Chen Data Science & AI for Neuroscience Summer School

CHEN TIANQIAO & CHRISSY INSTITUTE

Caltech

Model Engineering Principles

Sabera Talukder

What are your model engineering principles?

- GIGO ⟶ Garbage In Garbage Out

- GIGO → Garbage In Garbage Out
- KISS → Keep It Simple Stupid

- GIGO → Garbage In Garbage Out
- KISS → Keep It Simple Stupid
- If the simple stuff doesn’t work: the hard problems are hard because picking the right model (architecture) is hard

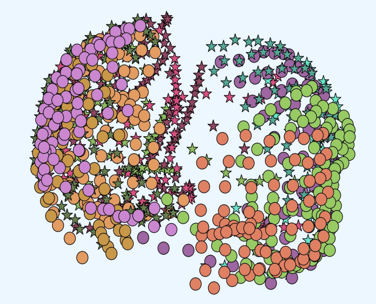
- GIGO → Garbage In Garbage Out
- KISS → Keep It Simple Stupid
- If the simple stuff doesn’t work: the hard problems are hard because picking the right model (architecture) is hard
- Your representation is only as good as your target task!

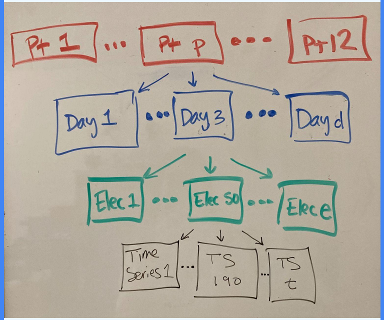
- GIGO → Garbage In Garbage Out
- KISS → Keep It Simple Stupid
- If the simple stuff doesn’t work: the hard problems are hard because picking the right model (architecture) is hard
- Your representation is only as good as your target task!
- Emphasizing Modularity (***with a big ML exception)

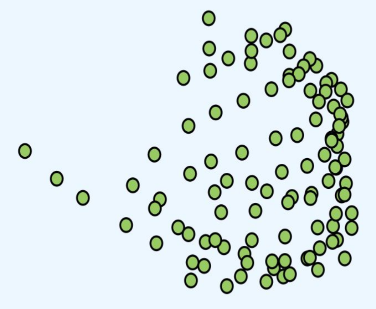
- GIGO → Garbage In Garbage Out
- KISS → Keep It Simple Stupid
- If the simple stuff doesn’t work: the hard problems are hard because picking the right model (architecture) is hard
- Your representation is only as good as your target task!
- Emphasizing Modularity (***with a big ML exception)
- Ground Truth / Baselines / Reproducibility

Garbage In Garbage Out

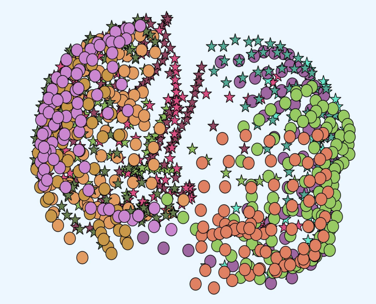
Garbage In Garbage Out

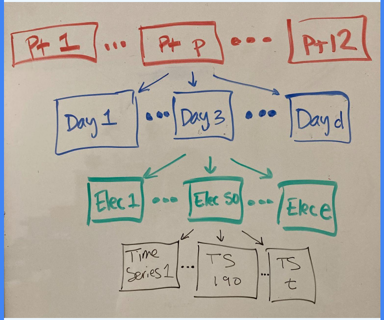


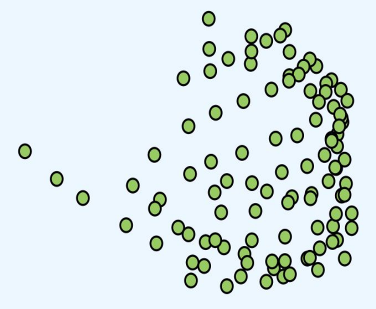




Garbage In Garbage Out





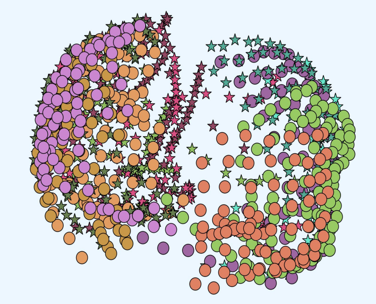


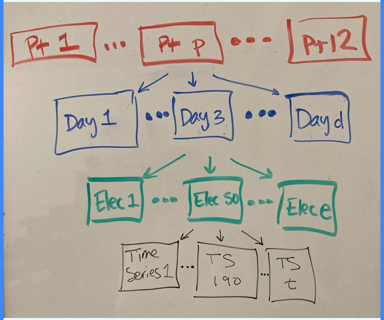
For Each
Patient Flatten
Data Along
Days and
Electrodes

