

GEOLOCALIZATION WITH MULTIMODAL LARGE LANGUAGE MODEL (MLLM)

Geo-Instruction Tuning

Srikumar Sastry

CSE 561A

OUTLINE

Task Definition

Dataset Creation

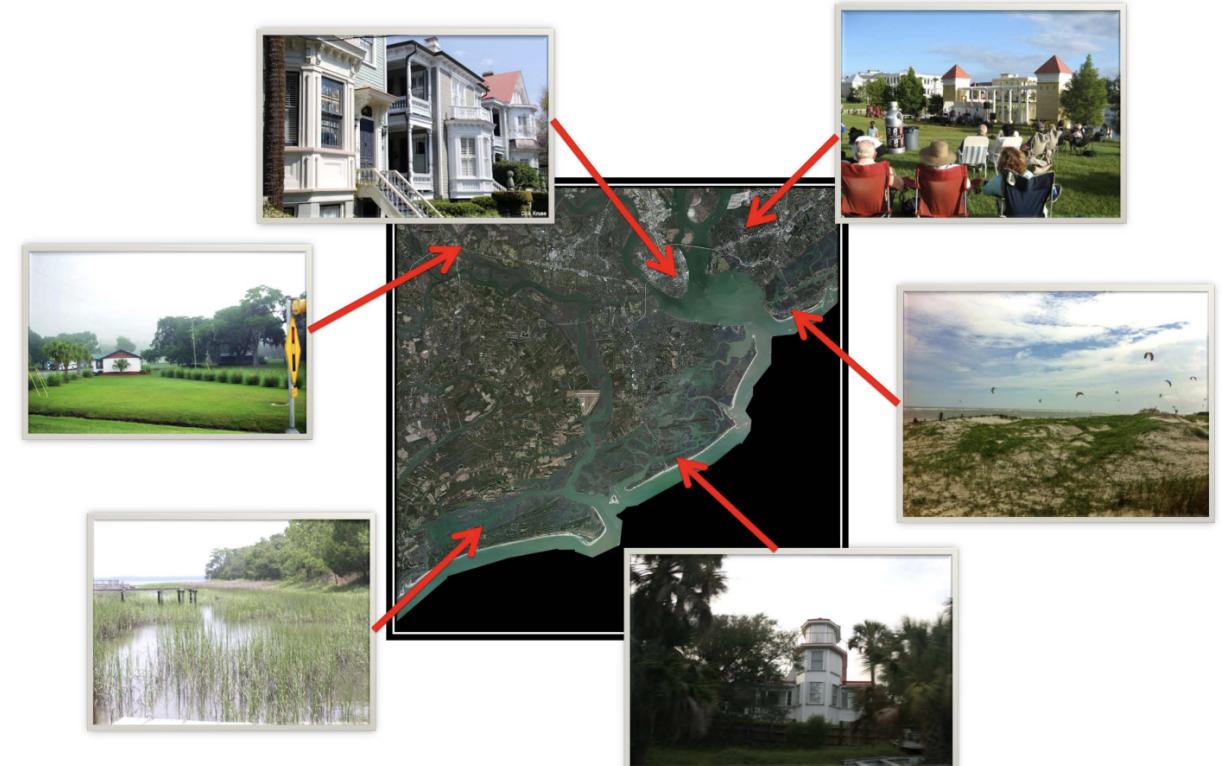
Geo-Instruction Tuning using
QLoRA

Hidden Representation
Regression

Contrastive Training

Future Works and Conclusions

GEOLOCALIZATION



Lin, Tsung-Yi, Serge Belongie, and James Hays. "Cross-view image geolocalization." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2013.

TASK DEFINITION

- Teaching MultiModal Large Language Model to reason about location/geography of an image.
- Current MLLM models insufficient at answering localization queries.

LLaVA Chatbot



Which US state this image comes from?

The image shows a highway with a body of water alongside it, which could be indicative of a coastal state. However, without more specific landmarks or signs, it's not possible to determine the exact state from this image alone. The highway design and the vegetation do not provide enough unique features to pinpoint a specific state. If you have more context or details, I might be able to provide a more accurate answer.

Try to guess a state. Be specific in your answer.

Based on the image, it's difficult to pinpoint a specific state without more context. The highway design, vegetation, and the presence of a body of water do not provide enough unique features to confidently identify a particular state. The image could be from any number of states along the coast, such as California, Florida, or New Jersey, among others. Without more specific details, it's not possible to make an accurate guess.

DATASET CREATION

[39.76025, -105.01829, 39.75998, -105.02099]

State: Colorado County: Denver



The image is a bird's eye view of a residential neighborhood, showcasing the layout of the houses and streets. The neighborhood is filled with trees, and there are several cars parked on the streets. The view is captured from above, providing a comprehensive perspective of the area.

[34.05156, -117.94014, 34.04886, -117.94284]

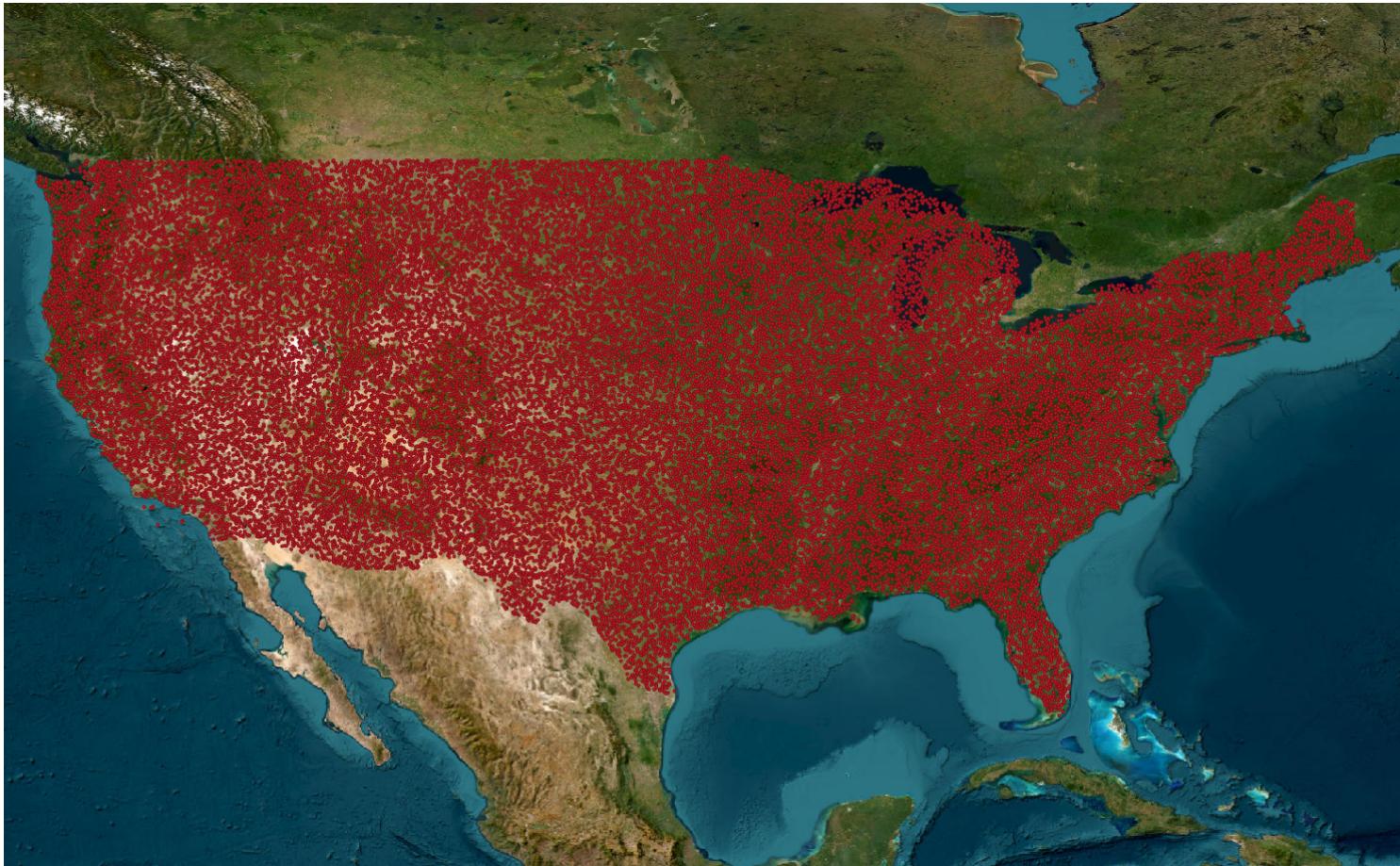
State: California County: San Diego



The image is a bird's eye view of a residential neighborhood, featuring a street with houses and trees. The street is lined with houses, and there are cars parked along the street. The scene captures the essence of a typical suburban neighborhood, with houses and trees surrounding the street.

DATASET SAMPLING

- Distance between samples: **1km**
- Total samples: **50k**
- Train: **45k**
- Validation: **2.5k**
- Test: **2.5k**



INSTRUCTION TUNING – LLAVA-7B

LIU, HAOTIAN, ET AL.
"VISUAL INSTRUCTION
TUNING." ADVANCES IN
NEURAL INFORMATION
PROCESSING SYSTEMS 36
(2024)

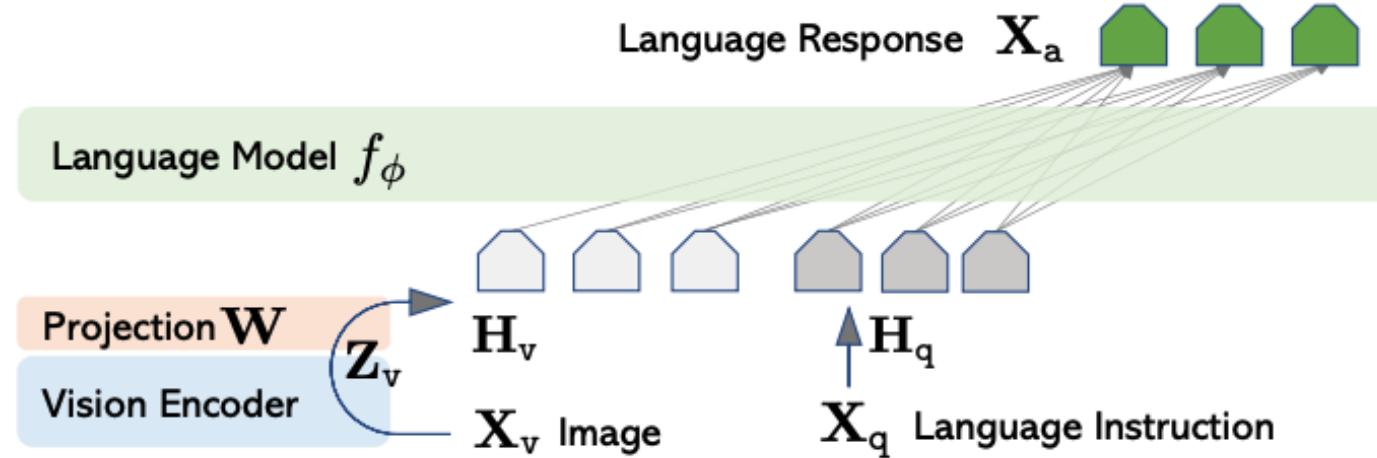


Figure 1: LLaVA network architecture.

