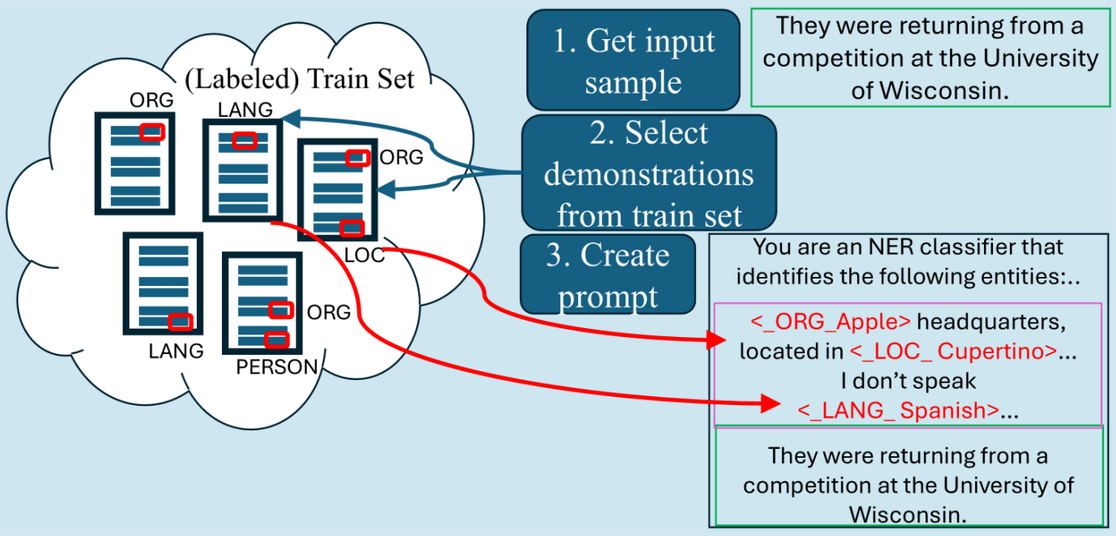
In-Context Learning on a Budget: A Case Study in Token Classification

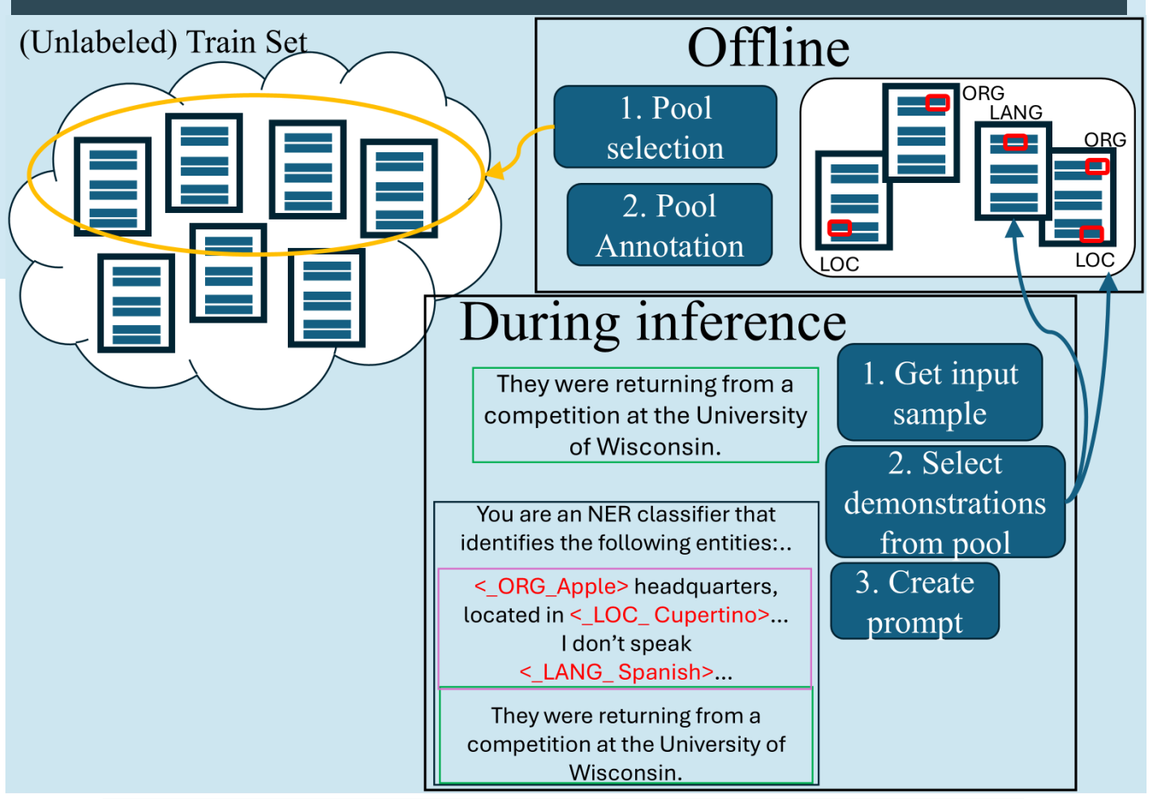
Uri Berger [1,2]
Tal Baumel [3]
Gabriel Stanovsky [1]

