CRCV REU Week #5

Project: Geo-HiPO

Advisors:

- Rohit Gupta
- Parth Kulkarni
- Dr. Shah

David Rivard

06/27/2025

Project Overview



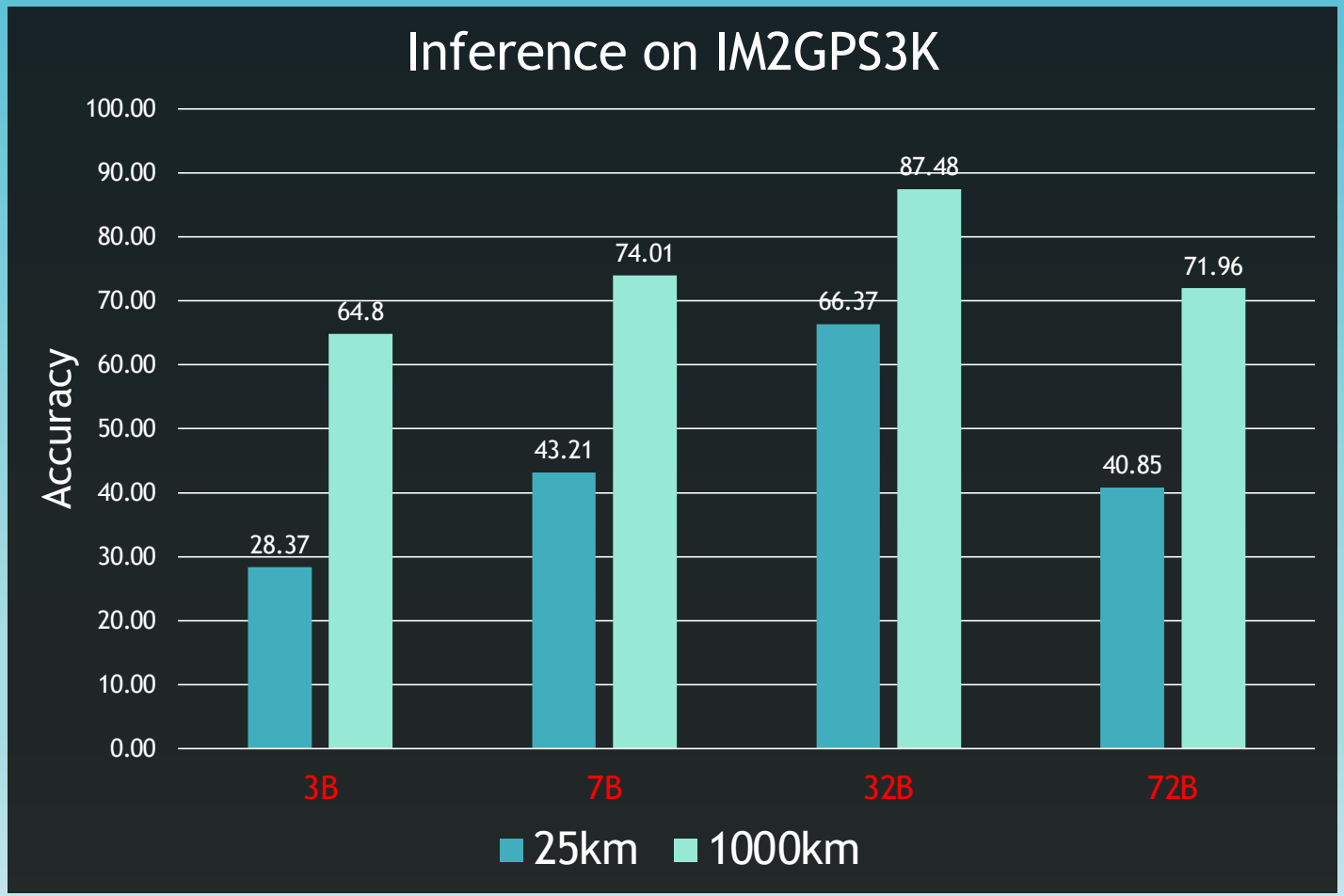
Motivation:

