a place of mind
THE UNIVERSITY OF BRITISH COLUMBIA

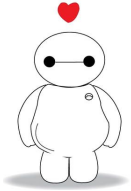
ACL 2025

Delta-KNN: Improving Demonstration Selection in In-Context Learning for Alzheimer's Disease Detection

Chuyuan Li, Raymond Li, Thalia S. Field, Giuseppe Carenini

Department of Computer Science
The University of British Columbia

chuyuan.li@ubc.ca



1 Healthcare Task: Alzheimer's Disease Detection

Dementia detection from patient-generated discourse

Alzheimer's Disease

- Severe neurodegenerative disorder, leads to **dementia**.
- Affecting **55 million** people worldwide.
- Among one of the **most costly** diseases.
- **Early prevention** is crucial.
  - Symptoms, e.g. **language disorders**

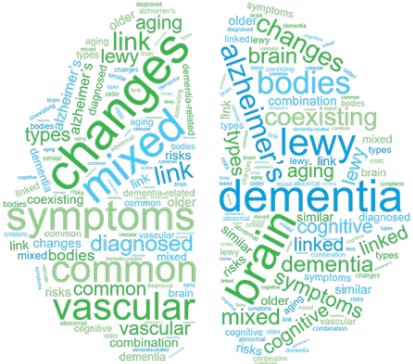
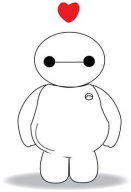


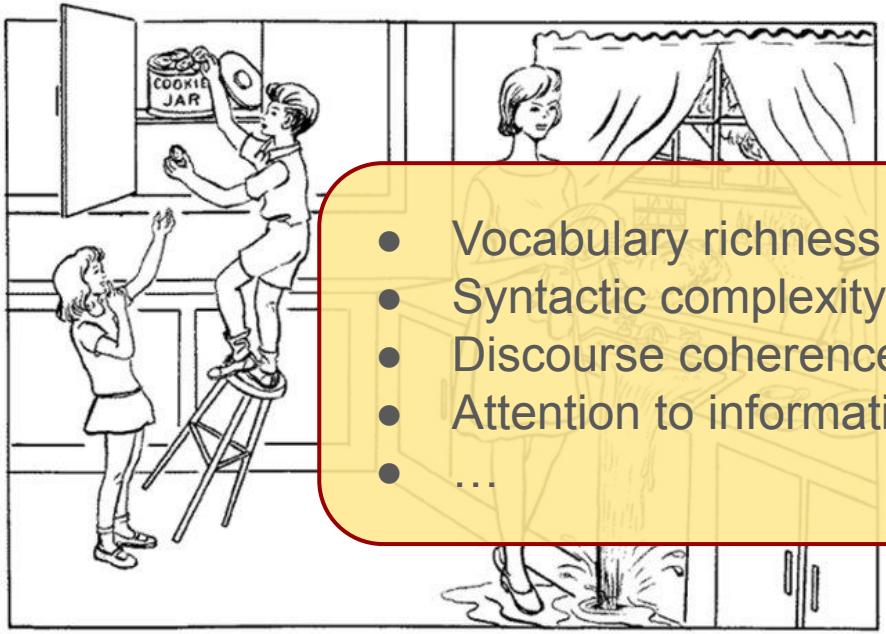
Image credit: https://leafcare.co.uk/blog/alzheimers-disease-symptoms-and-causes/