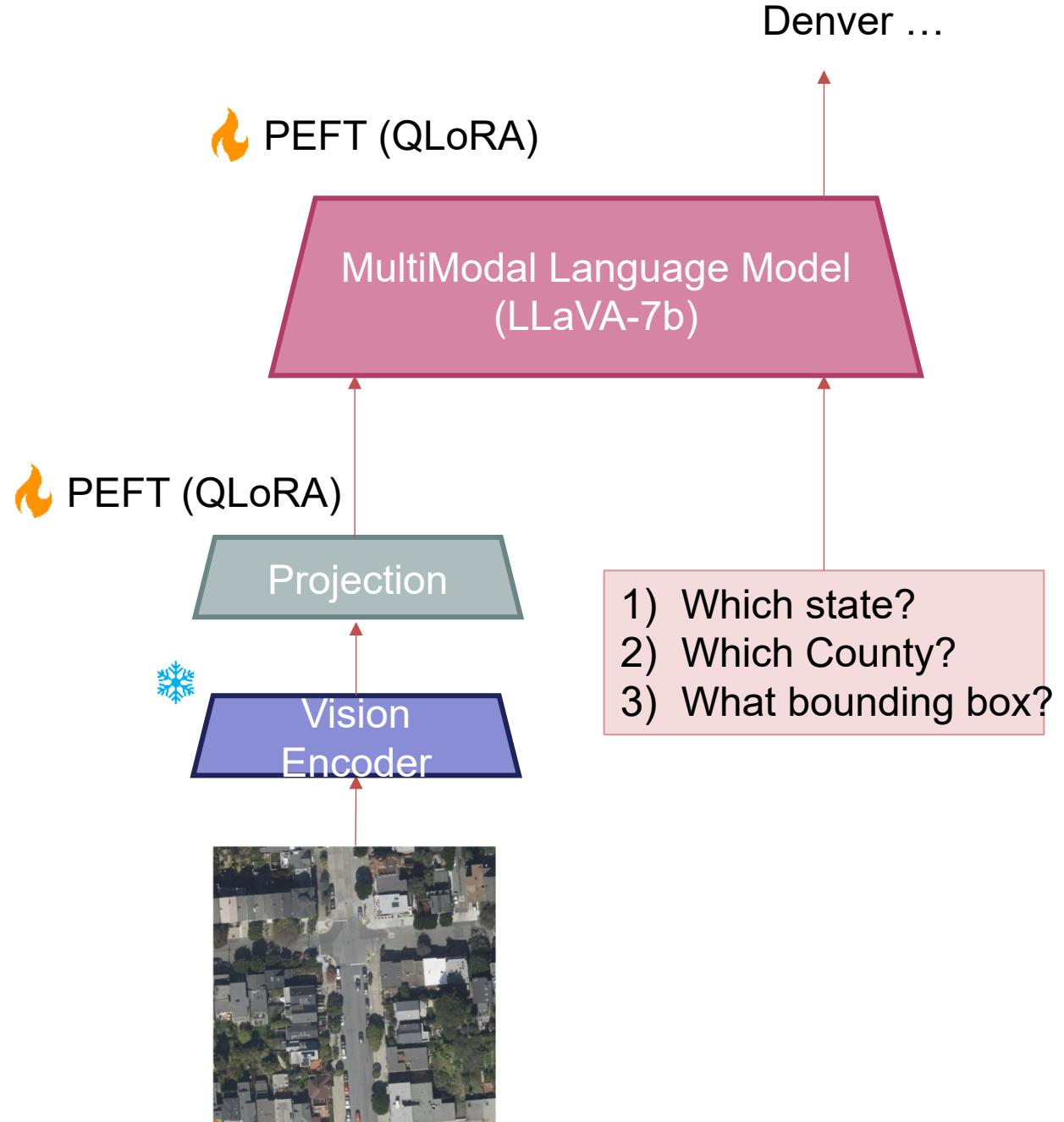
Two stage training framework:

1. Visual feature alignment with the LLM.
2. Fine-tuning End-to-End.

First stage: Freeze everything except the visual projection layer W .

Second stage: Freeze only the vision encoder and train end-to-end.

GEO- INSTRUCTION TUNING



GEO-INSTRUCTION TUNING

HYPERPARAMETER S

QLoRA Training Hyperparameters

Parameter	Value
lora_r	128
lora_alpha	256
bits	4
learning_rate	2e-5

Newly added #parameters: ~5.2 M

Original #parameters: ~7b

% new parameters: ~0.074%

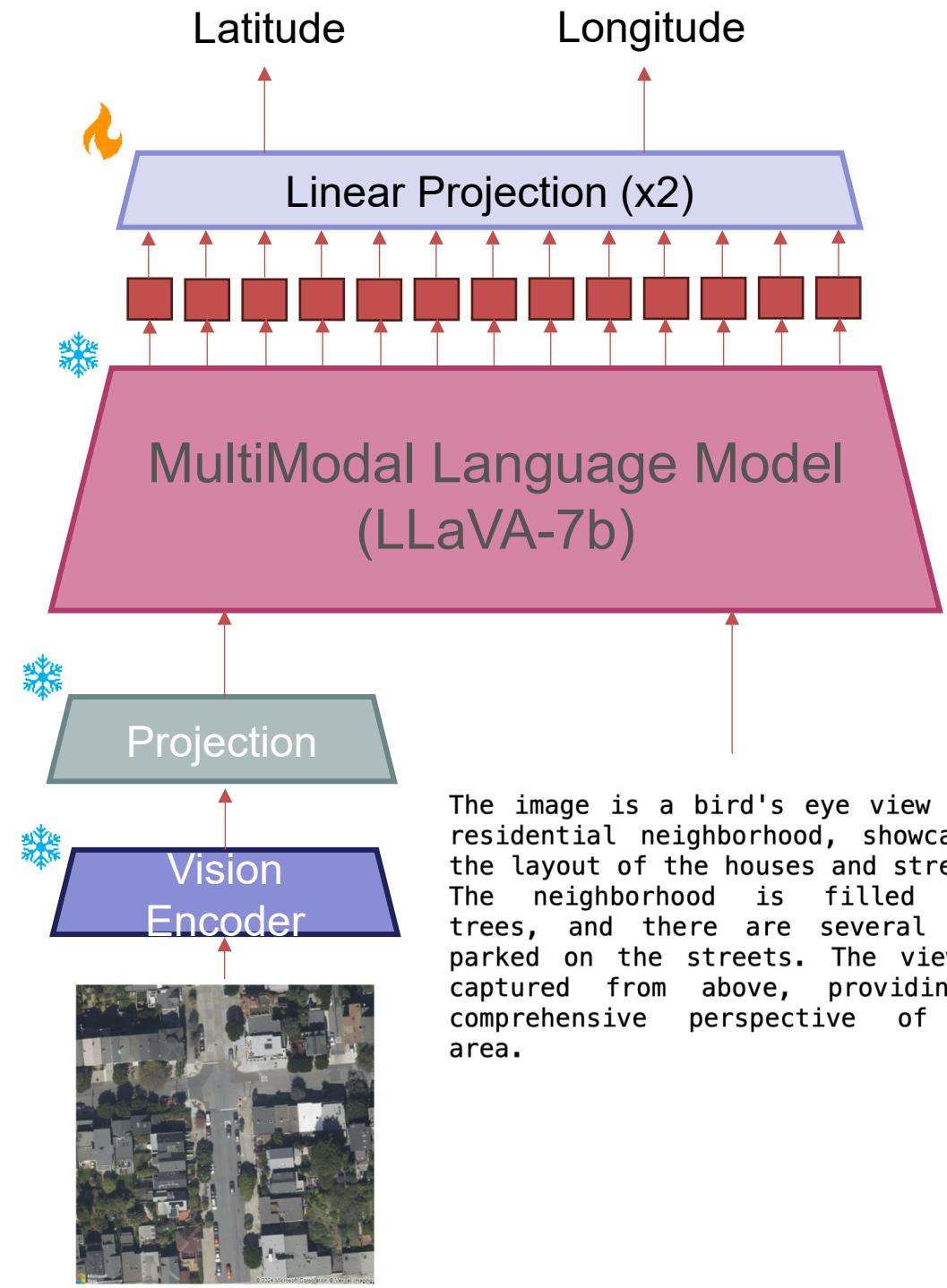
RESULTS

- Output bounding boxes are always between 0-1
- This was original format LLaVA was trained on
- This shows QLoRA is not enough

