

Bottom-up Approach for translating Human Language to High- Level Robot Language

CS613: Natural Language Processing | Prof. Mayank Singh
Mentor – Praveen Venkatesh

Problem Description

Our goal is to develop a bottom-up approach for translating English language instructions to High level Robot Instruction

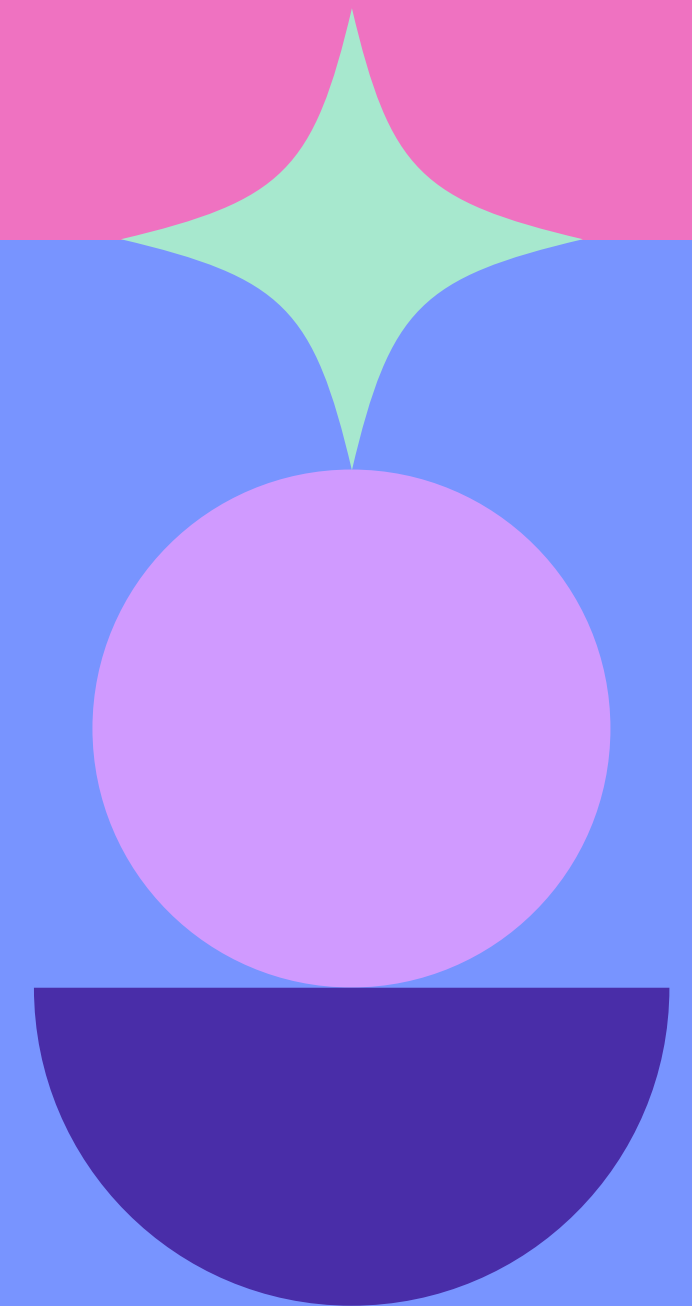
*"Exit Kitchen into corridor.
Turn right. Advance through
corridor to Hall-0. Enter
Hall-0 and proceed towards
corridor on opposite end of
Hall-0. Enter corridor and
advance to the first door on
your left. Turn left and
advance into Bathroom."*



