Frontiers and Open-Challenges

CS330
Logistics

Today: This is the last lecture!

Tuesday (Nov 30th): Project poster session

Wednesday (Dec 8th): Project due
Today: Frontiers!
and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?

Making RL a better data sponge

How to scale up data collection?
How to utilize broad offline data sources?
How to apply RL to lifelong settings?

Open Challenges
Today: Frontiers!
and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?

Making RL a better data sponge

How to scale up data collection?
How to utilize broad offline data sources?
How to apply RL to lifelong settings?

Open Challenges
What is special about large language models?

Zero-shot generalization!

OpenAI’s new language generator GPT-3 is shockingly good—and completely mindless

The GPT-3 hype is way too much. It’s impressive (thanks for the nice compliments!) but it still has serious weaknesses and sometimes makes very silly mistakes. AI is going to change the world, but GPT-3 is just a very early glimpse. We have a lot still to figure out.

11:45 AM · Jul 19, 2020 · Twitter Web App

825 Retweets 175 Quote Tweets 7,056 Likes
Today: Frontiers!

and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
Today: Frontiers!
and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?
Utilizing LLM for other domains

Solving math word problems

Question

Ali is a dean of a private school where John is also a dean of a public school in his school. Each class has 1/8 the capacity of the school which has the capacity of 120 students. What is the combined capacity of both schools?

Answer

Method: 175B Verification

Ali’s class has a capacity of 120 students.
Each of John’s classes has a capacity of 1/8 of 120 = 15 students.
The total capacity of John’s two classes is 15 x 2 = 30 students.
The combined capacity of the two schools is 120 + 30 = 150 students.

Code completion

```python
# Load the downloaded data into a dataframe, print results
import pandas as pd
df = pd.DataFrame(response.json()['results'])
print(df)

# Select only the rows with datatype 'TMAX', and convert from C to F. Also change dates to MM-DD
df = df[df['datatype'] == 'TMAX']
df['date'] = pd.to_datetime(df['date'])
df['date'] = df['date'].dt.strftime('%m-%d')
df['value'] = df['value'] * 1.8 + 32
print(df)
```
Today: Frontiers!
and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?

**Takeaways**

Transforming gradient update can allow us to edit models
Large language models can boost generalization in robotics
• New text commands
• New language
• New tasks
• More?
Today: Frontiers!
and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?

Making RL a better data sponge

How to scale up data collection?
How to utilize broad offline data sources?
How to apply RL to lifelong settings?

Open Challenges
Supervised Learning with Deep NN

A lot of data
Supervised Learning with Deep NN

Text Prompt: a snail made of corkscrew, a snail with the texture of corkscrew.

AI-Generated Images:
Reinforcement Learning with Deep NN

<table>
<thead>
<tr>
<th>Methods</th>
<th>MT10</th>
<th>MT50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-task PPO</td>
<td>30.5%</td>
<td>35.4%</td>
</tr>
<tr>
<td>Multi-task TRPO</td>
<td>31.3%</td>
<td>21.0%</td>
</tr>
<tr>
<td>Task embeddings</td>
<td>20.9%</td>
<td>11.8%</td>
</tr>
<tr>
<td>Multi-task SAC</td>
<td>68.3%</td>
<td>38.5%</td>
</tr>
</tbody>
</table>
Better sample complexity (aka less data) → Algorithms that utilize *more* data
Today: Frontiers! and open challenges

Utilizing zero-shot generalization of large (language) models
How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?

Making RL a better data sponge
How to scale up data collection?
How to utilize broad offline data sources?
How to apply RL to lifelong settings?

Open Challenges
Today: Frontiers! and open challenges

Making RL a better data sponge

How to scale up data collection?
How to utilize broad offline data sources?
How to apply RL to lifelong settings?

**Takeaways**

We can scale up data collection by eliminating resets

There are multiple ways to broaden data sources via:
- Incorporating non-robotic data
- Casting RL into a simpler problem formulation

We can extend RL to lifelong settings by retaining experiences
Today: Frontiers!
and open challenges

Utilizing zero-shot generalization of large (language) models

How to fine-tuning large models without negative transfer?
How to apply large (language) models to other domains?

Making RL a better data sponge

How to scale up data collection?
How to utilize broad offline data sources?
How to apply RL to lifelong settings?

Open Challenges
Open Challenges in Multi-Task and Meta Learning

(that we haven't previously covered)
Open Challenges in Multi-Task and Meta Learning

Addressing fundamental problem assumptions

- Generalization: Out-of-distribution tasks, long-tailed task distributions
The problem with long-tailed distributions.

We learned how to do few-shot learning...

...but these few-shot tasks may be from a different distribution.

We've seen some generalization to the tail:
- Editing LLM to a new distribution

Further hints might come from domain adaptation, robustness literature.
Open Challenges in Multi-Task and Meta Learning

Addressing fundamental problem assumptions

- Generalization: Out-of-distribution tasks, long-tailed task distributions
- Multimodality: Can you learn priors from multiple modalities of data?
Rich sources of prior experiences.

- visual imagery
- tactile feedback
- language
- social cues

Can we learn priors across multiple data modalities?

- Varying dimensionalities, units
- Carry different, complementary forms of information

Some hints might come from multimodal learning literature.

Open Challenges in Multi-Task and Meta Learning

Addressing fundamental problem assumptions

- Generalization: Out-of-distribution tasks, long-tailed task distributions
- Multimodality: Can you learn priors from multiple modalities of data?
- Algorithm, Model Selection: When will multi-task learning help you?

Benchmarks

- Breadth: That challenge current algorithms to find common structure
- Realistic: That reflect real-world problems
Some steps towards good benchmarks

**Meta-Dataset**
Triantafillou et al. ’19

**Meta-World Benchmark**
Yu et al. ’19

**Visual Task Adaptation Benchmark**
Zhai et al. ’19

**Taskonomy Dataset**
Zamir et al. ’18

**Goal:** reflection of real world problems + appropriate level of difficulty + ease of use
Open Challenges in Multi-Task and Meta Learning

Addressing fundamental problem assumptions

- Generalization: Out-of-distribution tasks, long-tailed task distributions
- Multimodality: Can you learn priors from multiple modalities of data?
- Algorithm, Model Selection: When will multi-task learning help you?

Benchmarks

- Breadth: That challenge current algorithms to find common structure
- Realistic: That reflect real-world problems

Improving core algorithms

- Theory: Develop a theoretical understanding of the performance of these algorithms
- Multi-Step Problems: Performing tasks in sequence presents challenges.

+ the challenges you discovered in your homework & final projects!
The Bigger Picture
Machines are specialists.
Humans are generalists.

Source: https://youtu.be/8vNxjwt2Aq
A Step Towards Generalists

Some of what we covered in CS330:

- learn **multiple tasks** (multi-task learning)
- leverage **prior experience** when learning new things (meta-learning)
- solve **sequential decision problems** (RL)
- learn **general-purpose models** (model-based RL)
- prepare for tasks **before** you know what they are (exploration, skill discovery)
- perform tasks **in sequence** (hierarchical RL)
- learn **continuously** (lifelong learning)

What’s missing?
Logistics

Today: This is the last lecture!

Tuesday (Nov 30th): Project poster session

Wednesday (Dec 8th): Project due