Category	Accuracy
State	33.28%
County	11.12%
Bounding Box	-

Model failed to predict correct bounding boxes

REGRESSION

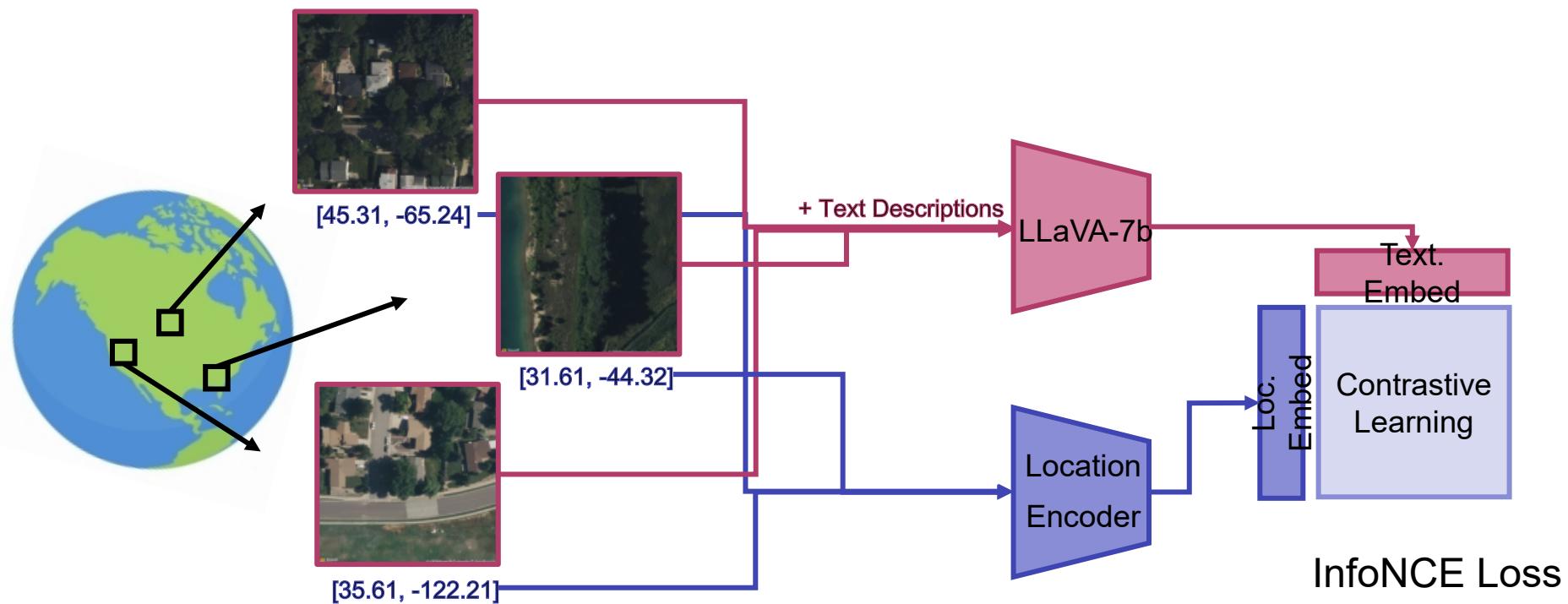




RESULTS

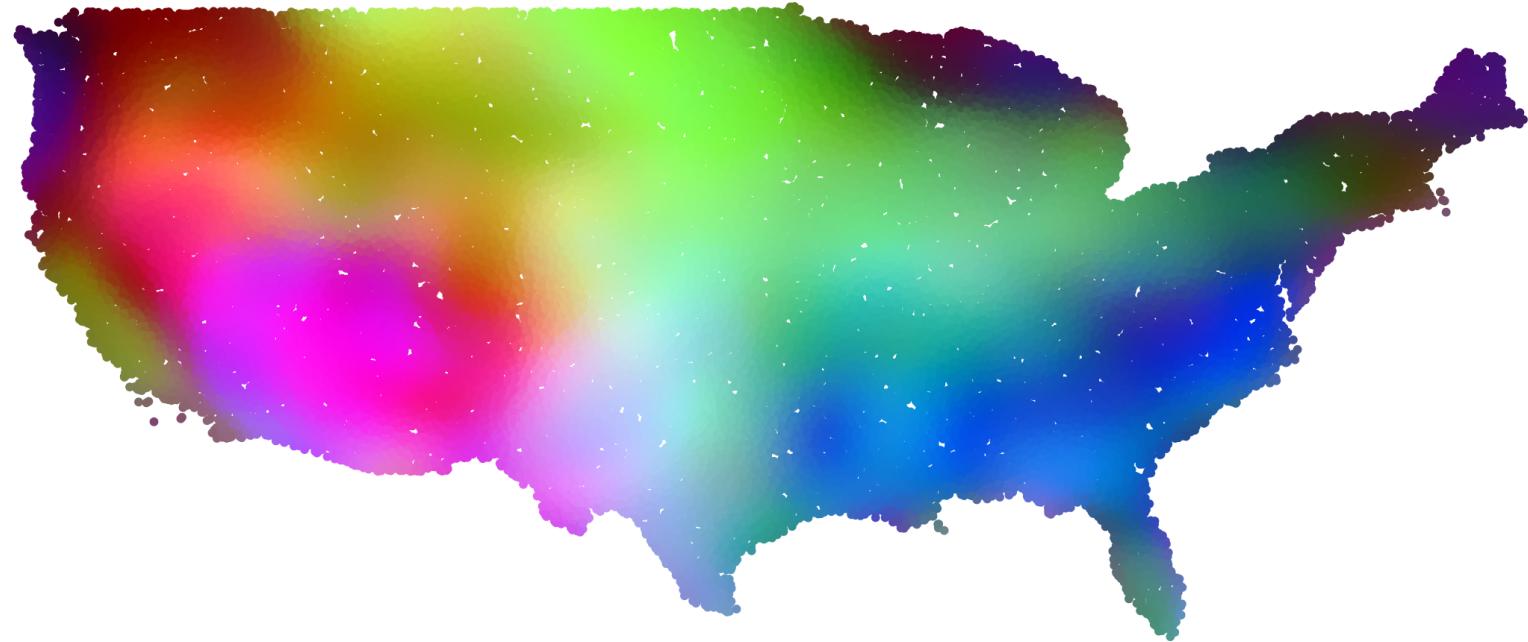
- Intermediate layer hidden representation used
- Loss used: Mean Squared Error
- Accuracy is measured as average geographic distance between prediction and true location.
- Final test accuracy – 2.2^0
- $2.2^0 = 244.44 \text{ km} = 152.77 \text{ miles}$
- Not reasonable results

CONTRASTIVE TRAINING



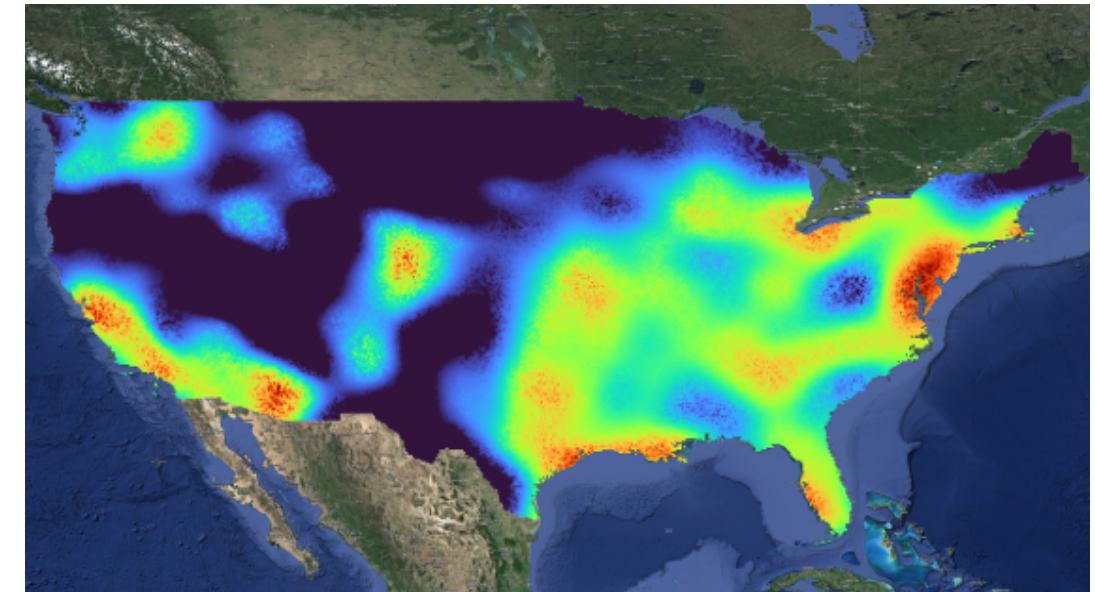
batch_size = 4096, learning_rate = 1e-5, text_masking_ratio = 0.25

ICA OF LOCATION EMBEDDINGS



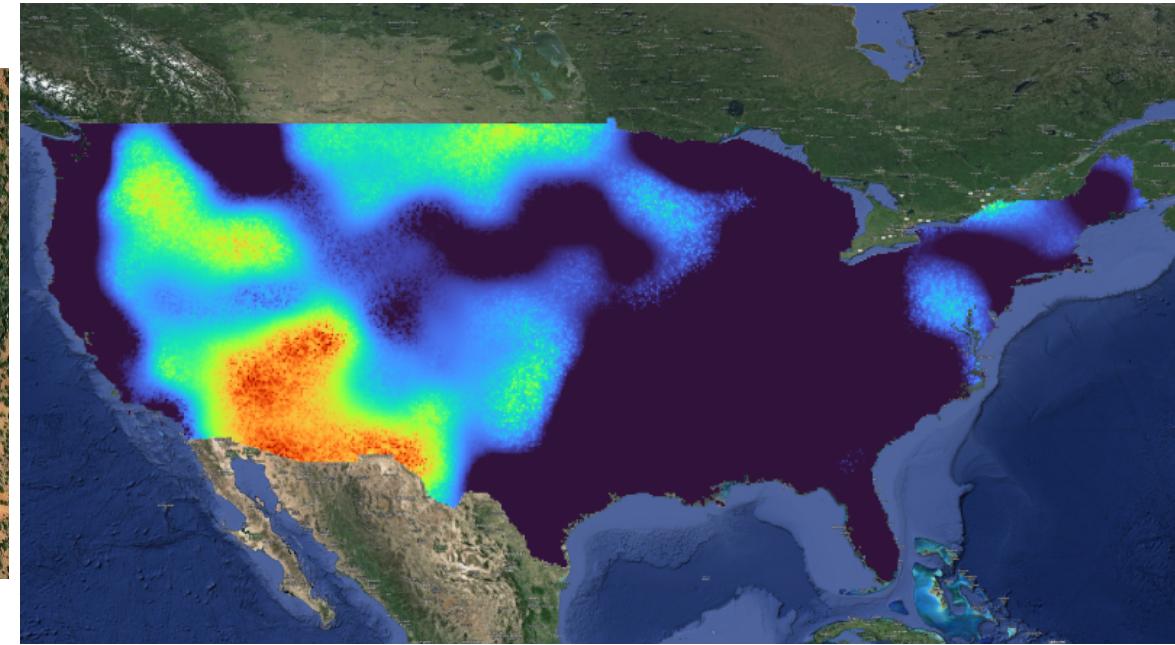
- Embeddings from State-of-the-art SatCLIP Location Encoder
- Low Frequency Information

MODEL OUTPUTS



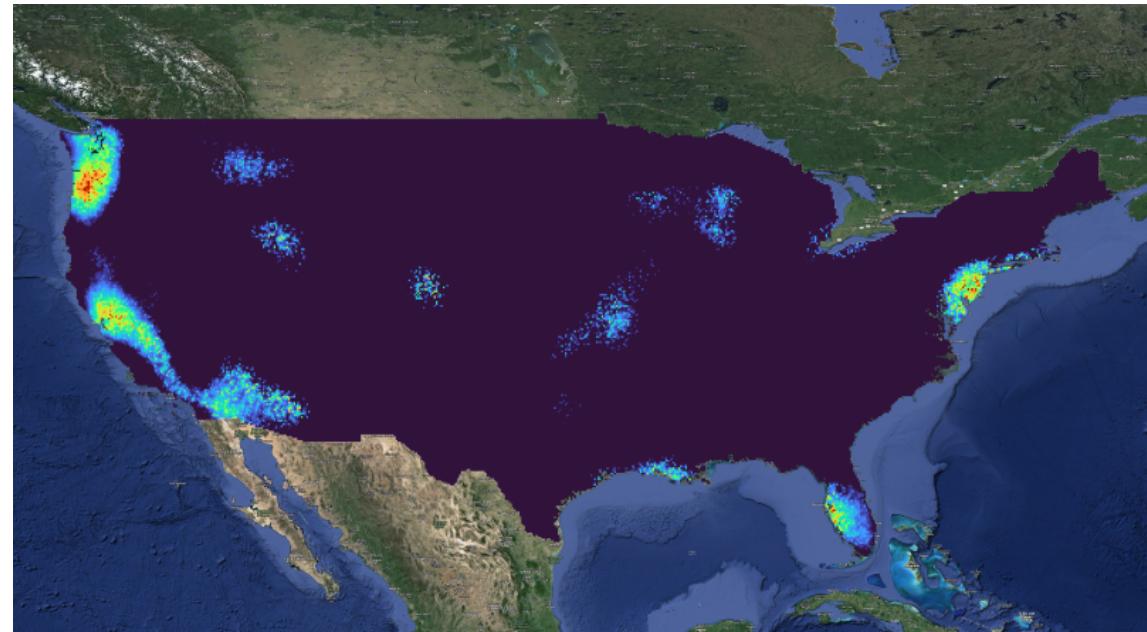
US Capitol

MODEL OUTPUTS



Grand Canyon

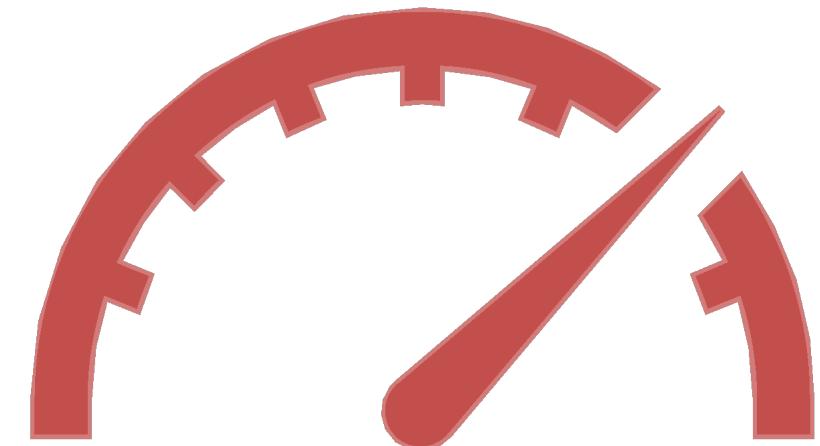
MODEL OUTPUTS



Times Square

RESULTS

- Retrieval Metrics : R@1 for 2500 gallery - 0.32.