Behavior Map

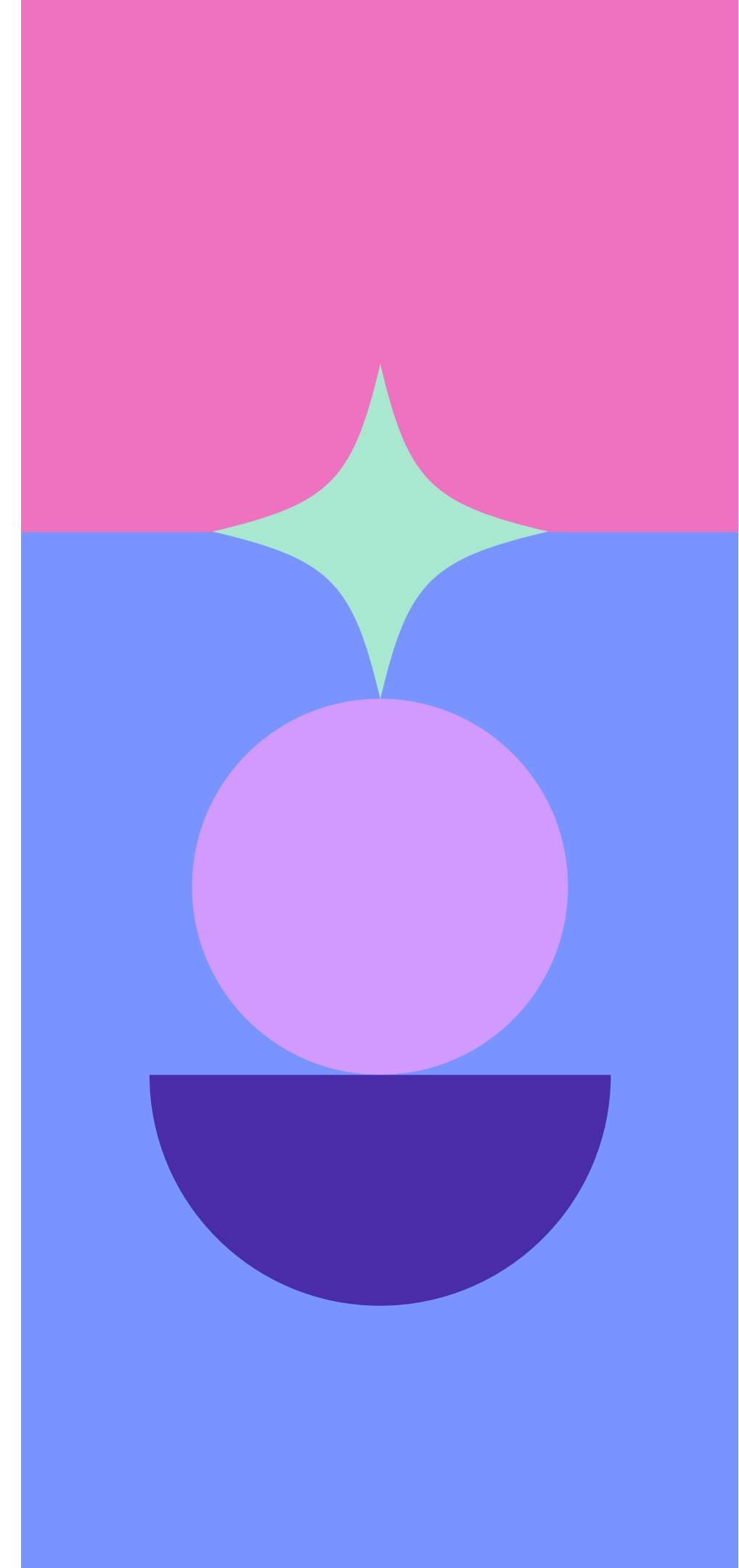


K-0 oor C-1 cf C-1 chs C-0 cf C-0 iol B-1



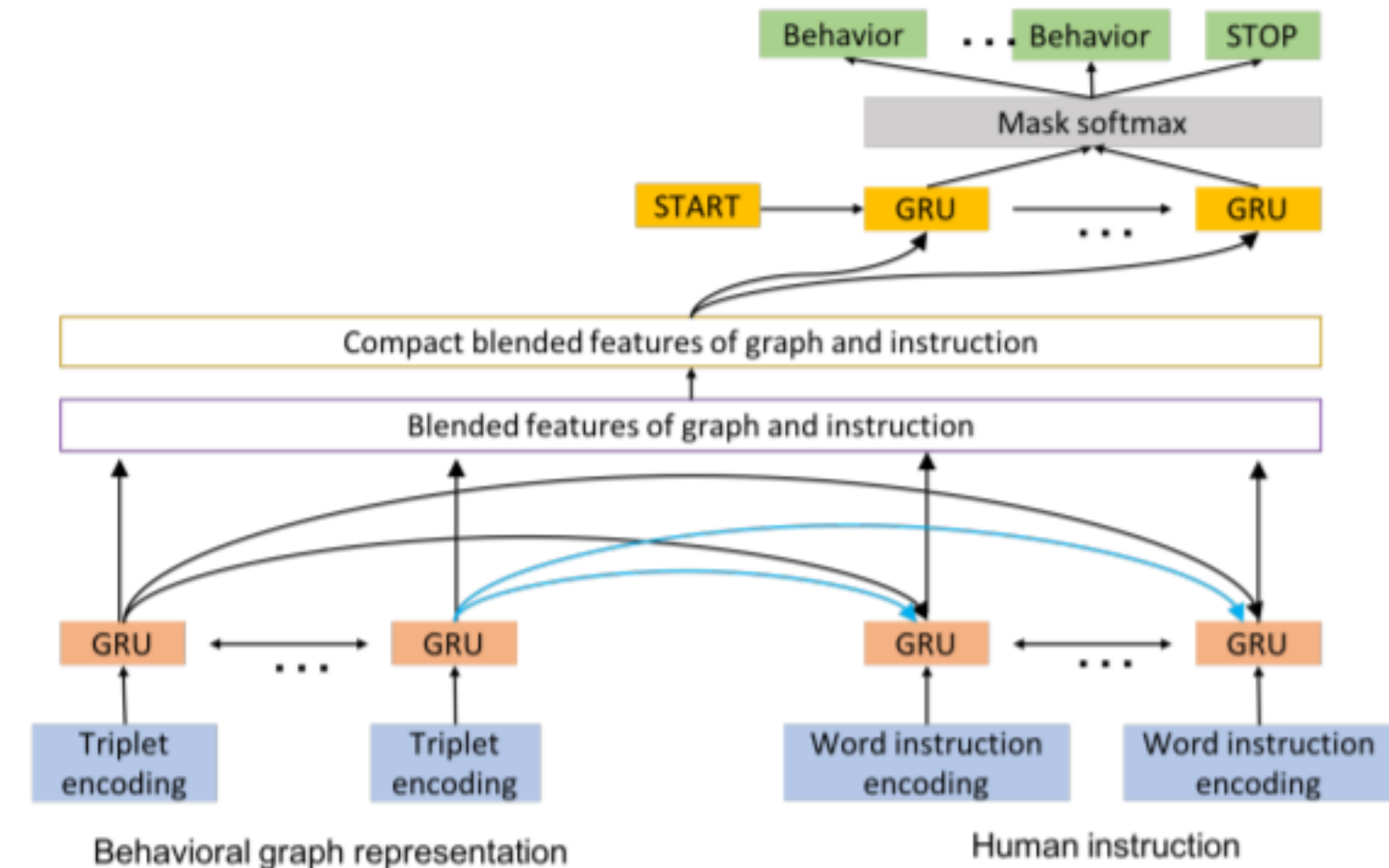
Importance

- Since the world is leading towards automation, leveraging NLP with robots can act as a huge transformation in the industry and other sectors.
- Autonomous robots can help in daily tasks, in factories, in medical fields and many more!

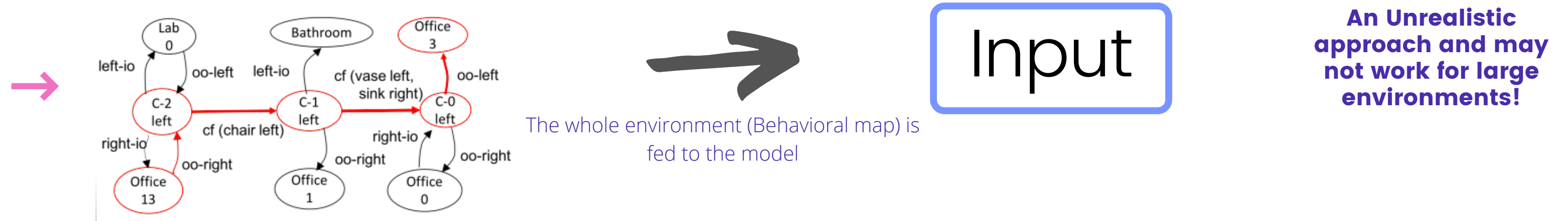


Baseline Approach

- Seq2seq translation model consists of 6 layers.
- In Embed layer, the instructions are embedded in GloVe vector and the input triplet components are encoded into vectors. The triplet components are $(p_i, b[attr], p_j)$
- In Encoder layer, there is GRU used for storing sequence memory.
- In Attention layer, one way attention used for accessing local information of the behavior graph.
- Decoder layer is a sequence of GRU's, takes the input from context matrix and behavior predicted by previous unit.