1 Healthcare Task: Alzheimer's Disease Detection

Dementia detection from patient-generated discourse

- The Cookie Theft picture description task



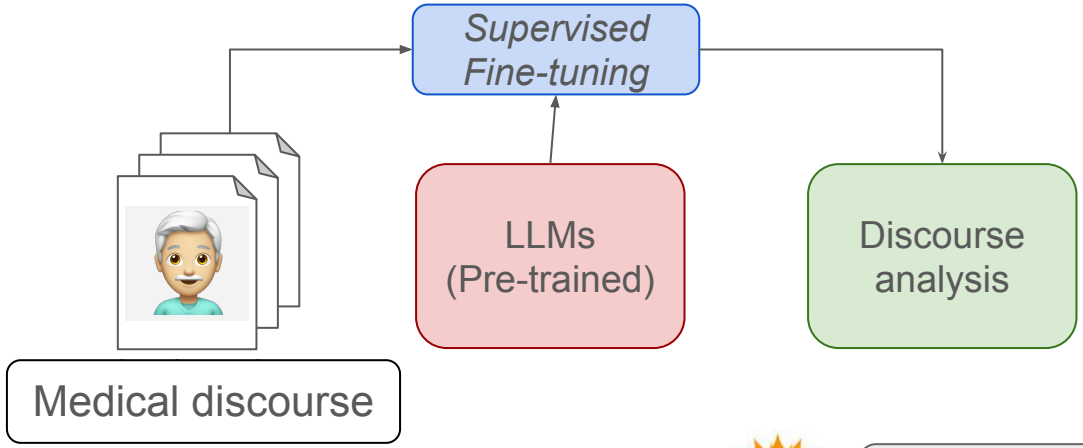
- Vocabulary richness
- Syntactic complexity
- Discourse coherence
- Attention to information units
- …

Please describe everything you see going on in this picture.

Well the sink is running over. She's drying the dishes. They're getting in the cookie jar …

2 Improving LLMs In-Context Learning for AD Detection

In-context learning for LLMs



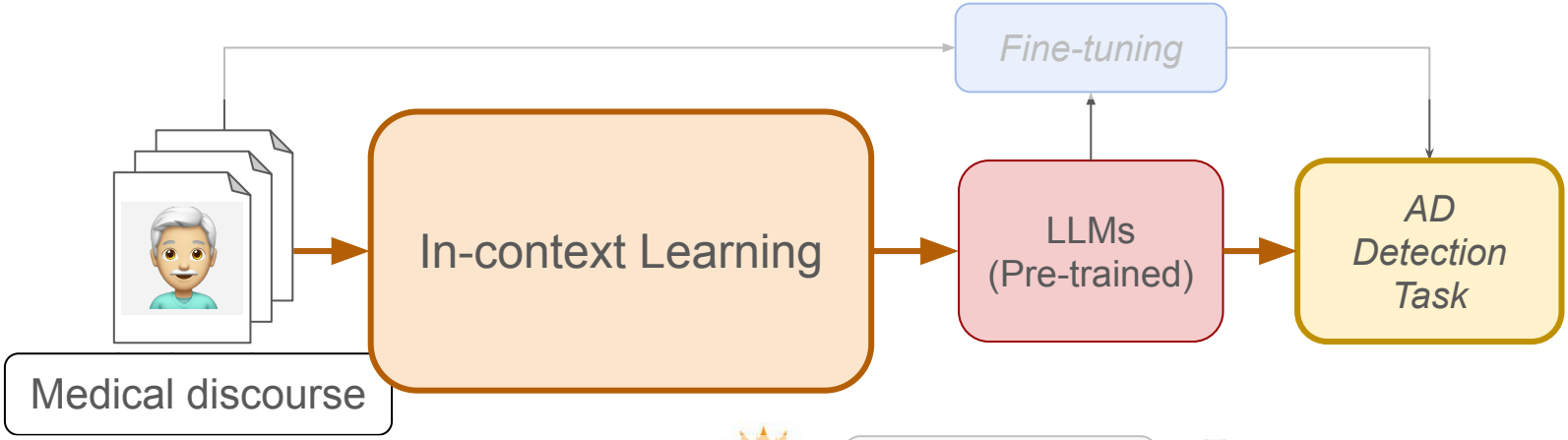
- Challenges:
  - Extreme limited data
  - Unstable in fine-tuning

		No. of patients:	*No. of* controls:
	ADReSS [1] (DementiaBank)	#78	#78
	Canary [2] Dataset (UBC)	#63	#67

[1] Luz, Saturnino, et al. "Detecting cognitive decline using speech only: The adresso challenge." In INTERSPEECH 2021. ISCA. 903.
[2] Jang, Hyeju, et al. "Classification of Alzheimer's disease leveraging multi-task machine learning analysis of speech and eye-movement data." Frontiers in Human Neuroscience 15 (2021): 716670.