FUTURE WORKS

- Full Finetuning of LLaVA-7b using full 32-bit precision
- Convert decoder-only model to encoder model (Parishad, et al. (2024))
- Train on the whole world
- Train using images of different zoom level

CONCLUSI S

- Presented three techniques for geolocation using MLLM.
- The most promising ones are geo-instruction tuning and contrastive learning technique.

Accelerating Inference in Pre-trained LLMs with Dynamic Branching

Alexander Wollam



Background & Motivation

- Transformer-based models have become the predominant model design for language tasks
 - Performance, Scaling
- Many pre-trained LLMs exist
 - GPT, MPT, LLama, ...
- LLMs are expensive to run
 - Parameter count
 - Attention scales quadratically
- Increasing push to making existing LLMs accessible to smaller devices
 - Consumer hardware limited
 - Need to reduce computational footprint of LLMs while retaining speed



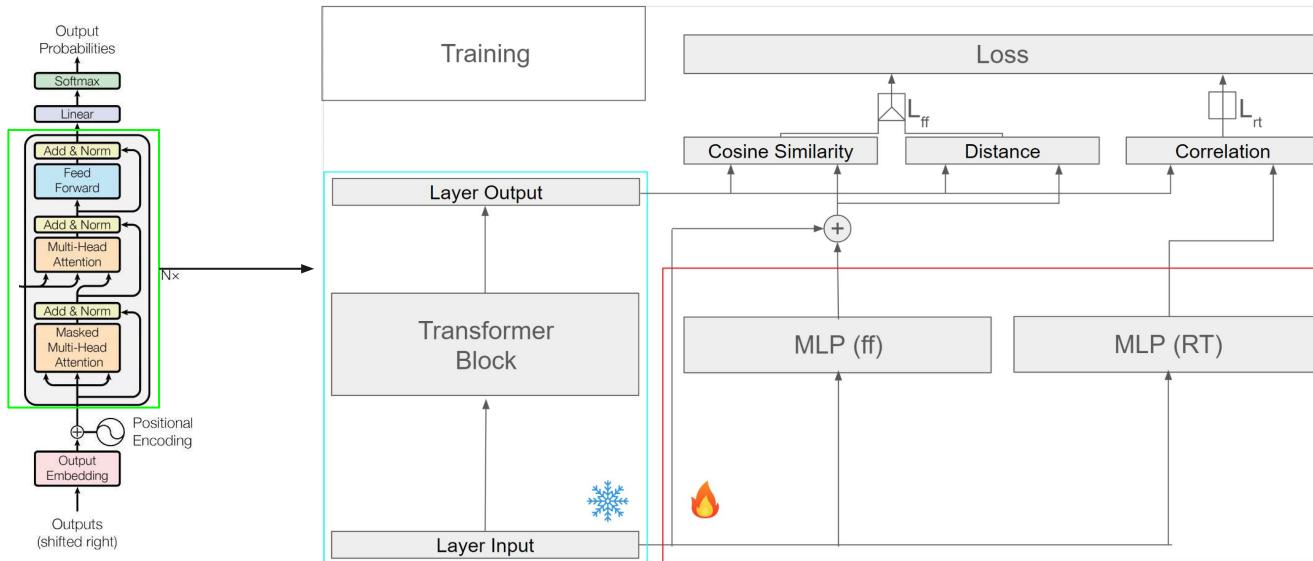
Related Work

- Early Exiting
 - E.g. Depth-Adaptive Transformer, DeeBERT
- Reduced Attention Spans
 - Fixed Spans
 - Learned Spans
- Layer Skipping
 - Residual Connections

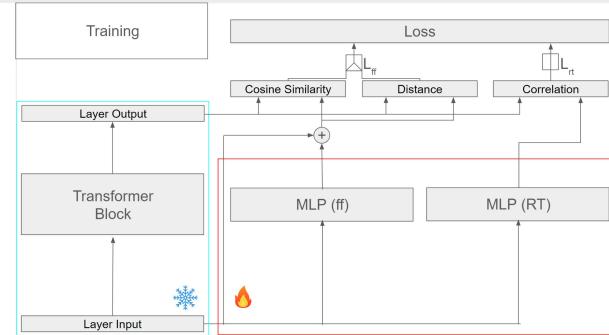
Contribution

- Accelerate inference via a Dynamic Branching Module (DBM)
 - DBM is added alongside each transformer block in LLM
 - Dynamically chooses between transformer block and cheaper MLP branch for each token
- Low-overhead training regime
 - Minimal compute requirements during training
 - Requires no finetuning of the LLM

Implementation: Training



Implementation: Training



Given Layer Input x of length n , Transformer Block TB , MLP ff , MLP RT :

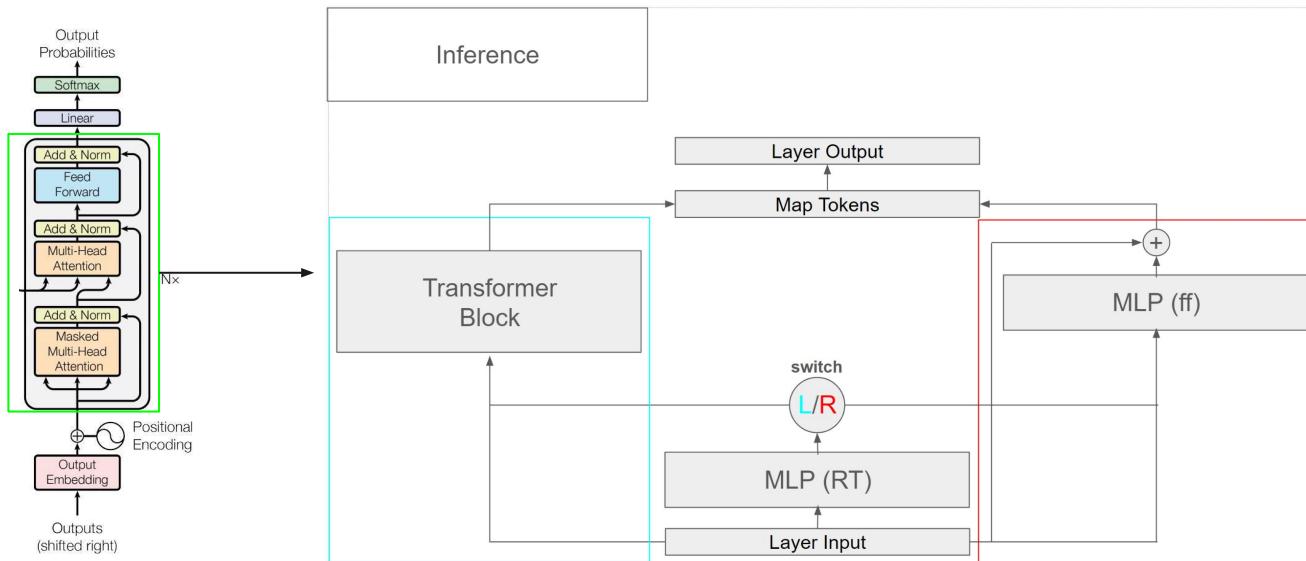
Let $FF(x) := x + ff(x)$; $CS(x) := \text{cosine_similarity}(FF(x), TB(x))$; $ED(x) := \text{euclidian_distance}(FF(x), TB(x))$

Then, $L_{ff} = -\lambda_1 * \frac{1}{n} \sum_{i \in n} (CS_i(x)) + \lambda_2 * \frac{1}{n} \sum_{i \in n} (ED_i(x))$

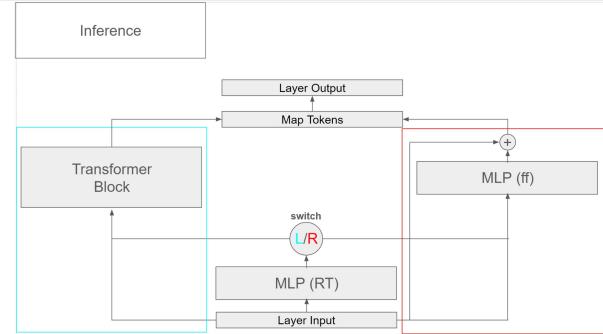
$$L_{rt} = -\lambda_3 * \text{correlation}(RT(x) + \varepsilon, CS(x)) \quad \varepsilon > 0, \varepsilon \in R^n$$

$$\text{Layer_Loss} = L_{ff} + L_{rt}$$

Implementation: Inference



Implementation: Inference



- Switch
 - Use RT output (score) to select whether to branch
 - Thresholding (e.g. branch if $rt_score > threshold$)
 - Probabilistic Sampling w/ Probability Prior (e.g. branch if $rand_num < rt_score * Prob_Prior$)
- Batch Unification
 - Different sequences likely branch different amounts at a given layer
 - Batched input requires same number
 - Solution: Using modified Probabilistic Sampling
 - Generate $rand_num$, and compute $Prob_thresh = rt_score * Prob_Prior$
 - Compute $rand_num < Prob_thresh$
 - Find median number k selected across batch
 - Select top-k of $Prob_thresh - rand_num$ across batch
- Mapping
 - Preserve token ordering when recombining branched and non-branched outputs

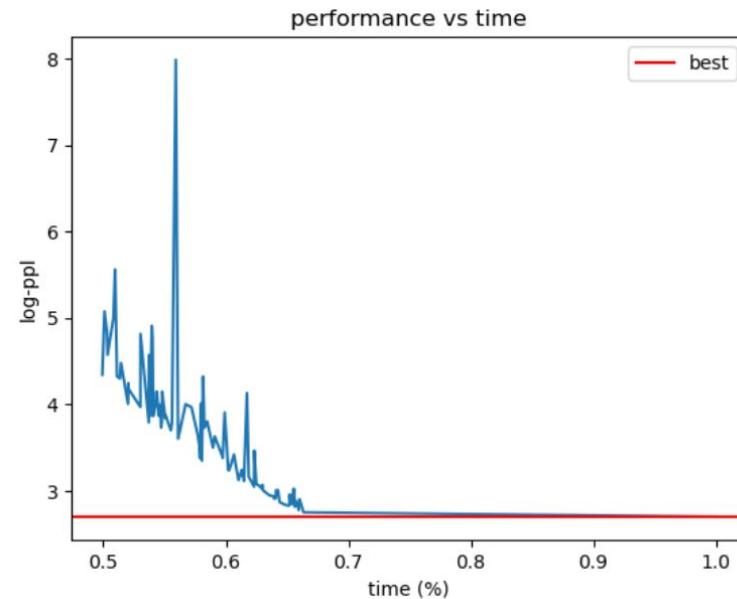


Experiments

- Model: GPT-2-Medium
- Training Dataset: WikiText 2
- Evaluation Datasets:
 - WikiText-2: text generation
 - CNN/Daily Mail: document summarization
 - Filtered to only examples with 1024 tokens total