Limitations



- There is no **visualisation module** built for the model where the behaviour map could be drawn and the translated path could be highlighted.

Datasets

- Dataset created by (Zhou and He, 2020) using the **Amazon Mechanical Turk** with various maps of simulated indoor environments.
- It is a **labelled** dataset
- Train – Test Split

Dataset	#Single	#Double	#Total
Training	4062	2002	8066
Test-Repeated	944	34	1012
Test-New	962	0	962

Components

Behavioral Graph

o-13 oor c-2 ; c-2 cf c-1 ;

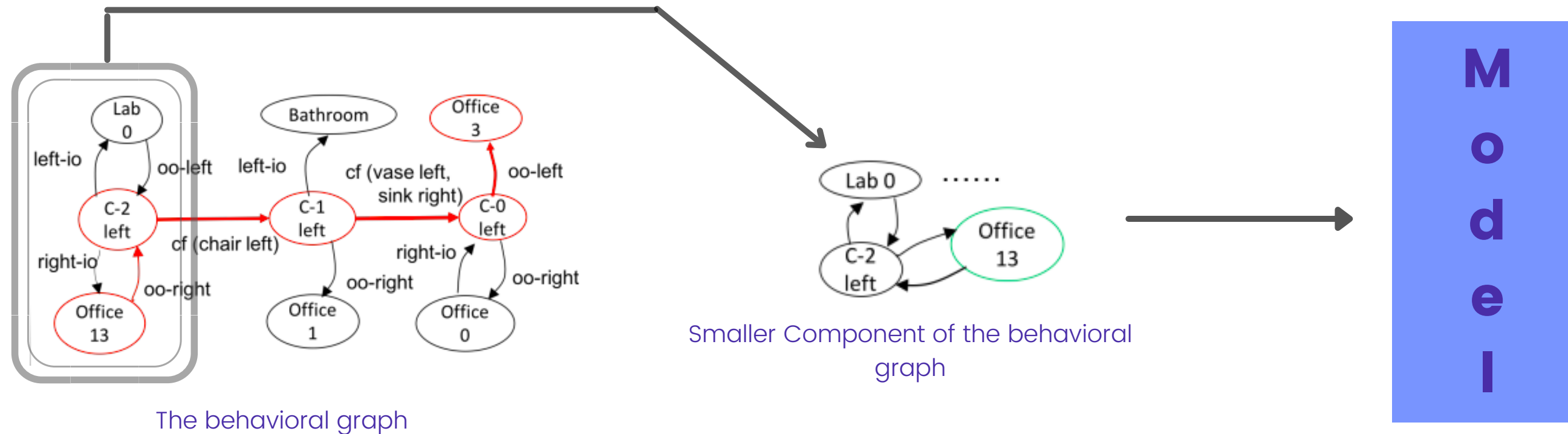
Instructions

”Exit Kitchen into corridor. Turn right. Advance through corridor to Hall-0”