2 Improving LLMs In-Context Learning for AD Detection

In-context learning for LLMs



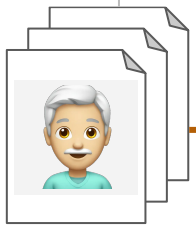
- Challenges:
  - Extreme limited data
  - Unstable in fine-tuning

ADReSS [3] (DementiaBank)	#78	#78
Canary [4] Dataset (UBC)	#63	#67

[3] Luz, Saturnino, et al. "Detecting cognitive decline using speech only: The adresso challenge." In INTERSPEECH 2021. ISCA. 903.
[4] Jang, Hyeju, et al. "Classification of Alzheimer's disease leveraging multi-task machine learning analysis of speech and eye-movement data." Frontiers in Human Neuroscience 15 (2021): 716670.

2 Improving LLMs In-Context Learning for AD Detection

In-context learning for LLMs



In-context Learning

LLMs (Pre-trained)

AD Detection Task

- **Few-shot learning**
  - Random sampling
    - General, representative
  - Similarity-based
    - Semantically similar
  - $H(Y|X)$ Text-understanding-based
    - Conditional entropy / perplexity

2 Improving LLMs In-Context Learning for AD Detection

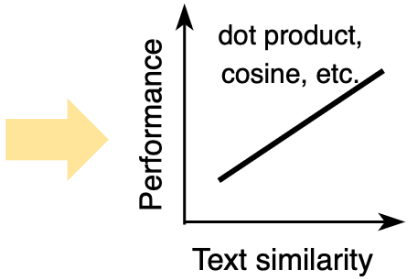
Method: In-context learning via demonstration selection

- Related work

ACL DeeLIO workshop 2022 [3]

What Makes Good In-Context Examples for GPT-3?

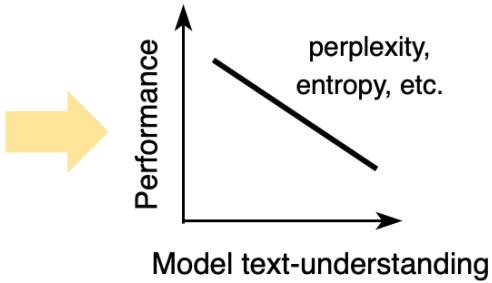
Jiachang Liu[1]*, **Dinghan Shen**[2], **Yizhe Zhang**[3], **Bill Dolan**[4], **Lawrence Carin**[1], **Weizhu Chen**[2]
[1]Duke University [2]Microsoft Dynamics 365 AI [3]Meta AI [4]Microsoft Research
[1]{jiachang.liu, lcarin}@duke.edu
[3]yizhe.zhang@hotmail.com
[2,4]{dishen, billdol, wzchen}@microsoft.com



ACL 2024 [4]

Revisiting Demonstration Selection Strategies in In-Context Learning

Keqin Peng[1], **Liang Ding**[2]*, **Yancheng Yuan**[3]*
Xuebo Liu[4], **Min Zhang**[4], **Yuanxin Ouyang**[1], **Dacheng Tao**[5]
[1]Beihang University [2]The University of Sydney [3]The Hong Kong Polytechnic University
[4]Harbin Institute of Technology, Shenzhen [5]Nanyang Technological University
keqin.peng@buaa.edu.cn, liangding@gmail.com