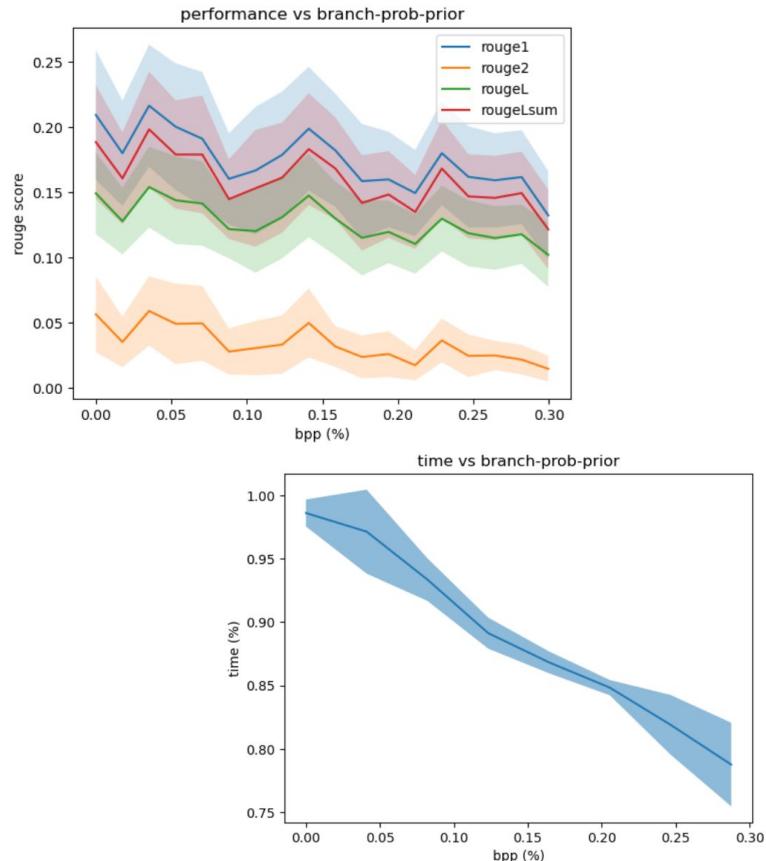
Evaluation: WikiText-2

- log-Perplexity vs Time
 - Perplexity measures the degree of uncertainty of a LM when it generates a new token
 - Time recorded as a fraction of that taken without any branching
- ~33% savings without significant loss in performance



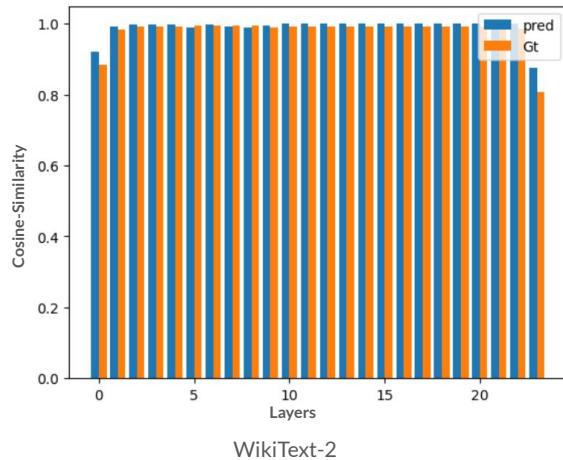
Evaluation: CNN/Daily Mail

- ROUGE-score vs Branch-Prob.-Prior
 - ROUGE scores measure word-similarity between pairs of text
 - Branch Probability Prior is the probabilistic sampling threshold from before
- Less time savings overall
- Performance largely retained with moderate branching

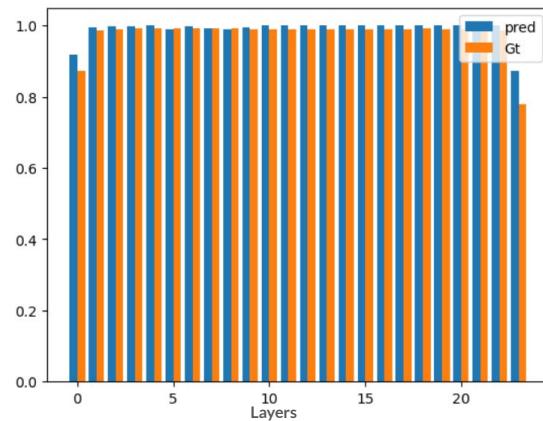


Layer-wise Similarity

- Layer-wise RT-scores comparable
- Branch outputs match similarly
- Attention most leveraged in first and last layer



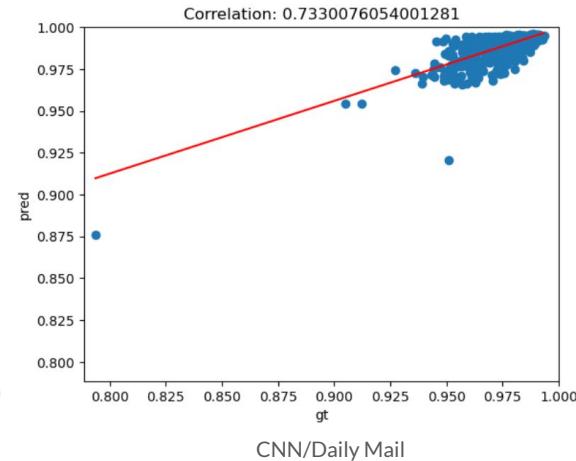
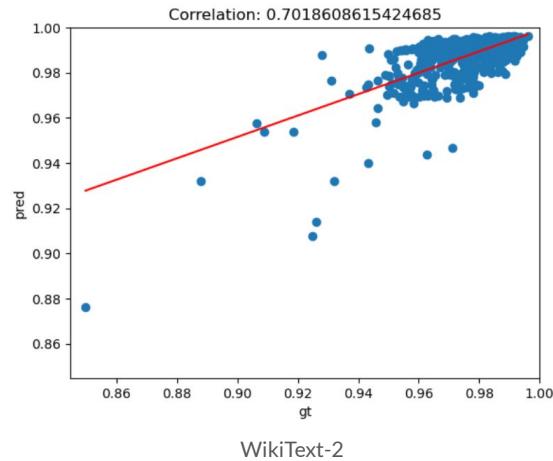
WikiText-2



CNN/Daily Mail

Overall Correlation

- Overall Correlation similar
- RT-scores extend to other tasks well



Conclusion + Future Work

- The proposed branching module successfully allows for significant inference-time savings at marginal performance cost
- It retains decent performance to tasks outside that which it was trained on
- It requires relatively inexpensive training
- Relatively easy to add to pre-trained models



1

Final project: sentiment analysis of customer review

Sihao Lin



CONTENTS

01

background

02

contribution

03

Result& analysis



01

background

Background

Appling LLM for costomers review:

Nowadays, shopping online is a very popular trend for customers and company. So in order to improve the service, understanding the customer review is a very important part to evaluate the quality. By accurately analyzing sentiments expressed in customer reviews, sellers can make informed decisions to enhance product features, customer service, and overall user experience, leading to increased customer satisfaction and loyalty.

Also, extracting detailed sentiment information from reviews can guide product development teams in tailoring new features or addressing shortcomings in existing products based on user preferences and expectations.

Background

Related work on finetuning LLM:

Prefix finetuning :

Prefix Tuning, proposed by Xiang Lisa Li and Percy Liang, is a form of fine-tuning that introduces task-specific information to a pre-trained language model. Unlike traditional fine-tuning, where all parameters of a model are adjusted, Prefix Tuning introduces a small, trainable module, known as the “prefix,” at the beginning of the model’s processing pipeline.

Adapter finetuning:

Adapter is centered around the idea of adding tunable layers to the various transformer blocks of an LLM, as opposed to only modifying the input prompts.

Lora finetuning: LoRA presented an effective solution to this problem by decomposing the update matrix during finetuning.

Background

Goal:

In this project, we use bert-base-cased model as a pretrained model to do the sentiments classification task and use LoRA with SVD as fine-tuning method to update the specific weight metrices in pretrained model.

Motivation:

By using LoRA method, we could finetuning the Bert2 based pretrained model on the down steam task: sentiment analysis of the customer reviews.

Contribution:

we applying SVD to optimize the matrix decomposition part to improve the efficiency of the LoRA finetuning method.



02

contribution

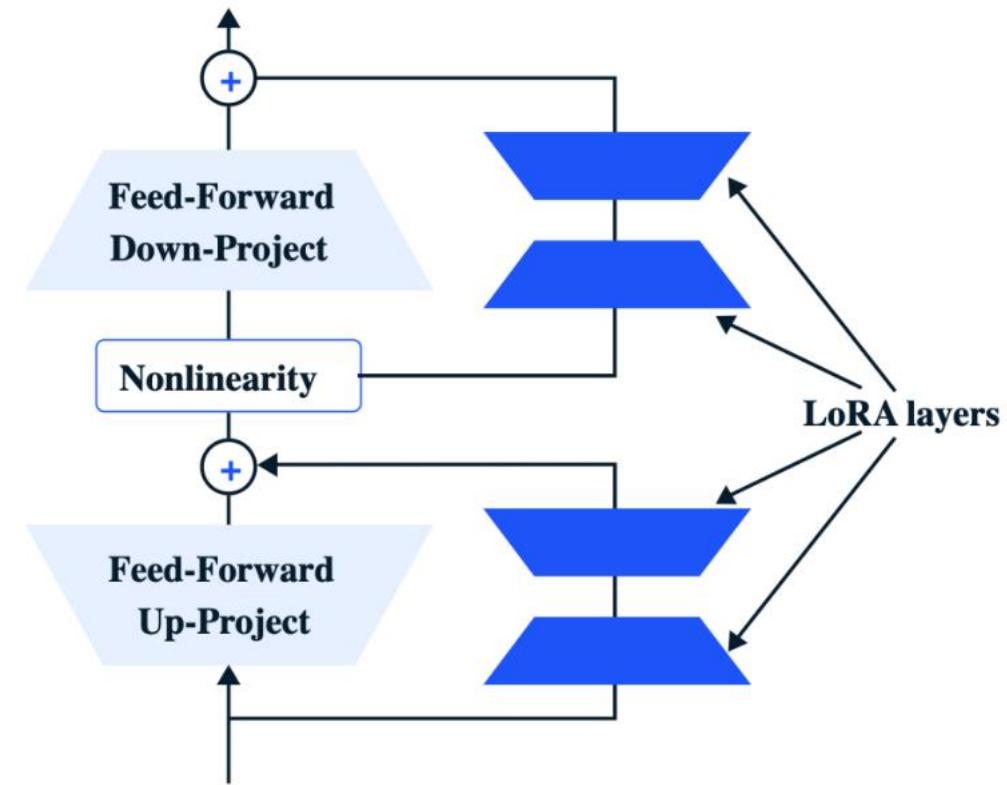
Method design

LoRA finetuning

$$h = W_0x + \Delta Wx = W_0x + BAx$$

By applying LoRA, we reduce the dimension of the updated weight matrix ΔW

The size of B and A is fixed and is chosen with experience.