- Current models use flat classification
  - Don’t generalize well
  - Use cross-entropy loss
  - No distance-based error

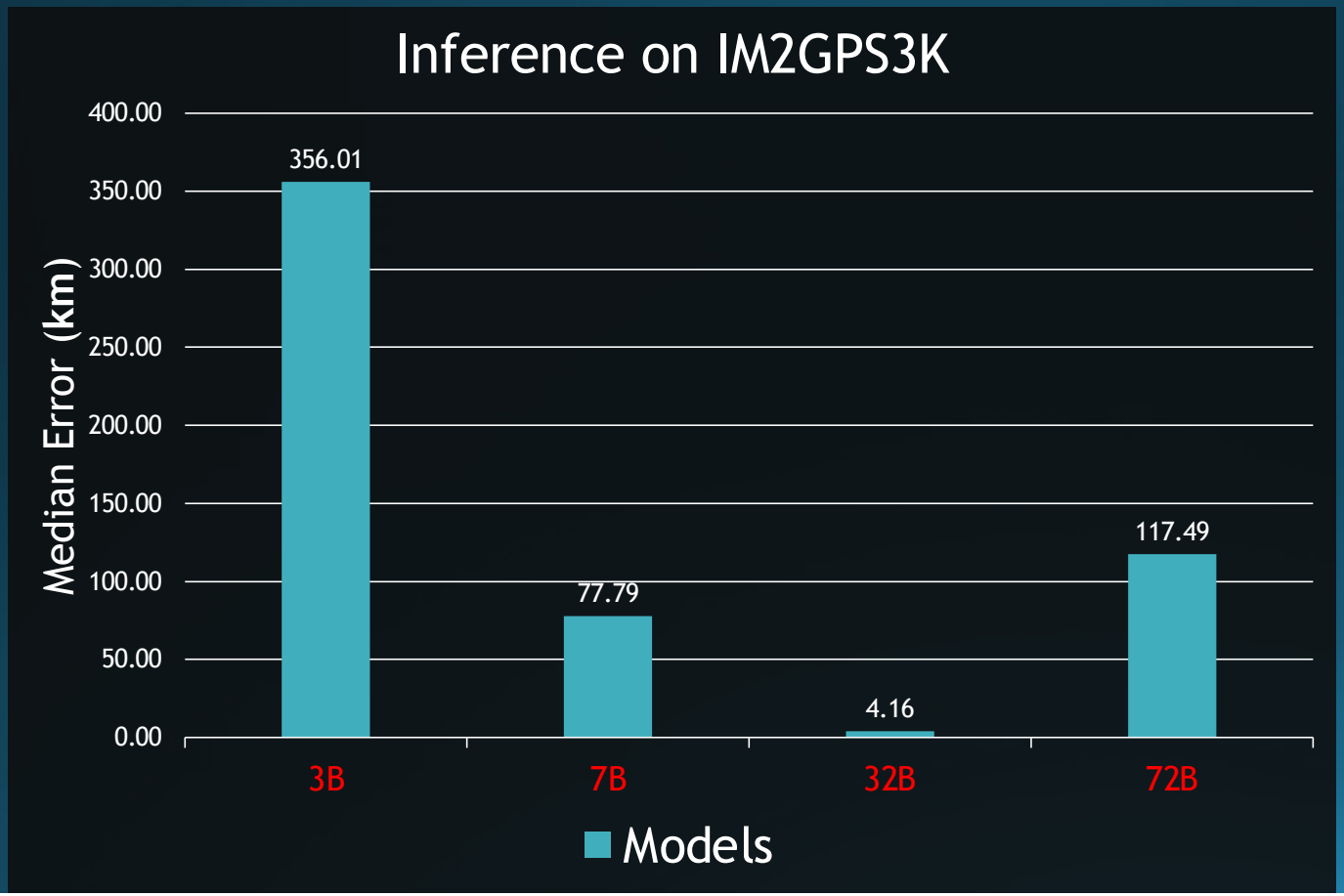
Research Question:

- Can we improve geolocation accuracy by **teaching vision-language models to reason location like humans**?