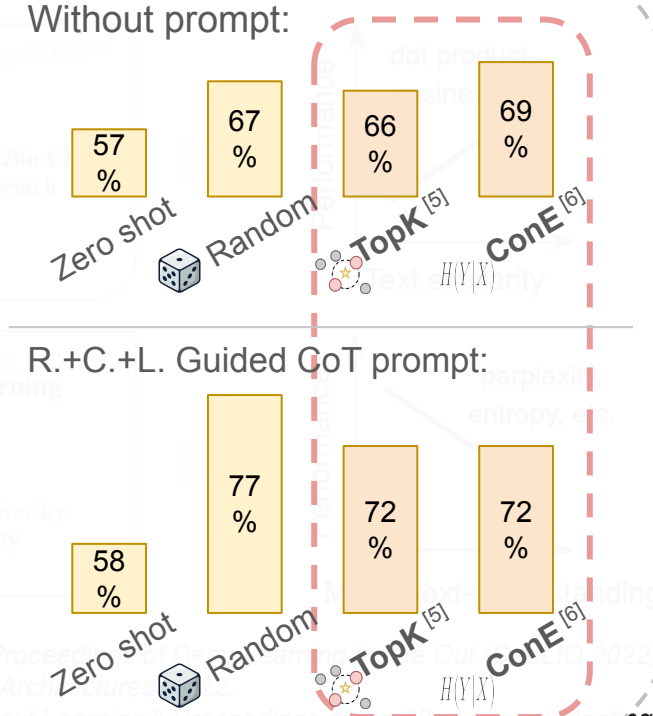


[3] Liu, Jiachang, et al. "What Makes Good In-Context Examples for GPT-3?." *Proceedings of Deep Learning Inside Out (DeeLIO 2022): The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*. 2022.
[4] Peng, Keqin, et al. "Revisiting Demonstration Selection Strategies in In-Context Learning." *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2024.

2 Improving LLMs In-Context Learning for AD Detection

Method: In-context learning via demonstration selection

- Related work
  - **Good performance on:**
    - ✓ Question answering
    - ✓ Commonsense reasoning
    - ✓ SQL generation
    - …
  - **Need to be studied on:**
    - ? AD Detection (medical discourse)
    - …



[3] L… Liu, Jiachang, et al. "What Makes Good In-Context Examples for GPT-3?" *Proceedings of Deep Learning Inside Out (DeeLIO 2022): The … Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*. 2022.

[4] Peng, Keqin, et al. "Revisiting Demonstration Selection Strategies in In-Context Learning." *Proceedings of the … Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2024.

2 Improving LLMs In-Context Learning for AD Detection

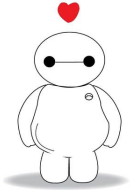
Method: In-context learning via demonstration selection

- Related work
  - **Good performance on:**
    - ✓ Question answering
    - ✓ Commonsense reasoning
    - ✓ SQL generation
  - **Need to be studied on:**
    - ? AD Detection (Medical discourse)

- Longer context
- Subtle linguistic differences
- Complex conceptual understanding

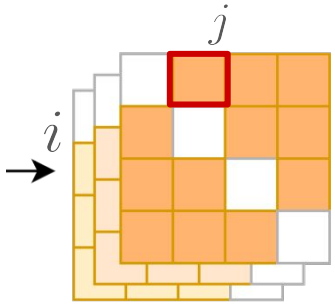
[5] Liu, Jiachang, et al. "What Makes Good In-Context Examples for GPT-3?" *Proceedings of The 3rd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*. 2022.

[6] Peng, Keqin, et al. "Revisiting Demonstration Selection Strategies in In-Context Learning." *Proceedings of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2024.



2 Improving LLMs In-Context Learning for AD Detection

Method: In-context learning via demonstration selection

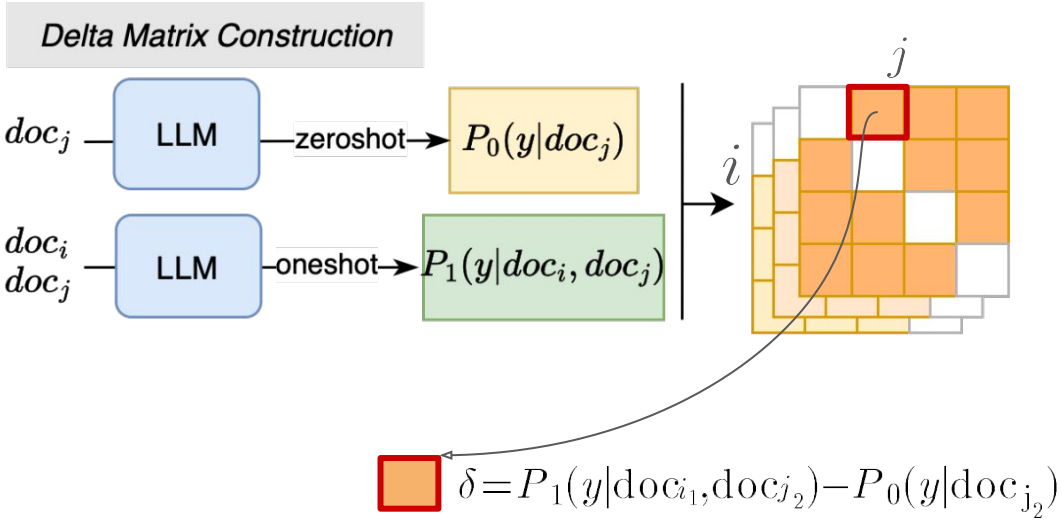


Delta Matrix

- A **“look-up”** table
- **Relative gain** contributed by a demo example *doc i* to a target *doc j*.