Answer

oor C-3 cf C-3 lt C-2 cf C-2
ior O-7

Our Approach



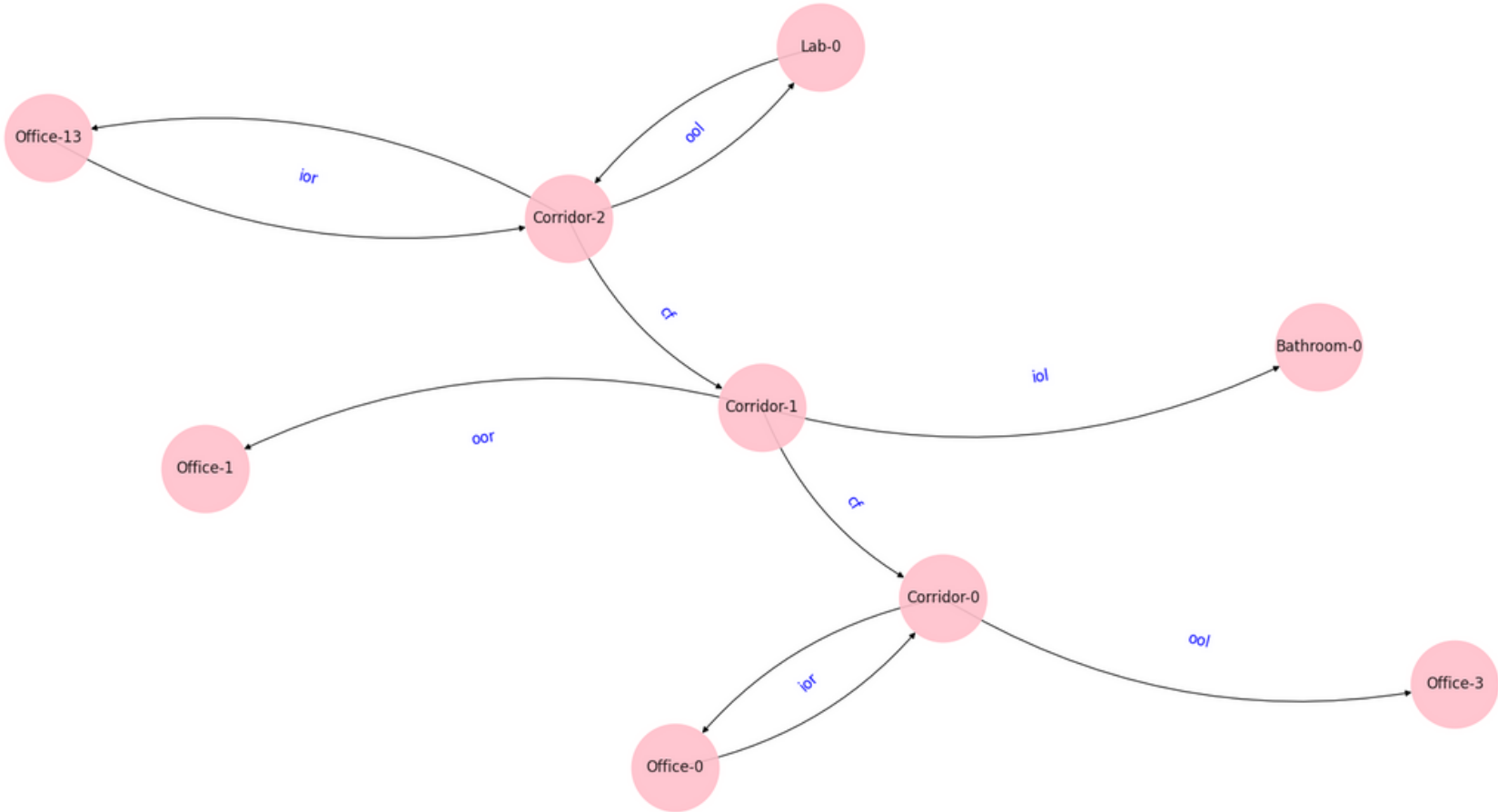
- We plan to break the behavioral map into **smaller chunks** and feed these chunks to the model instead of feeding the whole behavioral map unlike the baseline model
- This is a more **realistic bottom up** approach in learning the behavioral map and correlating the instructions with the map. However, the training process might take longer duration. Will be suited for large environments with large number of nodes and behaviors!