Model Scaling



Model Scaling



Prompts

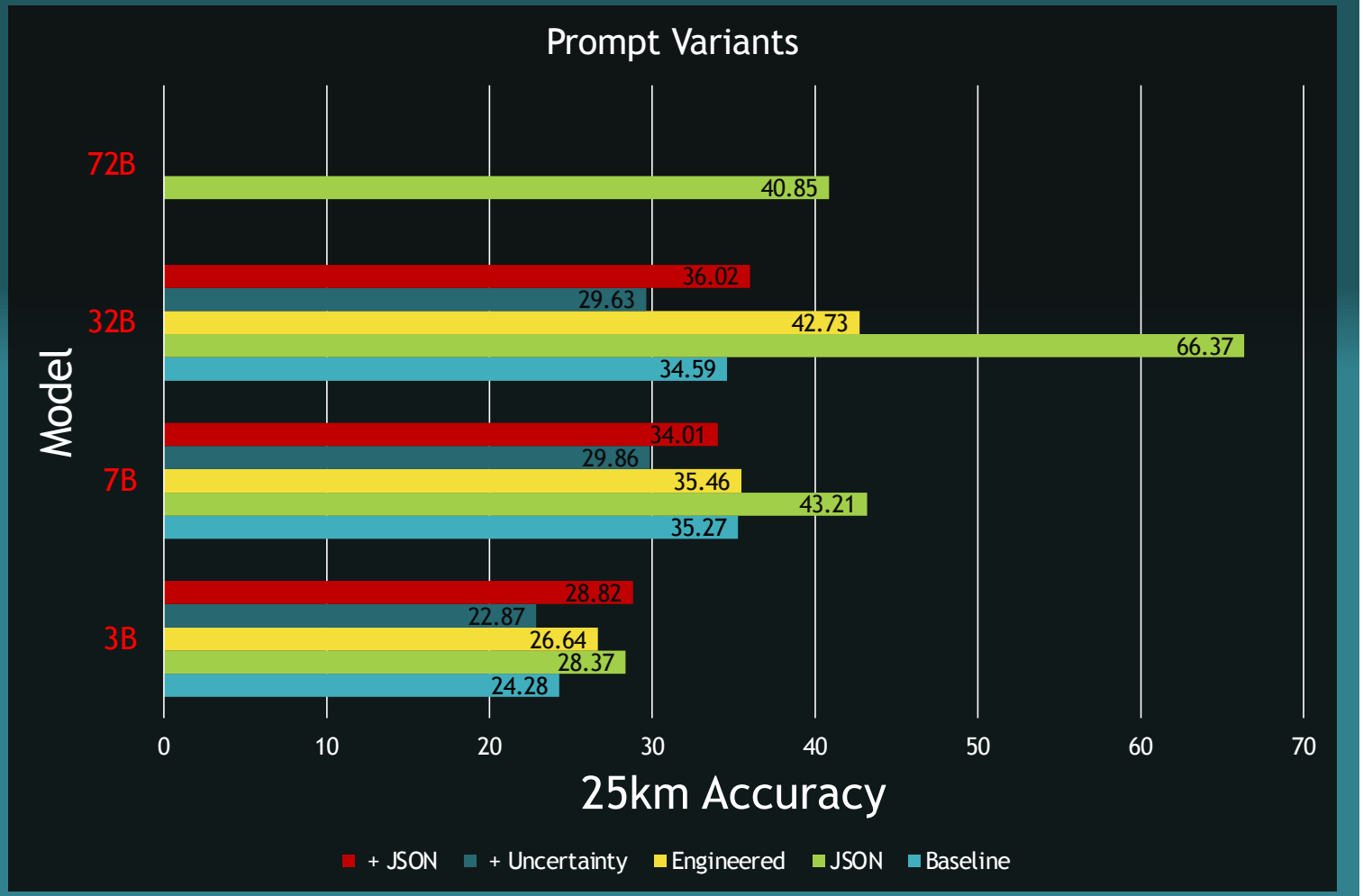
Baseline Prompt:
"<image>Where is this image taken? Respond with only the city and country."