Method design

LoRA finetuning with SVD

$$\Delta W = U \Sigma V^T$$

$$h = W_0 x + \Delta W x = W_0 x + BAx$$

- Start with a larger rank r , then use gradient descent algorithm to optimize $B(d \times r)$ and $A(r \times n)$
- each m -th iterations, do SVD to find U , Σ , and V . use threshold to cutoff the singular value. Pick first r' th largest singular values over threshold t .
- update the size of B to $(d \times r')$ and A ($r' \times n$)



03

Result and analysis

Result for experiment

Task : sentiment analysis on Amazon review dataset .

Model : bert2-base and training a classifier for the task.

```
100%|██████████| 668/668 [04:57<00:00, 2.25it/s]
Epochs: 1 | Train Loss: 0.763 | Train Accuracy: 0.324 | Val Loss: 0.665 | Val Accuracy: 0.461
100%|██████████| 668/668 [04:57<00:00, 2.25it/s]
Epochs: 2 | Train Loss: 0.579 | Train Accuracy: 0.623 | Val Loss: 0.388 | Val Accuracy: 0.924
100%|██████████| 668/668 [04:57<00:00, 2.25it/s]
Epochs: 3 | Train Loss: 0.317 | Train Accuracy: 0.908 | Val Loss: 0.202 | Val Accuracy: 0.962
100%|██████████| 668/668 [04:57<00:00, 2.25it/s]
Epochs: 4 | Train Loss: 0.183 | Train Accuracy: 0.971 | Val Loss: 0.123 | Val Accuracy: 0.969
100%|██████████| 668/668 [04:57<00:00, 2.25it/s]
Epochs: 5 | Train Loss: 0.115 | Train Accuracy: 0.981 | Val Loss: 0.088 | Val Accuracy: 0.971
```

Example of the output

truth: {'label': 1, 'title': 'Stuning even for the non-gamer', 'content': 'This sound track was beautiful! It paints the senery in your mind so well I would recomend it even to people who hate vid. game music! I have played the game Chrono Cross but out of all of the games I have ever played it has the best music! It backs away from crude keyboarding and takes a fresher step with grate guitars and soulful orchestras. It would impress anyone who cares to listen! ^_-'}

predict: {'label': ' POSITIVE', 'score': 0.9990288019180298}

Result for experiment

Accuracy of LoRA

	Before fine-tuning	LoRA with fixed size	LoRa with SVD
Accuracy	0.972	0.971	0.972

Analysis:

- we try the fixed size of the decomposed matrices columns as 4 and the result after applying LoRA shows the accuracy on validation set remained slightly different.
- when we try to do the SVD, the result becomes unstable and sometimes performing worse than the pretrained model.
- The main reason might be SVD perform not so well on the sparse matrix or the threshold is not chosen properly.

A stylized globe of Earth is centered in the background, showing a network of glowing blue and yellow lines that represent global communication or data flow. The globe is set against a dark, star-filled space with several bright stars visible. In the lower right quadrant, the words "THANK YOU" are written in large, bold, white capital letters.

THANK YOU

Improving Medical Error Detection and Correction

Kriti Bhattacharai

CSE 561 Large Language Models

Table of Contents

- Motivation
- Literature
- Methods
- Experimental Results
- Discussion/ Limitations
- Conclusion

Motivation

- Ensuring the accuracy of clinical text is important for various stages of natural language processing (NLP) model development, from pre-training and fine-tuning to inference, ultimately impacting patient care.
- Physicians operate under time constraints, and may introduce spelling errors or other inaccuracies into their written records
- Moreover, the increase of language model-generated text introduces another layer of potential errors, with the risk of hallucinated or inaccurate information being incorporated into clinical text.
- Critical to identify and rectify errors in clinical text for reliability of NLP models deployed in medical settings.

Literature

A Resource for Detecting Misspellings and Denoising Medical Text Data

Enrico Mensa*
Università di Torino,
Dipartimento di Informatica

Gian Manuel Marino[†]
Università di Torino,
Dipartimento di Informatica

Davide Colla*
Università di Torino,
Dipartimento di Informatica

Matteo Delsanto*
Università di Torino,
Dipartimento di Informatica

Daniele P. Radicioni*
Università di Torino,
Dipartimento di Informatica

*{firstname.surname}@unito.it; [†]marino.jnf@gmail.com

CREAK: A Dataset for Commonsense Reasoning over Entity Knowledge

Yasumasa Onoe, Michael J.Q. Zhang, Eunsol Choi, Greg Durrett
The University of Texas at Austin
{yasumasa, mjqzhang, eunsol, gdurrett}@cs.utexas.edu

SemEval-2020 Task 4: Commonsense Validation and Explanation

**Cunxiang Wang^{1,2*}, Shuailong Liang^{3*}, Yili Jin⁴, Yilong Wang¹,
Xiaodan Zhu⁵ and Yue Zhang²**

¹Zhejiang University, China; ²School of Engineering, Westlake University, China

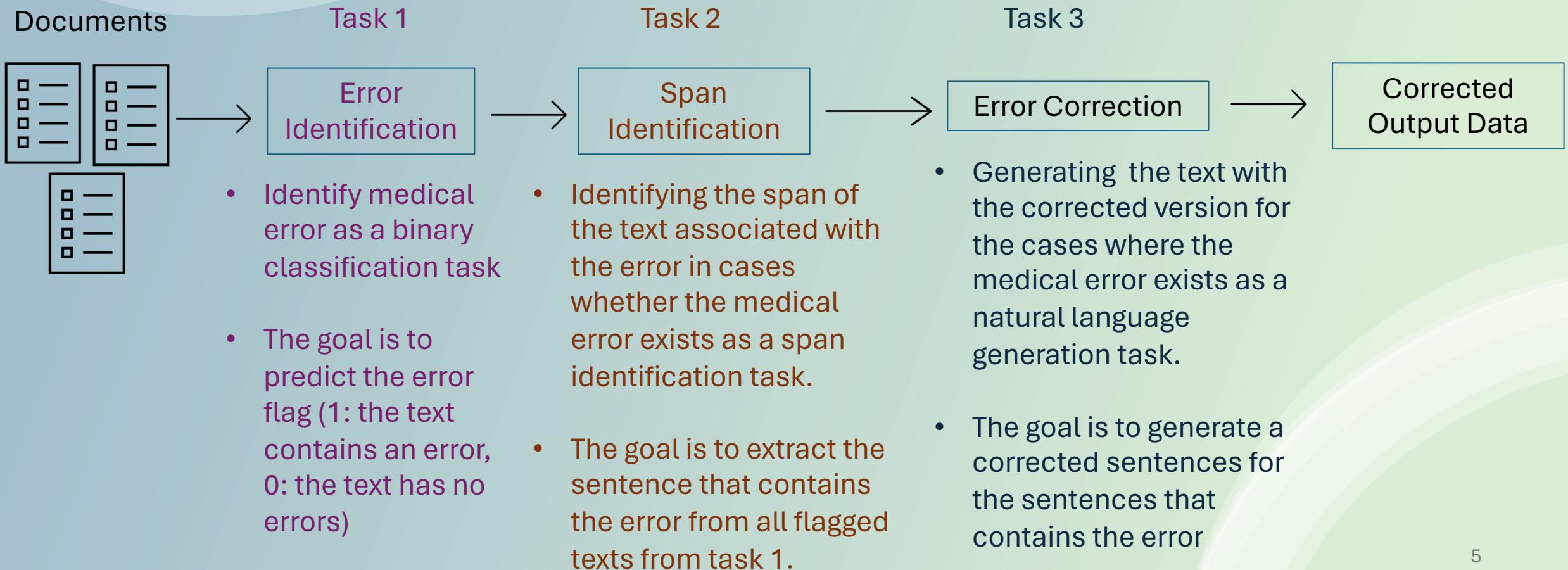
³Singapore University of Technology and Design, Singapore;

⁴Sun Yat-Sen University, China; ⁵Queen's University, Canada

wangcunxiang, zhangyue@westlake.edu.cn,
shuailong_liang@mymail.sutd.edu.sg, zhu2048@gmail.com

Framework

Improving error detection by identifying and correcting medical errors in clinical text



Dataset

- Deidentified data released by Microsoft Health Futures and University of Washington as part of NAACL 2024
- Contains 2,189 clinical texts to be used for training , 574 clinical texts for validation, another 160 texts for validation

Sample Data

Text-id	Text	Sentences	Error Flag	Error Sentence ID	Error Sentence	Corrected Sentence	Corrected Text
ms-train-0	"A 53-year- old male patient comes to the physician because...He works as a commercial fisherman on Lake..."	<p>0 A 53- 1 year-old male pa- tient comes to the physician because...</p> <p>1 He works as a com- mercial fisherman on hospital...</p>	1	6	<i>He works as a commercial fisherman on hospital...</i>	<i>He works as a commercial fisherman on the Lake</i>	<i>A 53-year-old male patient comes to the physician because...He works as a commercial fisherman on Lake</i>

Task 1: Error Detection

- **Problem Formulation:**

Given a text segment X_i consisting of N sentences, we aim to predict the error flag y_i (0 or 1) indicating whether there is an error in the text.

- Implementation

- Tokenize
- Utilize a pre-trained BERT model (e.g., clinicalBioBERT) for error detection

```
clinicalBioBERT(inputText)=Transformer Encoder(Token Embeddings(input))
```

- Fine-tune BERT-based error detection model using annotated data
- Inference / Test to predict the error flag y_i

Task 2 : Span Identification

- Split the text by sentences
- Prepare the tokenized sequences as input to the BERT model, including special tokens [CLS] at the beginning and [SEP] between sentences.

```
clinicalBioBERT(inputText)=Transformer Encoder(Token Embeddings(input))
```

- The model predicts the probability of each sentence containing an error
- Assemble the extracted sentences from all flagged texts to obtain the final set of sentences containing errors