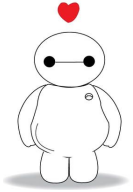
2 Improving LLMs In-Context Learning for AD Detection

Method: In-context learning via demonstration selection



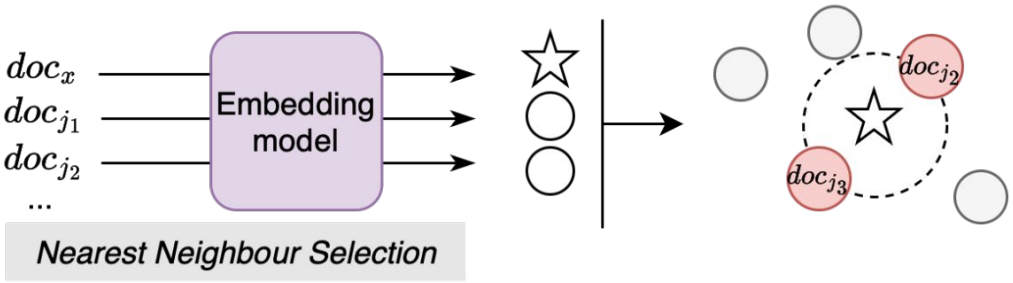
Delta Matrix

- A **"look-up"** table
- **Relative gain** contributed by a demo example *doc i* to a target *doc j*.
- **Gain = Delta score**: difference between one-shot (P_1) and zero-shot (P_0) performance.



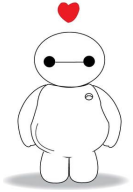
2 Improving LLMs In-Context Learning for AD Detection

Method: In-context learning via demonstration selection



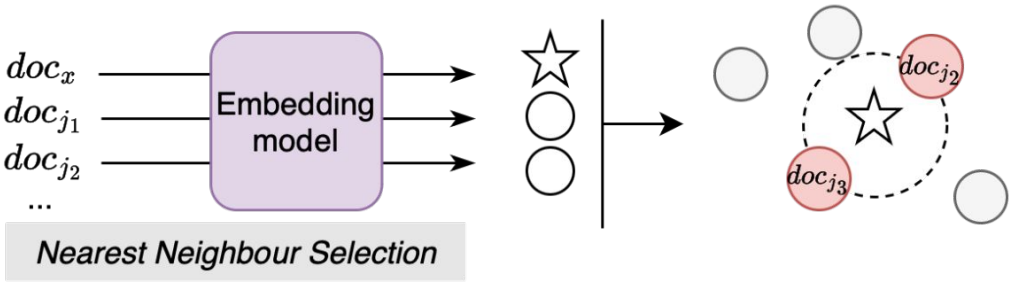
Delta-KNN retriever

- Identify *"**representatives**"* for the target *doc x*.
- Searching for its **nearest neighbors** in the training set.



