on the waitlist into the class.
 - I will be teaching the same course in next semester, so please consider to take next one if you didn't make it this time
- ❑ Recitation on computing basics (next Monday at 3-4:30p by Robert):
 - Env Set-up
 - MSI Set-up
 - NN basics
 - Pytorch Basics (basic concepts, custom dataloader, torch.nn, training loop)
 - Recording will be available too