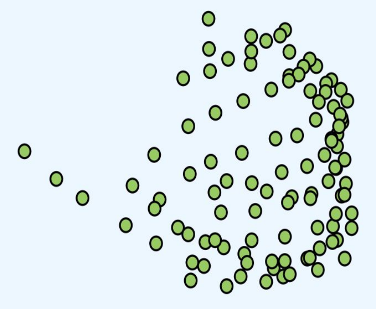
+

PCA

Garbage In Garbage Out







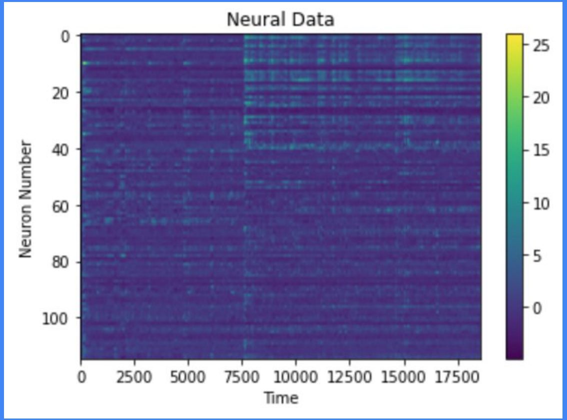
For Each Patient Flatten Data Along Days and Electrodes