2 Improving LLMs In-Context Learning for AD Detection

Method: In-context learning via demonstration selection



Delta-KNN retriever

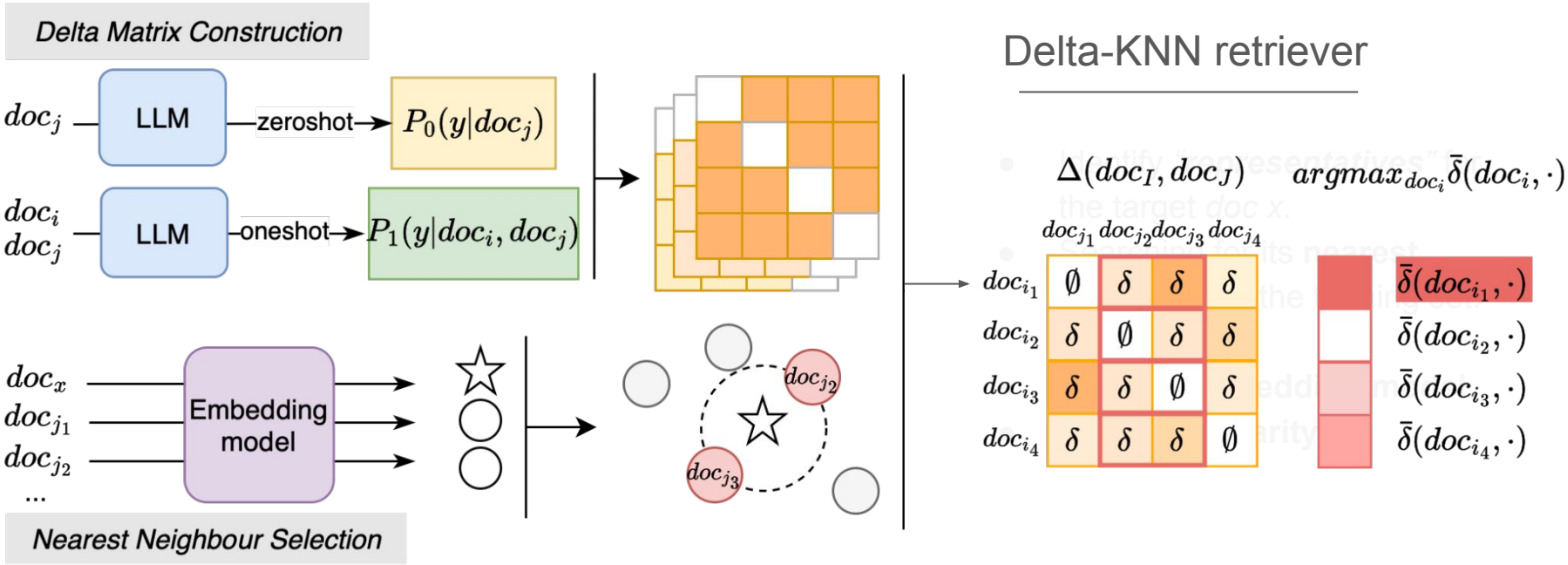
- Identify *"**representatives**"* for the target *doc x*.
- Searching for its **nearest neighbors** in the training set.

- Open-AI **embedding model** [1].
- **Cosine similarity**.

1. The latest text-embedding-3-large model (https://openai.com/index/new-embedding-models-and-api-updates/)

2 Improving LLMs In-Context Learning for AD Detection

Method: In-context learning via demonstration selection



2 Improving LLMs In-Context Learning for AD Detection

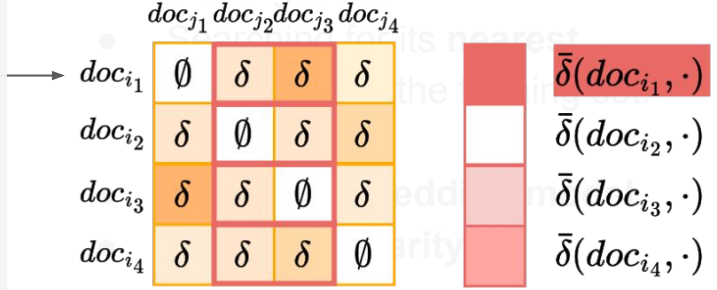
Method: In-context learning via demonstration selection

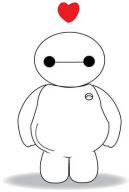
Combination

- **Average delta score** of each demo example (doc_i) over the *"representatives"* of the target input x.
  - → ***Expected* gain**
- **Winning example: maximizes** average delta score.

Delta-KNN retriever

$\Delta(doc_I, doc_J)$ $argmax_{doc_i} \bar{\delta}(doc_i, \cdot)$





2 Improving LLMs In-Context Learning for AD Detection

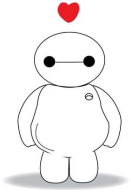
Method: In-context learning via demonstration selection

Combination

- **Average delta score** of each demo example (doc_i) over the "representatives" of the target input
  - → *Expected* gain
- **Winning example: maximizes** average delta score.

