

Phase 5: Experiments and Evaluation (Some Short Long Term Ideas)

## **Direct Performance Test**

**Evaluate Student and Teacher** Models on the task that was used for training the teacher.

## **Ablation Study**

Turn features on and off to figure out which features were most influential. Potentially further simplify the L2D dataset and check whether distillation improvements hold.

Close

**Embedding** 

## **Generalization of Embeddings**

Use embeddings to train/test a model on additional tasks that they were not explicitly trained for.

**Group episodes from the two** datasets based on embedding distance and check for semantic proximity

**No Distillation** 

Train and test the student model

without distillation as a baseline.

**Clustering & Visualization** 

Close

**Embedding** 

# **Analyze Transformer Attention Weights**

Make some visualizations to understand which features were more influential.

**Phase 2: Transformer for Node Embeddings** Transformer takes in features and images (processed with some CNN?) for each individual node and makes an embedding for each, which will be used as node features.

**Transformer** 

Phase 3: Teacher Model Training Distant **Embedding** Graph Embedding **Triplet Loss with** Same dataset **Embeddings** 

- Training is done with more complex dataset (nuScenes)
- Loss is determined by downstream task
- Transformer (Phase 2) and GNN are jointly trained.
- Teacher model is frozen and final graph embeddings are stored

Phase 4: Student Model Training / Knowledge Distillation Distant **Embedding** Graph **Embedding Triplet Loss with Teacher Embeddings** 

- Jointly train student GNN and Transformer with Triplet loss using the teacher embeddings
- Training is done with less complex dataset (L2D)
- Requires matching graphs (from the two datasets) based on semantic similarities from the Semantic tags and graph structure

OR