+

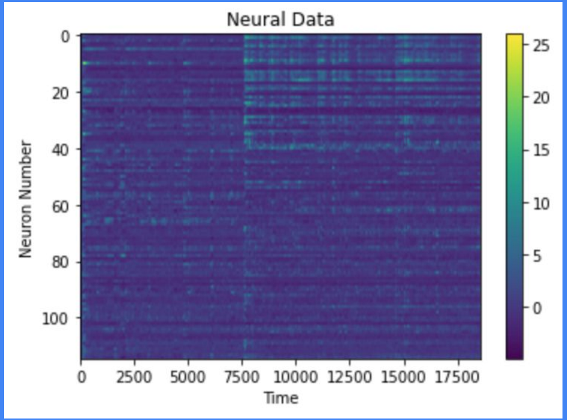
PCA

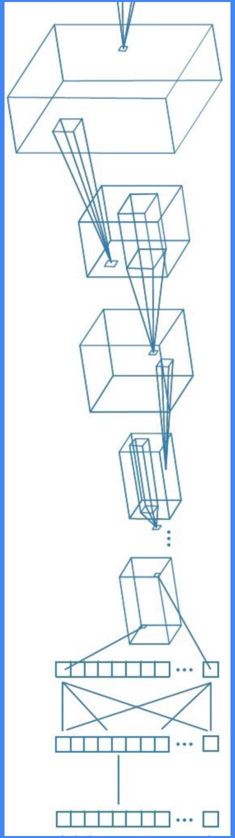
Keep It Simple Stupid

Keep It Simple Stupid

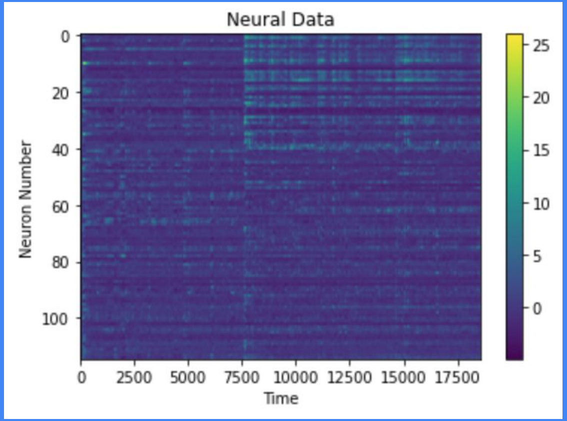


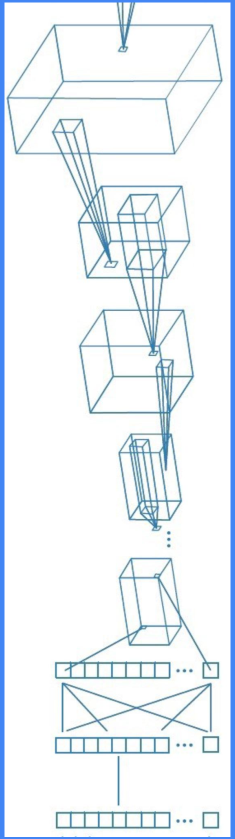
Keep It Simple Stupid



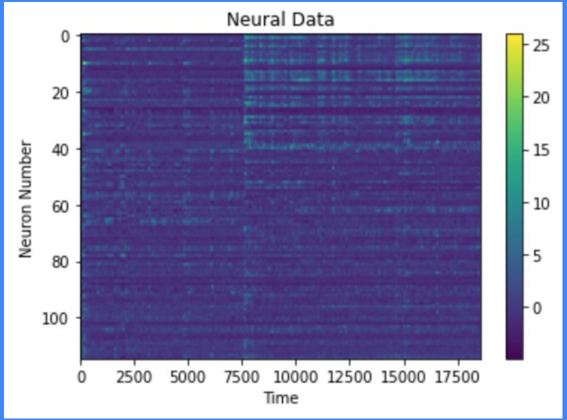


Keep It Simple Stupid

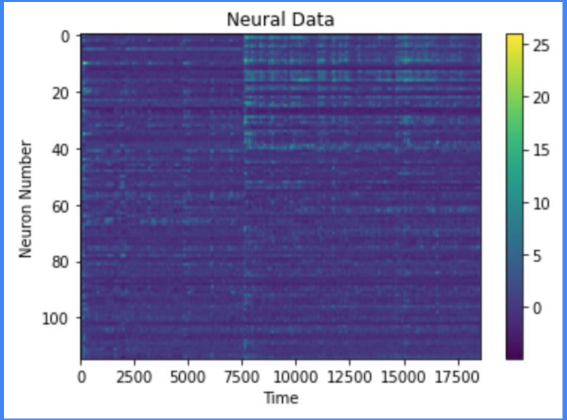




Keep It Simple Stupid

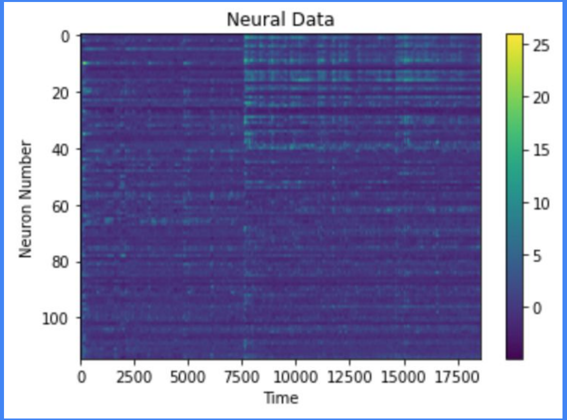


Keep It Simple Stupid

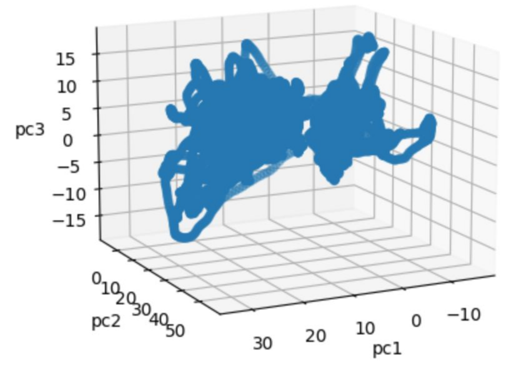


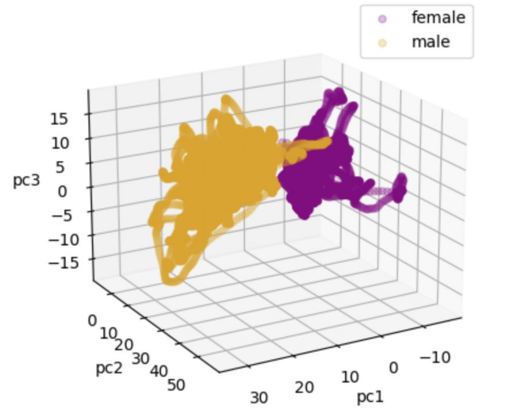
PCA

Keep It Simple Stupid



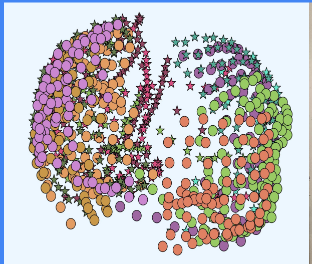
PCA

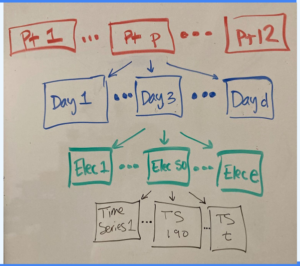


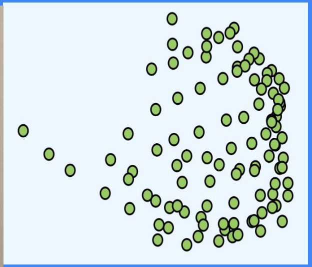


If the simple stuff doesn’t work: the hard problems are hard because picking the right model (architecture) is hard!