[1] The Hebrew University of Jerusalem
[2] University of Melbourne
[3] Microsoft Corporation

Classic In-Context Learning



In-Context Learning on a Budget



Pool Selection Methods

Central
Select the K samples that are closest to the center of the train set

Vote-K
Select diverse samples (Su et al. 2022)

Cluster
Cluster train set to K clusters and select one sample per cluster

Random
Randomly select K train samples

Experiments

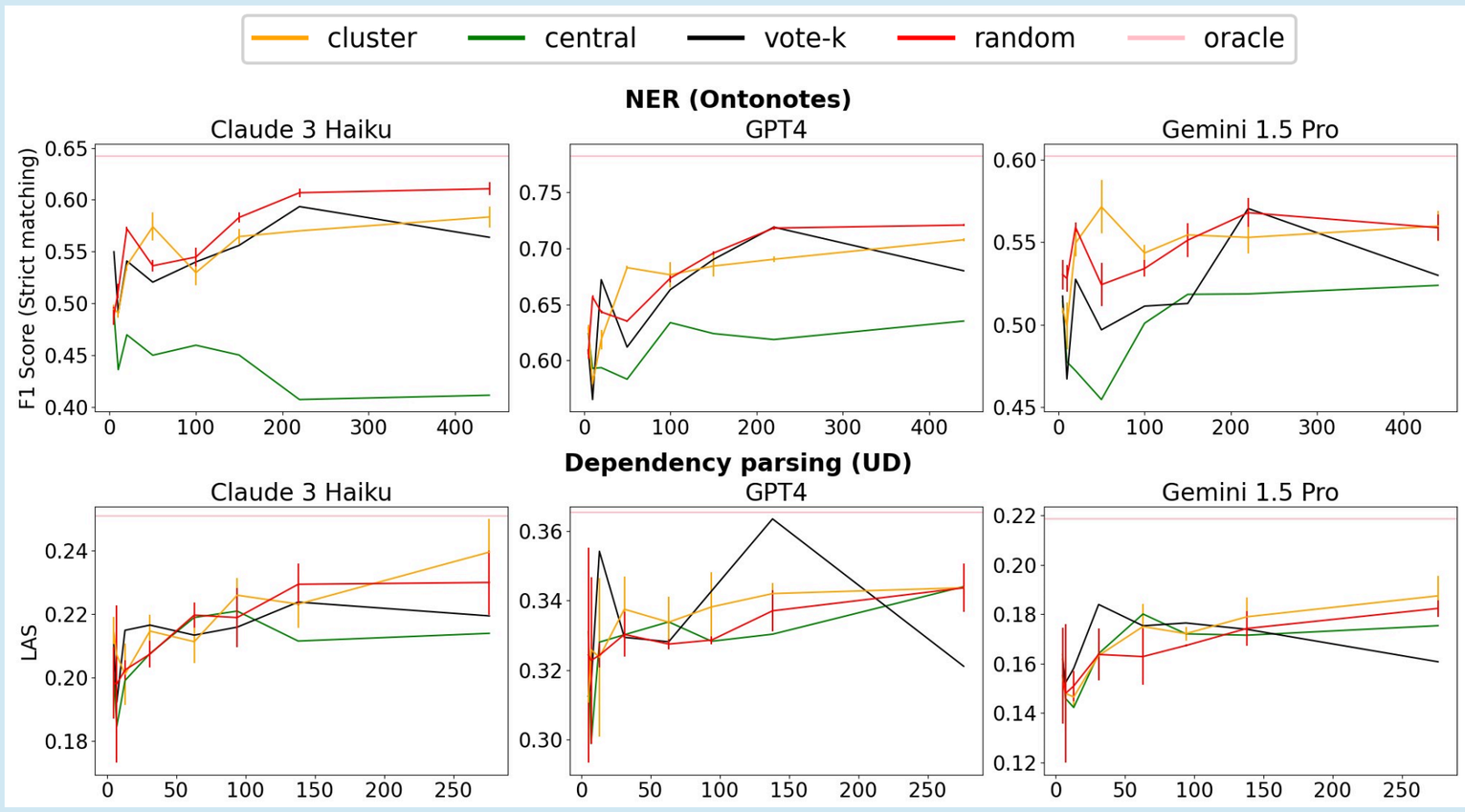
Tasks (token classification)

- Named entity recognition
- Dependency parsing
- Part-of-speech tagging

Models

- Claude 3 Haiku
- GPT4
- Gemini 1.5 Pro

Results



Contact

SCAN ME

References

Su et al. 2022. Selective annotation makes language models better fewshot learners. ArXiv preprint, abs/2209.01975.