Our Approach

Visualisation

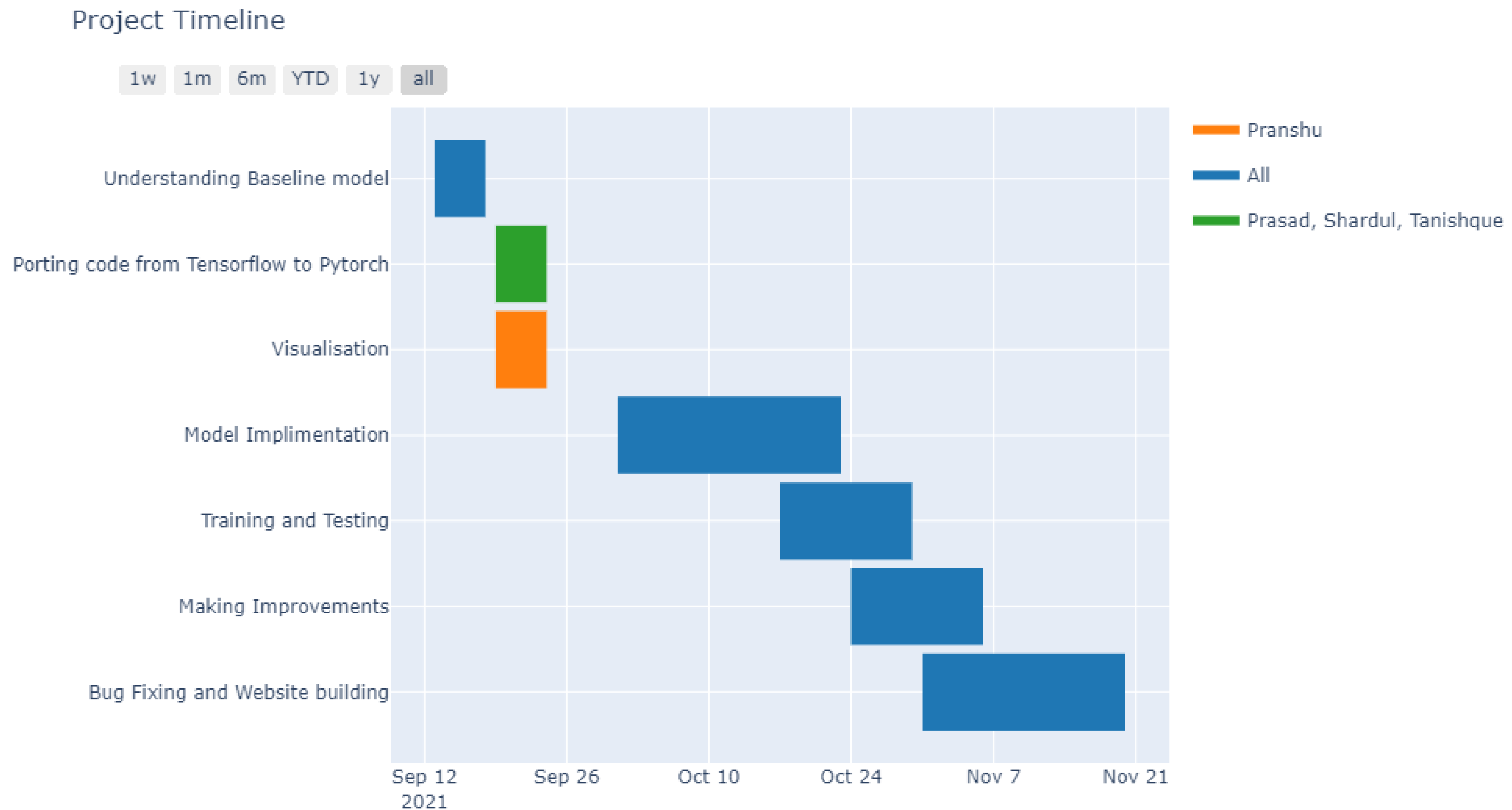
Encoded Behavioral Map

o-13 oor c-2 ; c-2 cf c-1 ; c-1 iol
b-0 ; c-2 iol l-0 ; l-0 ool c-2 ; c-
2 ior o-13 ; c-1 oor o-1 ; c-1 cf c-
0 ; c-0 ool o-3 ; c-0 oor o-0 ;
o-0 ior c-0 ;