If the simple stuff doesn't work: the hard problems are hard because picking the right model (architecture) is hard!

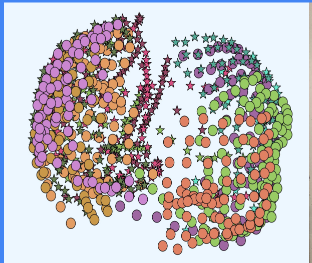


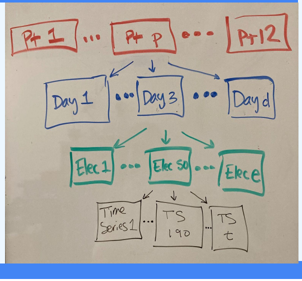


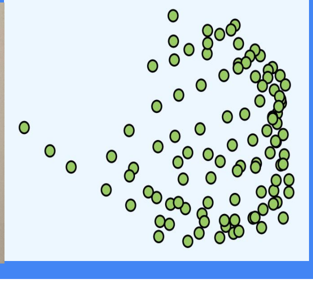


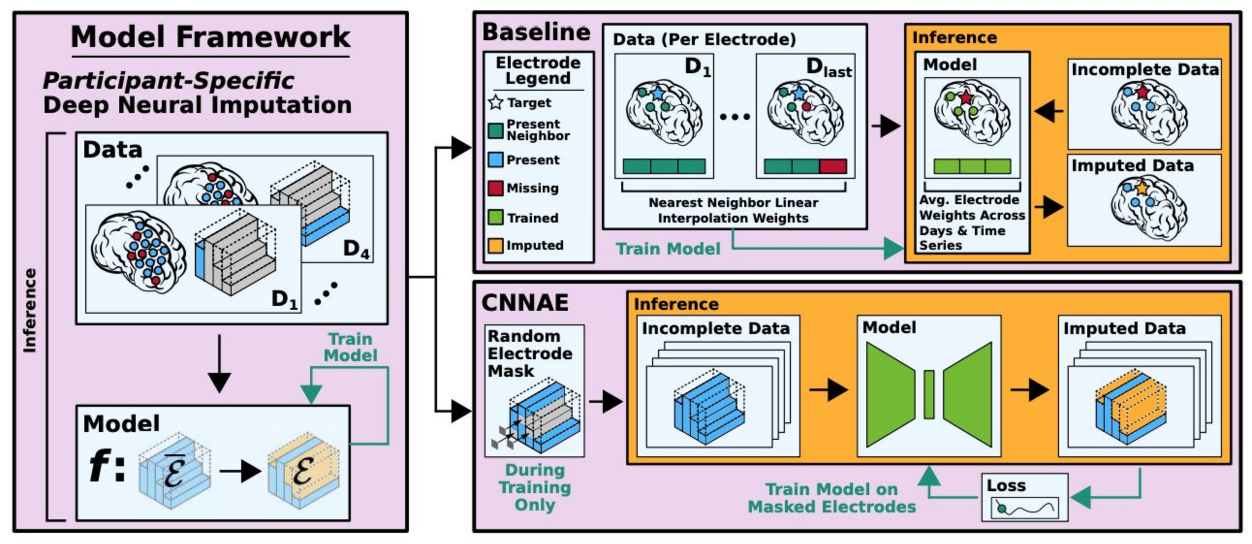
Talukder et al. 2022

If the simple stuff doesn't work: the hard problems are hard because picking the right model (architecture) is hard!