Task 3 : Error Correction

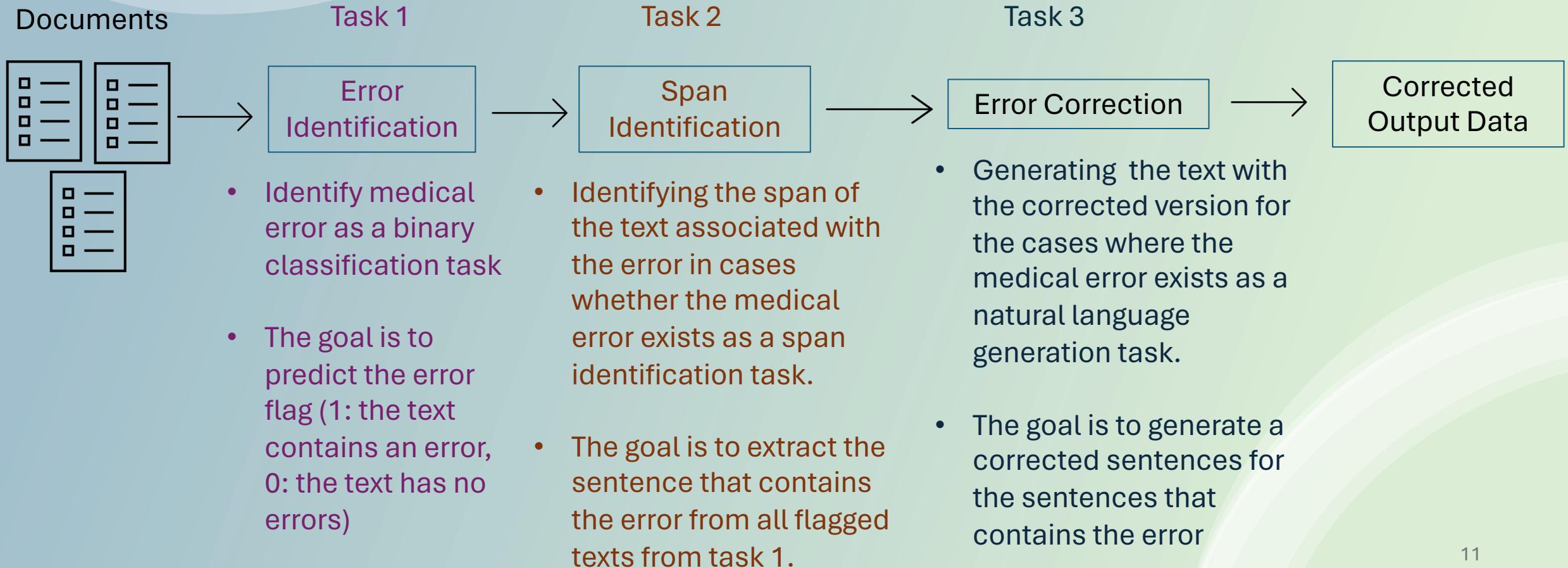
- We use generative model in proposing corrections for erroneous text.
- After identifying errors in the generated text, model can suggest corrections by applying edit operations to transform erroneous text into correct text.
- Inference task:
 - Specify prompts to replace the error in text with the correct text
 - GPT-3.5, GPT-4, Flan-T5, Flan-T5-xxl, Flan-T5-xl

$$S_{\text{corrected}} = \text{LLM}(S_{\text{error}}, \text{Prompt})$$

$$P(S_{\text{corrected}} | S_{\text{error}}, \text{Prompt}) = \prod_{i=1} P(w_i | w_{1:i-1}, S_{\text{error}}, \text{Prompt})$$

Framework

Improving error detection by identifying and correcting medical errors in clinical text



Model Baselines

- We compared the performance across multiple large language models including GPT-4, GPT-3.5 and Flan-T5 for error detection, span identification and error correction
- Evaluation:
 - Error Flag accuracy
 - Error Sentence Detection Accuracy
 - Composite ROUGE-1-F, composite rouge-2-F, composite rouge-l-F

Performance results

Table 2. Error Flag Accuracy and Error Sentence Detection Accuracy for a subset of the data as a few-shot task

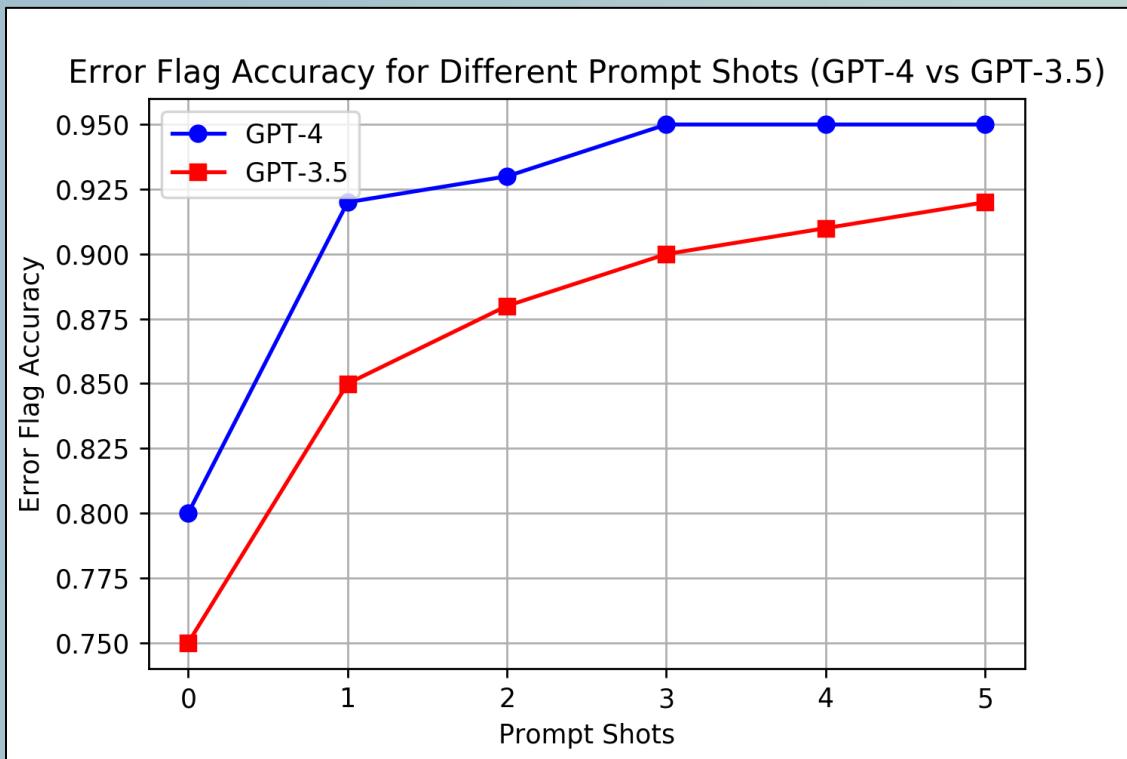
Model	Accuracy	
	Error Flag	Error Sentence
GPT-4	<u>0.95</u>	<u>0.90</u>
GPT-3.5	0.92	0.88
Flan-T5-xxl	0.88	0.85
Flan-T5-xl	0.86	0.85

Table 3. Composite ROUGE Scores for the same subset of the data

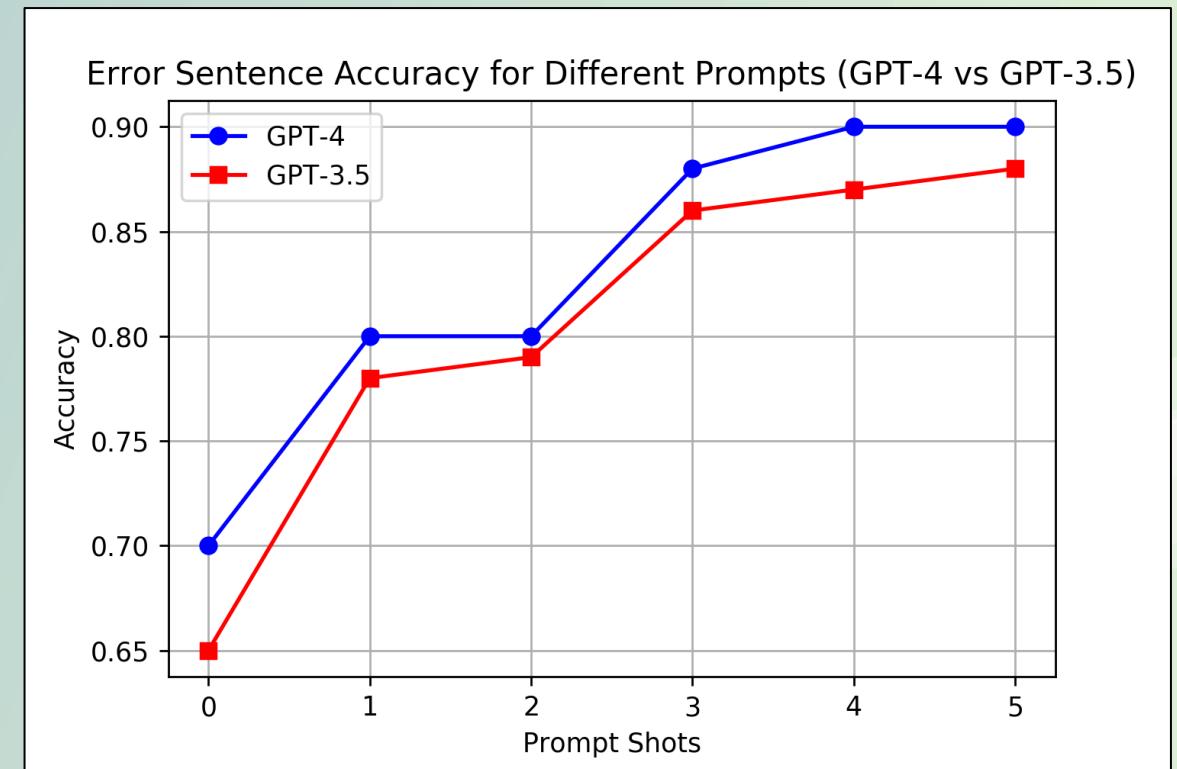
Model	ROUGE-1-F	ROUGE-2-F	ROUGE-L-F
GPT-4	0.85	0.75	0.80
GPT-3.5	0.80	0.70	0.75
Flan-T5-xxl	0.75	0.65	0.70
Flan-T5-xl	0.77	0.65	0.65

Analysis

Error Flag Accuracy



Error Sentence



Limitation

- Dependency on the cost-intensive pre-trained language models
- Limited training data to capture all contexts
- Complexity of error correction
- Evaluation metrics limitation

Discussion

- Built a framework for medical error detection and correction
- We use clinicalBioBERT and large language models to identify the incorrect phrases within the text and correct them
- With the most recent LLMs, we are better able to identify and correct the medical errors

References

- [1] Zhao Y., et al. Gemini in Reasoning: Unveiling Commonsense in Multimodal Large Language Models. *Arxiv*. 2023.
- [2] Zhang Y., et al. SemEval-2020 Task 4: Commonsense Validation and Explanation. *ACL*. 2023
- [3] Natarjan V., et al. Large language models encode clinical knowledge. *Nature*. 2023.
- [4] Jain A., et al. Structured information extraction from scientific text with large language models. *Nature*. 2024
- [5] Zhou L., et al. Automated misspelling detection and correction in clinical free-text records. *Journal of Biomedical Informatics*. 2015
- [6] Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Res* 2004; 32: D267-270
- [7] Kasmaiee Sa., et al. Correcting spelling mistakes in Persian texts with rules and deep learning methods. *Nature*. 2023.
- [8] Hu Y., et al. Misspelling Correction with Pre-trained Contextual Language Model. *Arxiv*. Jan 2021

Thank you!



News through the Prism of LLMs: Comprehending Complex Contexts to Construct Commonsense Knowledge Graphs

Yinxu Tang, Haoyu Yin and Chengsong Huang

Machines that Understand Humans

It's going to snow. I'll have to wake up 30 minutes earlier.



- Leave out important information
- Rely on shared intuitive inferences to establish context
- Reason about commonsense knowledge to reach understanding

Machines that Understand Humans

It's going to snow. I'll have to wake up 30 minutes earlier.