Delta-KNN retriever

- **Grounded in empirical evidence of performance gains**
  → **More reliable.**

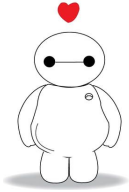


2 Improving LLMs In-Context Learning for AD Detection

Experiments and Results

Llama-3.1-8B-Instruct

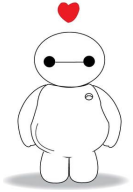
	Accuracy
Zero-shot Prompting	58.0%
ICL Random Sampling	77.7%
ICL Top-*k* Selection	72.4%
$H(Y\|X)$ ICL Conditional entropy-based Selection	72.4%
SVM Classifier (linguistic features)	79.9%
BERT Fine-tuned Classifier	79.3%
δ ICL Delta-KNN (ours)	83.6%



2 Improving LLMs In-Context Learning for AD Detection

Experiments and Results

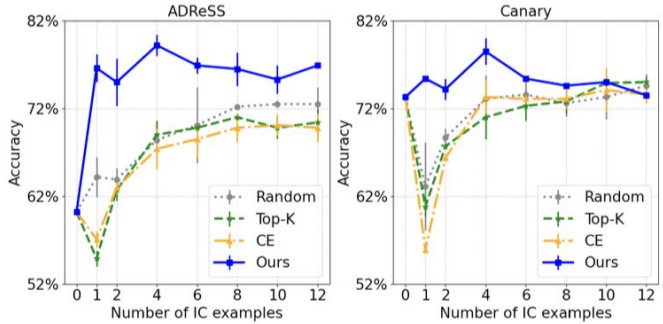
	ADReSS-train	ADReSS-test	Canary
Mistral-7B-Instruct-v0.3			
Zero-shot	$52.3_{0.5}$	$67.7_{1.0}$	$63.1_{0.8}$
Random	$62.0_{2.8}$	$70.8_{2.1}$	$55.0_{0.4}$
Top-k Select.	$53.2_{2.3}$	$63.5_{3.1}$	$62.3_{0.0}$
CE Select.	$61.1_{1.9}$	$66.7_{4.2}$	$58.8_{3.5}$
Ours	$\mathbf{69.9_{1.4}}$	$\mathbf{76.0_{5.2}}$	$\mathbf{72.3_{0.4}}$
Qwen2.5-7B-Instruct			
Zero-shot	$61.6_{0.5}$	$66.8_{2.2}$	$63.5_{0.4}$
Random	$62.0_{2.8}$	$57.3_{1.0}$	$64.6_{3.8}$
Top-k Select.	$58.8_{1.4}$	$66.7_{2.1}$	$53.1_{6.2}$
CE Select.	$58.8_{0.5}$	$65.8_{5.3}$	$60.0_{1.5}$
Ours	$\mathbf{63.4_{0.5}}$	$\mathbf{67.7_{0.0}}$	$\mathbf{66.1_{2.7}}$

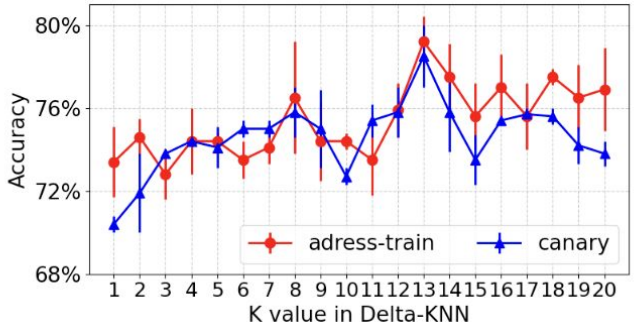


2 Improving LLMs In-Context Learning for AD Detection

Further Investigation

- Impact of **in-context examples N**: Ours shows immediate advantage at N=1, peaking at N=4, after it stabilizes.
- Impact of **Demonstration Ordering**: Ours achieves higher maximum and average accuracy across 24 possible orderings in the 4-shot setting, with lower standard deviation.
- Impact of **Prompt Engineering**: Seven prompt variations, ours consistently outperforms ICL baselines.
- Impact of **k value in Delta-KNN**: Varying k from 1 to 20 on train sets, found k=13 yields the best results on both datasets.





Thank you!

Chuyuan Li, Raymond Li, Thalia S. Field, Giuseppe Carenini

Welcome to our poster if you have any question or would like to learn more!