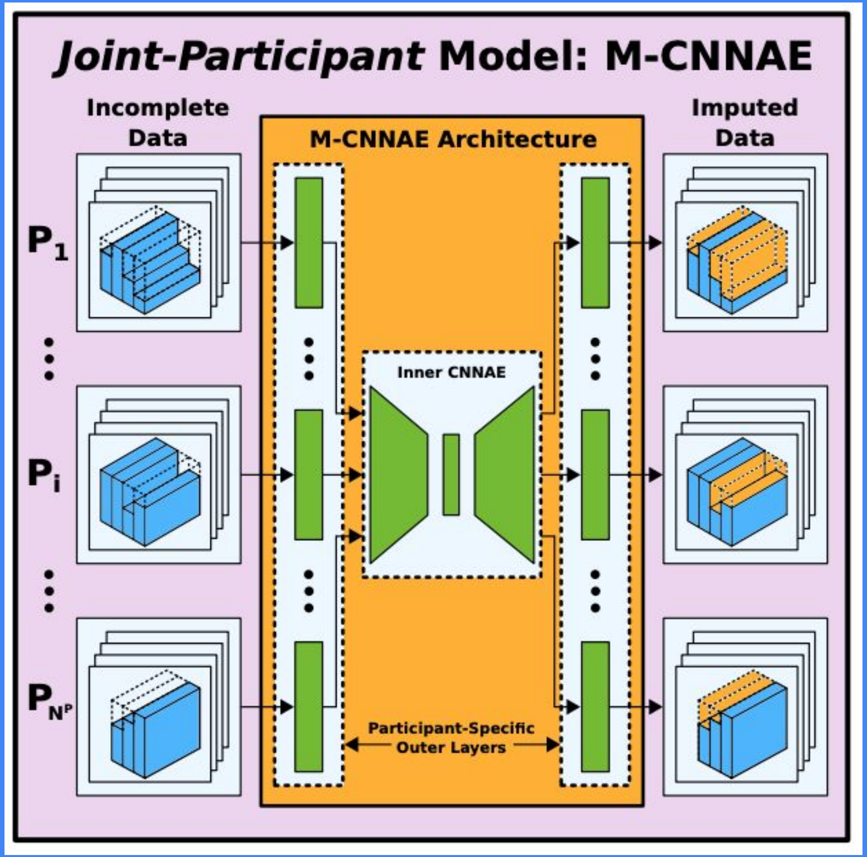






Talukder et al. 2022

If the simple stuff doesn't work: the hard problems are hard because picking the right model (architecture) is hard!



Talukder et al. 2022

Your representation is only as good as your target task

Your representation is only as good as your target task

Let's play a game called: ***How good is this representation!***

Jennifer Sun

Your representation is only as good as your target task

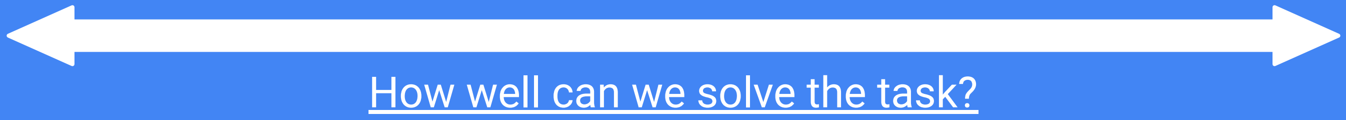
Let's play a game called: ***How good is this representation!***

How well can we solve the task?

Jennifer Sun

Your representation is only as good as your target task

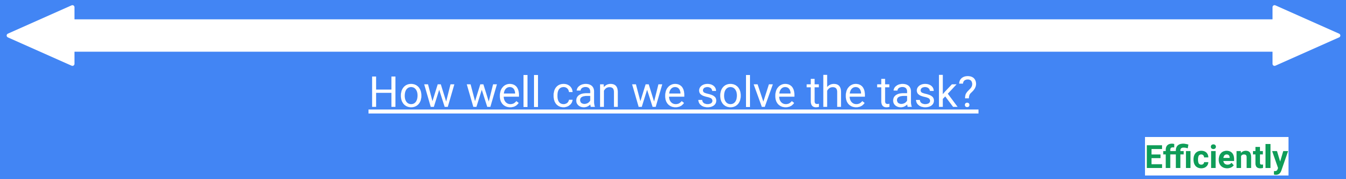
Let’s play a game called: ***How good is this representation!***



How well can we solve the task?

Your representation is only as good as your target task

Let’s play a game called: ***How good is this representation!***

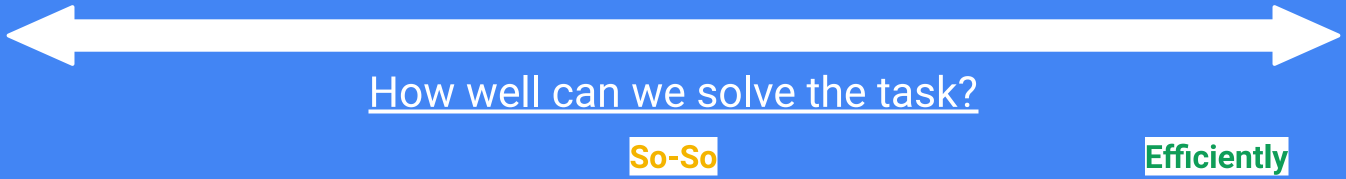


How well can we solve the task?