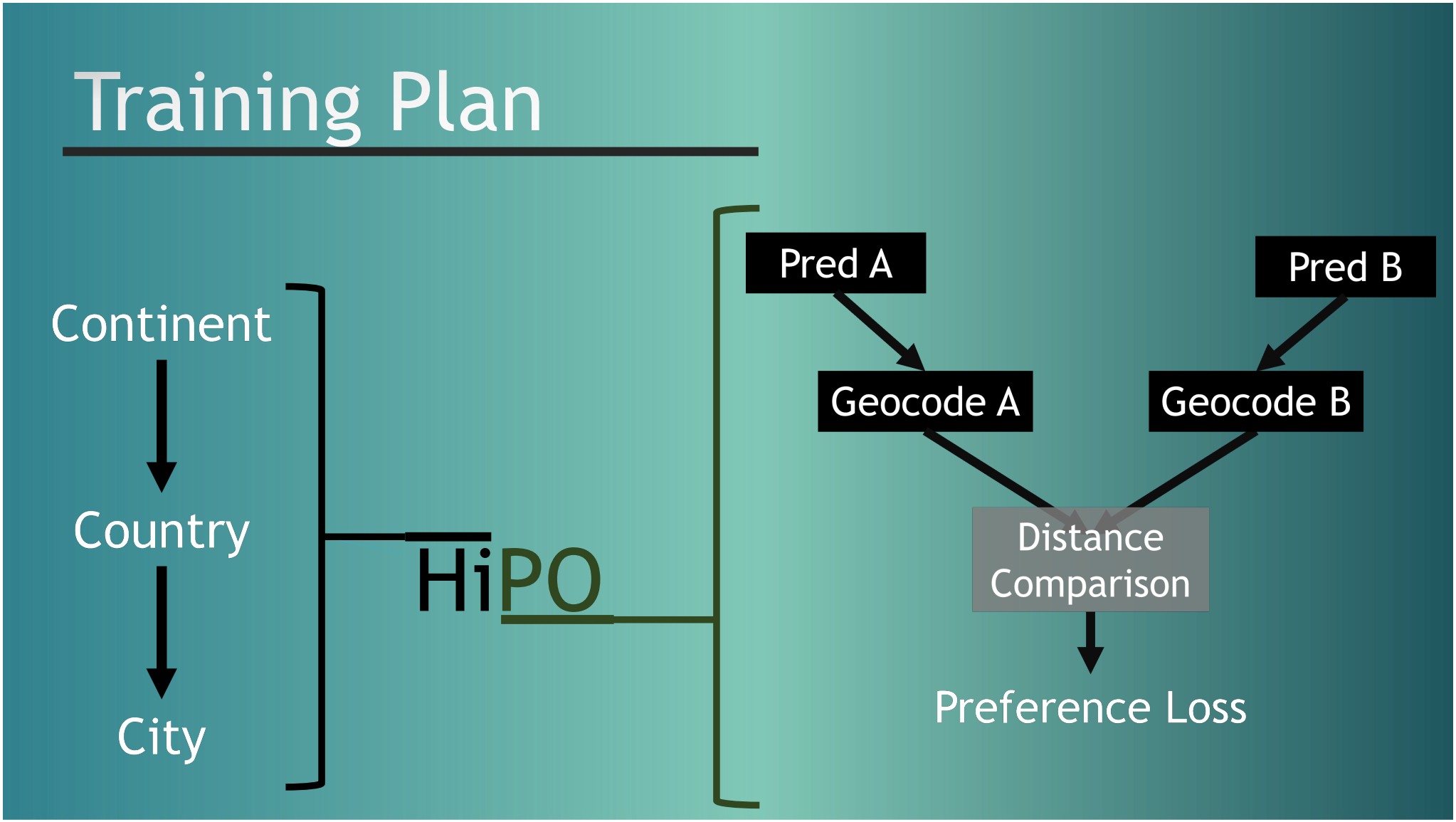
Engineered Prompt:
[Role Assignment Prompt] + [Baseline Prompt]