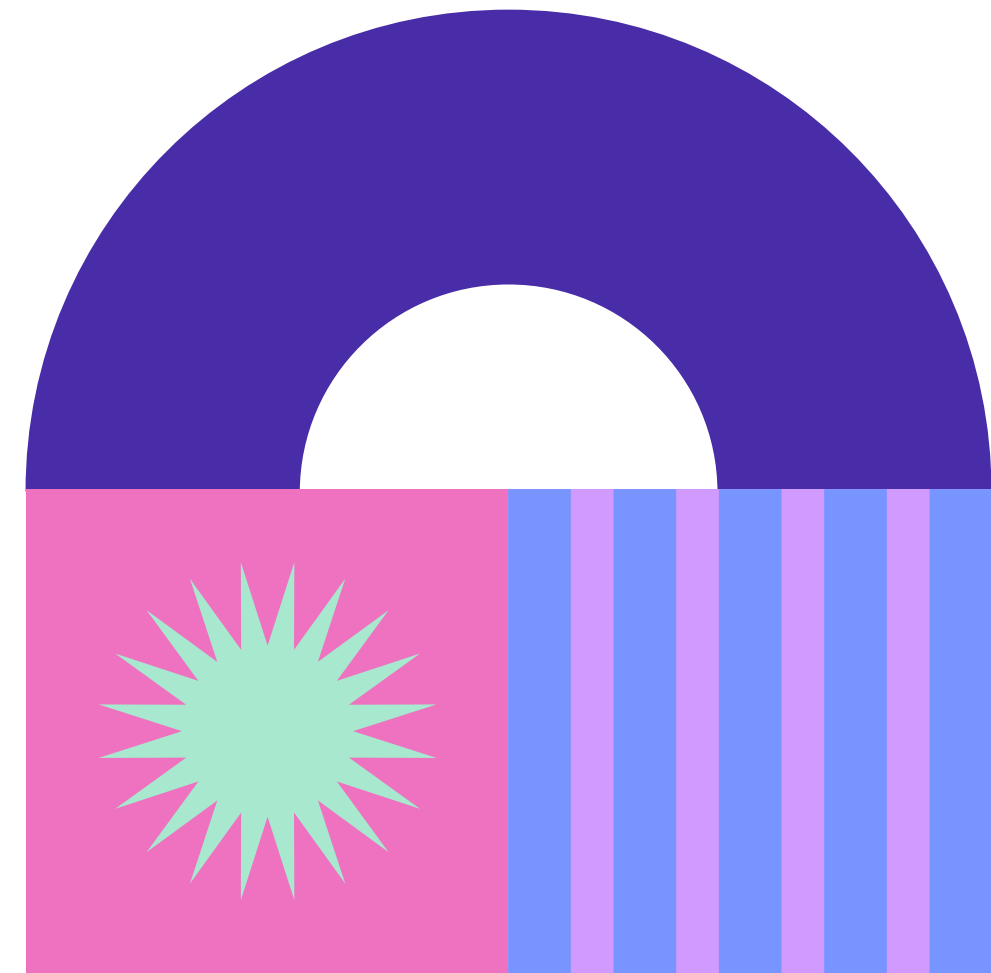
Behavioral map creating using our visualisation module

Proposed Timeline

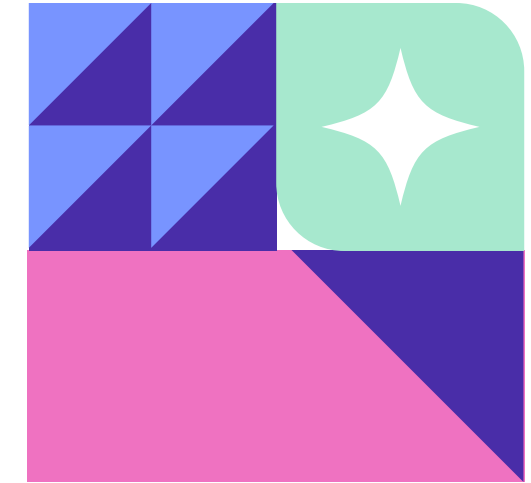


Challenges

- Hyperparameter tuning of the created model
- Converting TensorFlow to Pytorch (Bug fixing and training)
- Formulating the loss function and objective function for the proposed model and designing the model itself



Individual contribution



Pranshu

- Understanding the baseline model
- Worked on the visualization module



Tanishque

- Understanding the baseline model.
- Worked with prasad on conversion of tensorflow to pytorch.



Prasad

- Understanding the baseline model
- Tensorflow to pytorch conversion



Shardul

- Understanding the baseline model
- Worked with prasad on conversion of tensorflow to pytorch.



Thank you!

Team Members



Pranshu Kumar Gond



Prasad Athave



Tanishque Zaware



Shardul Kulkarni