Efficiently

Your representation is only as good as your target task

Let’s play a game called: ***How good is this representation!***

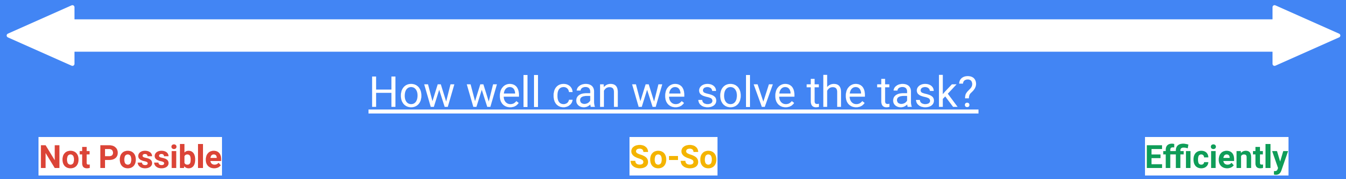


How well can we solve the task?

Jennifer Sun

Your representation is only as good as your target task

Let’s play a game called: ***How good is this representation!***



How well can we solve the task?

Not Possible

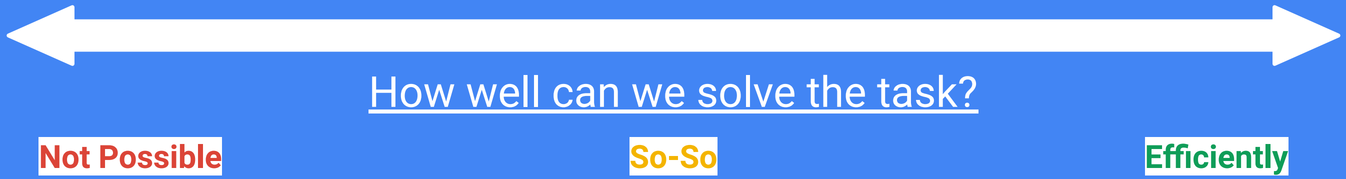
So-So

Efficiently

Your representation is only as good as your target task

Let’s play a game called: ***How good is this representation!***

Round 1: What is the species?



How well can we solve the task?

Not Possible

So-So

Efficiently

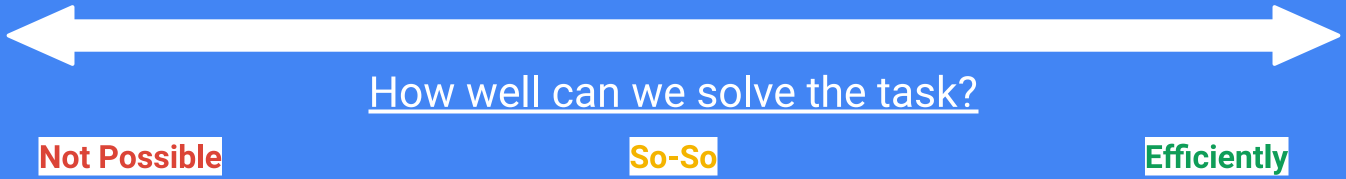
Jennifer Sun

Your representation is only as good as your target task

Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG



How well can we solve the task?

Not Possible

So-So

Efficiently

Your representation is only as good as your target task

Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

How well can we solve the task?

Not Possible

So-So

Efficiently

Jennifer Sun

Your representation is only as good as your target task

Let’s play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

How well can we solve the task?

Not Possible

So-So

Efficiently

Jennifer Sun

Your representation is only as good as your target task

Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

Round 2: What is the color of the dog?

How well can we solve the task?

Not Possible

So-So

Efficiently

Your representation is only as good as your target task

Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

Round 2: What is the color of the dog?

DOG

How well can we solve the task?

Not Possible

So-So

Efficiently

Jennifer Sun

Your representation is only as good as your target task

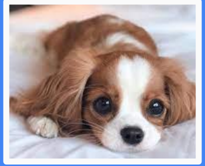
Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

Round 2: What is the color of the dog?

DOG



How well can we solve the task?

Not Possible

So-So

Efficiently

Jennifer Sun

Your representation is only as good as your target task

Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

Round 2: What is the color of the dog?

DOG

Each pixel color of image

How well can we solve the task?

Not Possible

So-So

Efficiently

Jennifer Sun

Your representation is only as good as your target task

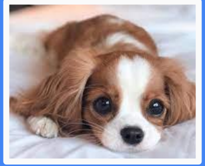
Let's play a game called: ***How good is this representation!***

Round 1: What is the species?

