Baselines and Analysis

First Last 1[*]   **First Last 2**[*]   **First Last 3**[*]   **First Last 4**[*]

{ID1, ID2, ID3, ID4}@andrew.cmu.edu

1 Models (2 pages)

Please explain the baselines you are using. This includes (but is not limited to):

1. Unimodal baselines

2. Very simple multimodal models

3. Alternate choices for modules (e.g. encoders/decoders)

1.1 Unimodal Baselines

We include the following unimodal baselines: ... we expect these to capture ...

1.2 Simple Multimodal Baselines

We include a very simple multimodal baseline that ... ignores history... doesn't have attention ... doesn't require pretraining ....

1.3 Competitive Baselines

We run N competitive baselines available at public repos. These include:

1. System 1 is a ... whose key insight is ...

2. System 2 is a ... whose key insight is ...

3. System 3 is a ... whose key insight is ...

---

[*]Everyone Contributed Equally – Alphabetical order

Methods	Dev	
	Accuracy ↑	L_2 Error ↓
Unimodal 1 ()		
Unimodal 2 ()		
Unimodal 3 ()		
Simple Multimodal 1 ()		
Simple Multimodal 2 ()		
Simple Multimodal 3 ()		
Previous Approach 1 ()		
Previous Approach 2 ()		
Previous Approach 3 ()		

2 Results (1 page)

Replace columns with the correct metrics for your task (extrinsic).

Methods	Dev	
	Detections ↑	Invalid Actions ↓
Unimodal 1 ()		
Unimodal 2 ()		
Unimodal 3 ()		
Simple Multimodal 1 ()		
Simple Multimodal 2 ()		
Simple Multimodal 3 ()		
Previous Approach 1 ()		
Previous Approach 2 ()		
Previous Approach 3 ()		

3   Analysis (2 pages)

This section should include plots. For example, how key metrics vary with a specific hyperparameter, task complexity, etc.

3.1   Intrinsic Metrics

These are not the task itself (they might overlap with auxiliary losses) but are skills the model should have.

3.2 Qualitative Analysis and Examples (full page table)

A full page table that includes: Example input, comparison of model outputs, theory about reason for failure. Every model presented should have at least one failure condition

4   Team member contributions

Member 1   contributed ...

Member 2   contributed ...

...   contributed ...