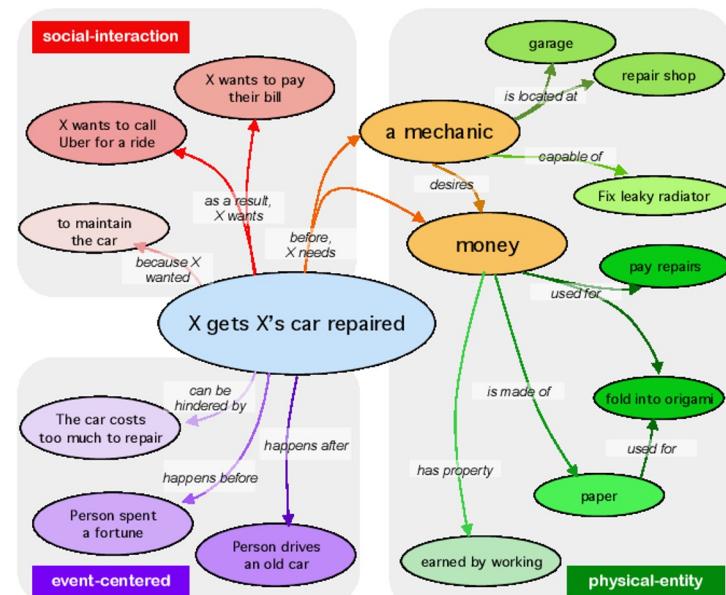
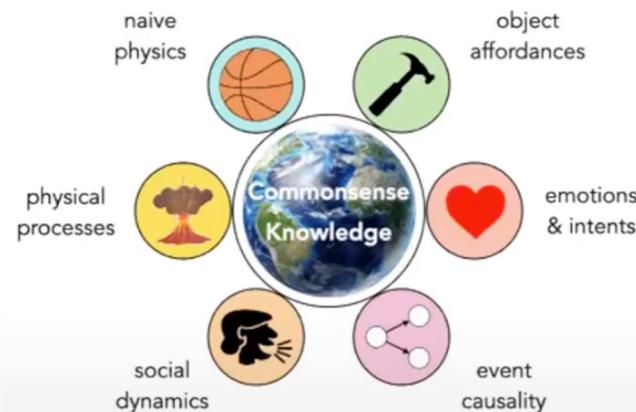


- Leave out important information
- Rely on shared intuitive inferences to establish context
- Reason about commonsense knowledge to reach understanding

How can we get machines to read, understand, and use language the same way humans do?

Commonsense Knowledge Graphs (CSKG)

Connecting components of world knowledge
to understand situations.



Traditional Methods

Traditional Methods for KG construction:

supervised learning, requiring extensive annotated data and adaptation for new domains.

Issues:

- High cost and time consumption.
- Limited scalability across different domains.

Modern Approaches

Methods using LLMs:

automatically process vast amounts of textual data to extract knowledge.

Issues:

- Lack of contextual understanding.
- Handling uncertainty and brittleness in reasoning.
- Limited commonsense knowledge coverage.

Constructing Commonsense Knowledge Graphs

Goal:

- Introduce a structured pipeline for creating a commonsense KG using advanced LLM techniques.
- Focus on leveraging causal relations extracted from the Causal News Corpus (CNC).

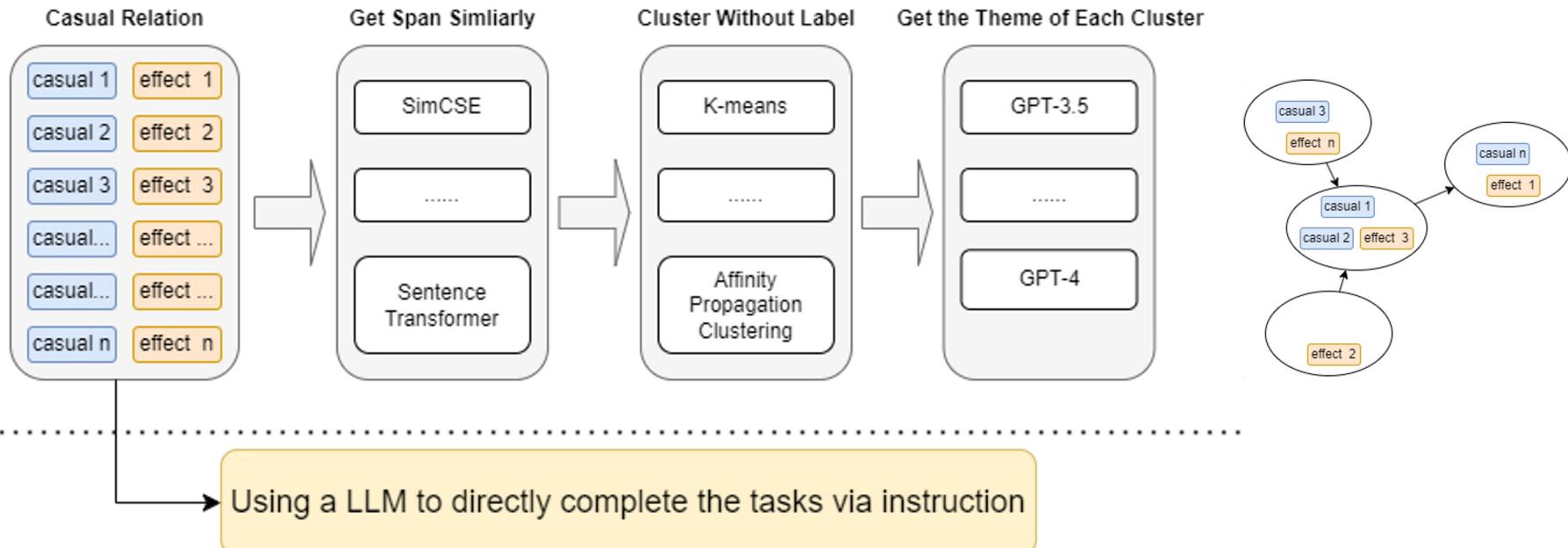
Key Tasks:

- Data Acquisition
- Extraction of Commonalities
- KG Construction

Impact:

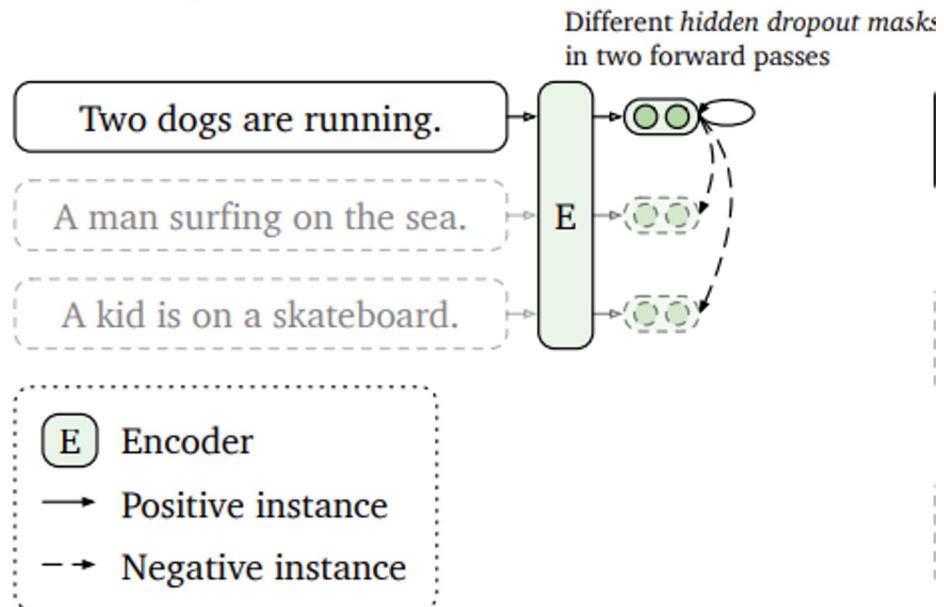
Allow for more sophisticated, natural interactions between AI systems and users, mirroring human cognitive processes.

Method – Pipeline



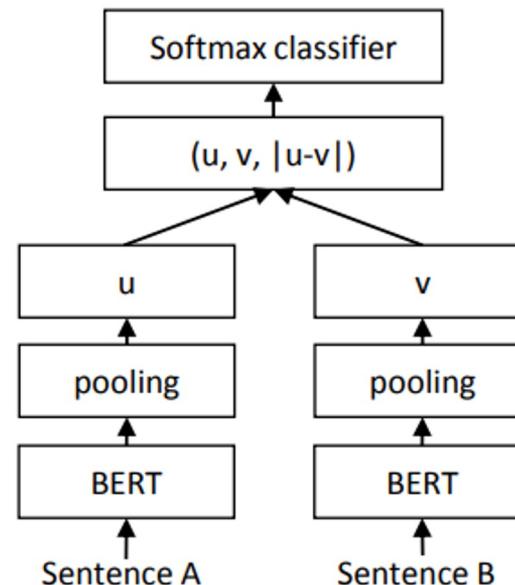
Method – SimCSE

We use a pretrained RoBERTa trained on SimCSE as the first methods to get the span representation



Method – Sentence Transformer

The second model we used is sentenceBERT, which is trained on some sentence-pair-level tasks like NLI.



Get the Theme via LLMs

We designed a prompt to instruct the LLMs to give the theme of a cluster of sentences.

Analyze the following phrases to extract and summarize the underlying semantic themes, the output should be in json format, please only add the themes in the json, you should only output `` `one` ` ` span including two or three words to show the commonalities of `` `all these sentences`` `:

- 1 a service delivery protest
- 2 Dissatisfied with the package
- 3 protests over poor service delivery



ChatGPT

json

Copy code

```
{  
  "common_theme": "service delivery protest"  
}
```

Experiment Details

- Data Collection:
 - Causal News Corpus (CNC)
 - Each data can be split according to causal relationship
- Baselines:

Sentence Transformer

SentenceBERT
SimCSE
SnowflakeArctic



Cluster

Affinity Propagation (AP)
K-means



Theme Summarization

GPT-3.5

K is decided by Elbow Method

Experiment Results

<i>Baseline Methods</i>	<i>Sentence Transformer</i>	<i>Cluster</i>	<i>Theme Summarization</i>	<i># of Causal Clusters</i>	<i># of Effect Clusters</i>
1	SentenceBERT	AP	GPT-3.5	254	249
2		K-means		150	150
3	SimCSE	AP	GPT-3.5	212	181
4		K-means		130	140
5	SnowflakeArctic	AP	GPT-3.5	217	192
6		K-means		110	120

Experiment Results – A Good Example

Boycotting the classes and medical services, Lack of medical services, rivalry among nursing homes, has affected hundreds of patients, against the increasing attacks on the medical community, the government appearing to be in no mood to meet the demand of the doctors of the health service, hospital development

Challenges in healthcare system

Cluster Results

Theme Summarization

Experiment Results – A Bad Example

discussions with officials, she explained her views, hearing about the incident, the information was spread in the area, to discuss the incident, inquiry, to enquire about the incident and meet the injured party workers ., to get an urgent interdict, to address issues of concern emanating from the incident, to stress her point, to show that its image had not been affected in any manner, to press their demands for setting up AIIMS at Raha and a judicial inquiry into the death of a person in clashes between police and protesters there three days back, to discuss their course of action

Cluster Results

discussions, explanations, incident inquiry, urgent action, addressing concerns, stressing points, demands pressing, course of action

Theme Summarization



Future Work

- Evaluation:
 - The effectiveness of cluster
 - The trade-off between the granularity and simplicity of a KG
- Use LLM to complete a end-to-end pipeline