DOG

Round 2: What is the color of the dog?

DOG



Each pixel color of image

How well can we solve the task?

Not Possible **So-So** **Efficiently**

Jennifer Sun

Your representation is only as good as your target task

In the end, we like this representation because it works for many tasks!
(and it’s cute 😍)

Jennifer Sun

Emphasize Modularity

Emphasize Modularity

How would you implement modularity into your development pipeline?

Emphasize Modularity

- Big & recent example: representations that once built can be used for many downstream tasks.

Emphasize Modularity

- Big & recent example: representations that once built can be used for many downstream tasks

- Separating the data processing pipeline from your model development pipeline.

Emphasize Modularity

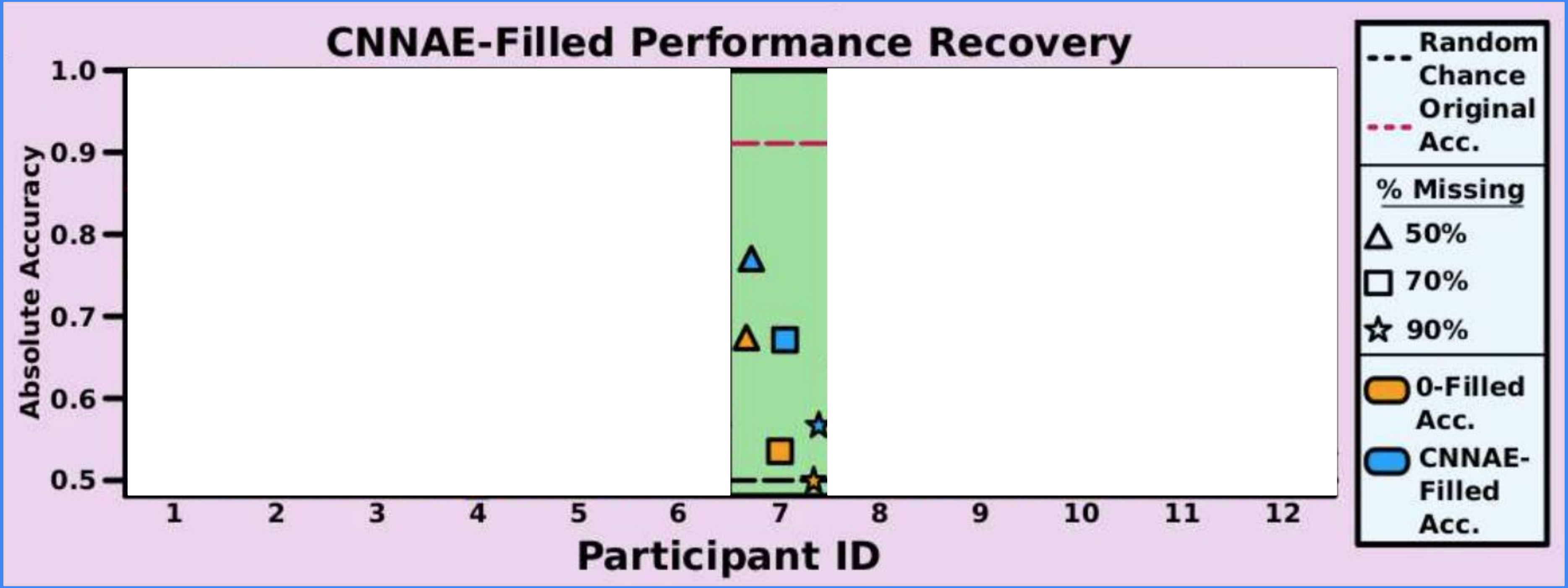
- Big & recent example: representations that once built can be used for many downstream tasks

- Separating the data processing pipeline from your model development pipeline.

- ***Exception*** in machine learning end-to-end modeling is all the rage! Allows the model to learn the data features to solve the task without intervention.