Uncertainty Prompt:
[Role Assignment Prompt] + [Baseline Prompt] +
"If you are unsure, respond with 'None'"

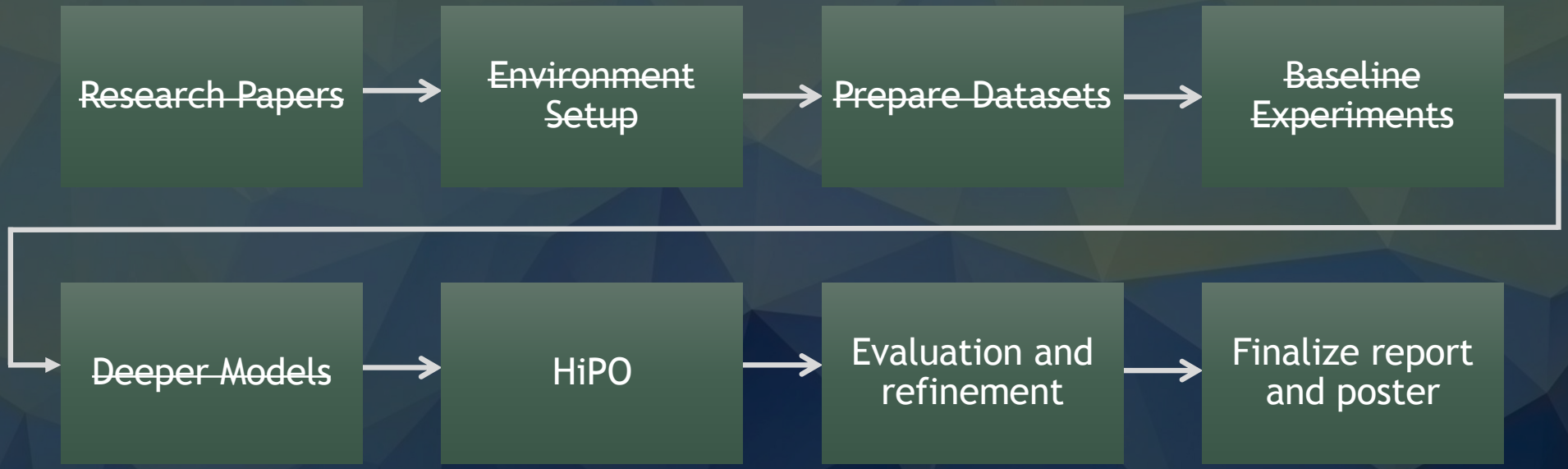
JSON Prompt:
[Role Assignment Prompt] + [Baseline Prompt] + "At the end of your answer, output the following in JSON format:…"

Prompts


Prompt Variants
Model
72B
32B
7B
3B
40.85
36.02
29.63
42.73
66.37
34.59
34.01
29.86
35.46
43.21
35.27
28.82
22.87
26.64
28.37
24.28
0
10
20
30
40
50
60
70
25km Accuracy
+ JSON
+ Uncertainty
Engineered
JSON
Baseline


Training Plan
Continent
Country
City
HiPO
Pred A
Pred B
Geocode A
Geocode B
Distance Comparison
Preference Loss

Timeline


~~Research Papers~~
~~Environment Setup~~
~~Prepare Datasets~~
~~Baseline Experiments~~
~~Deeper Models~~
HiPO
Evaluation and refinement
Finalize report and poster