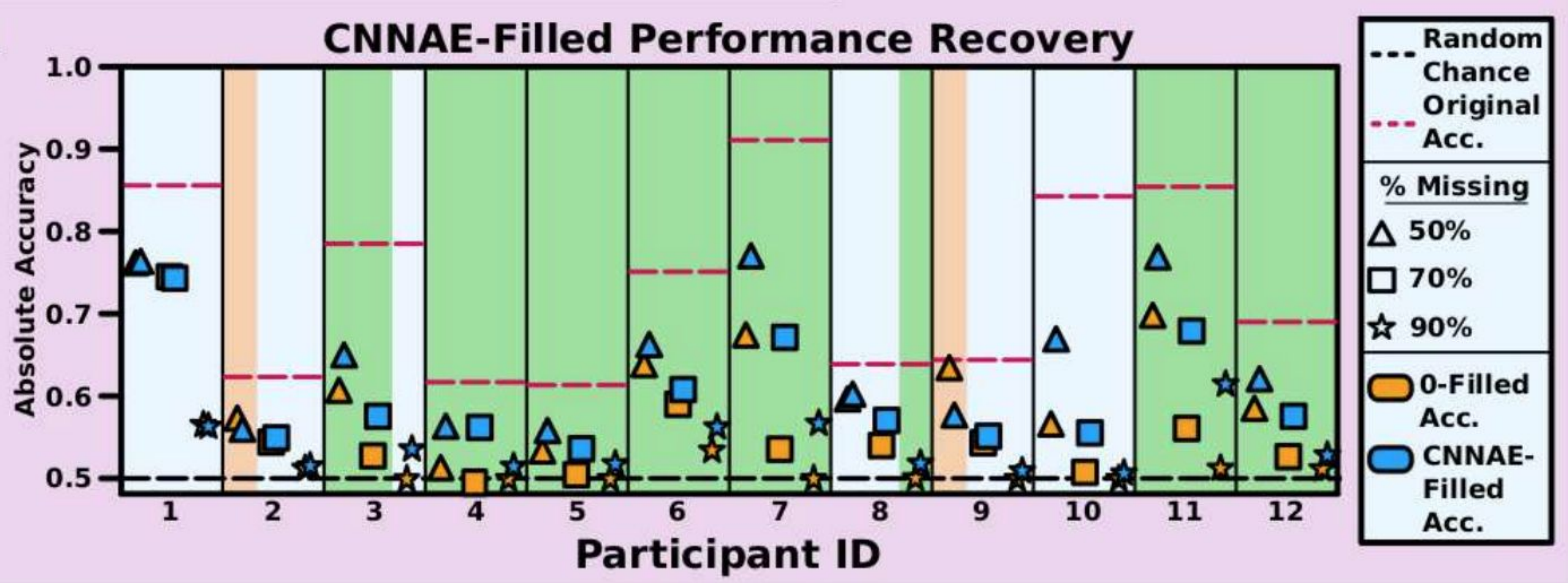
Ground Truth / Baselines / Reproducibility

Ground Truth / Baselines / Reproducibility



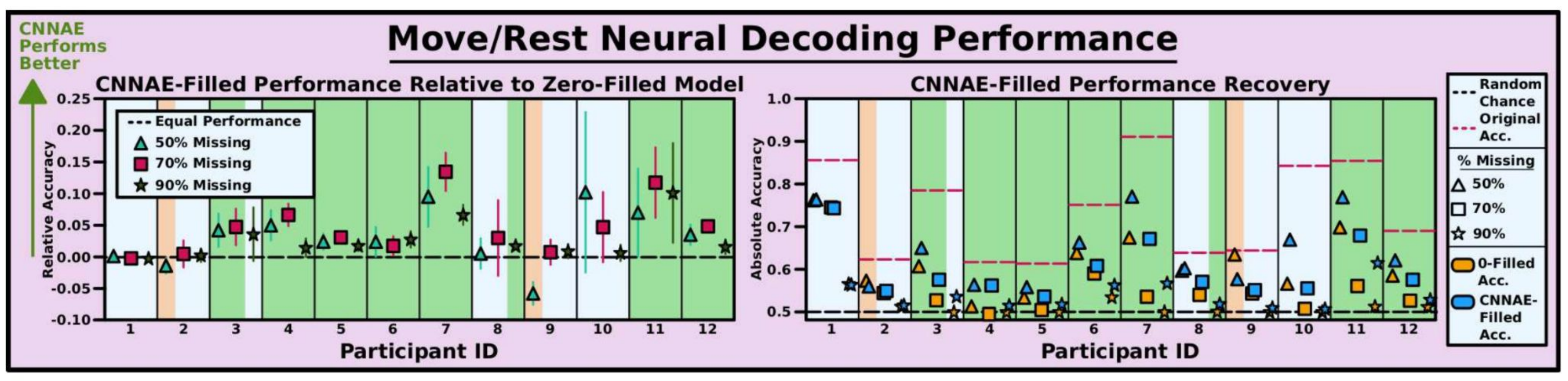
Talukder et al. 2022

Ground Truth / Baselines / Reproducibility



Talukder et al. 2022

Ground Truth / Baselines / Reproducibility



Talukder et al. 2022