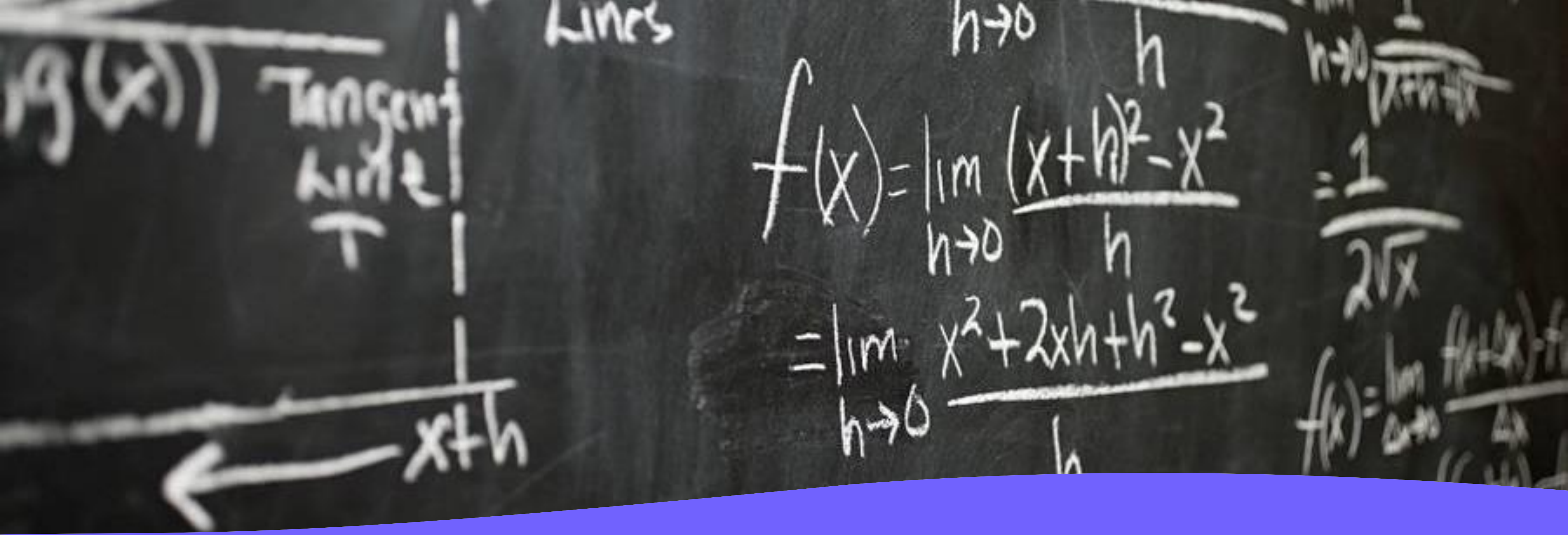
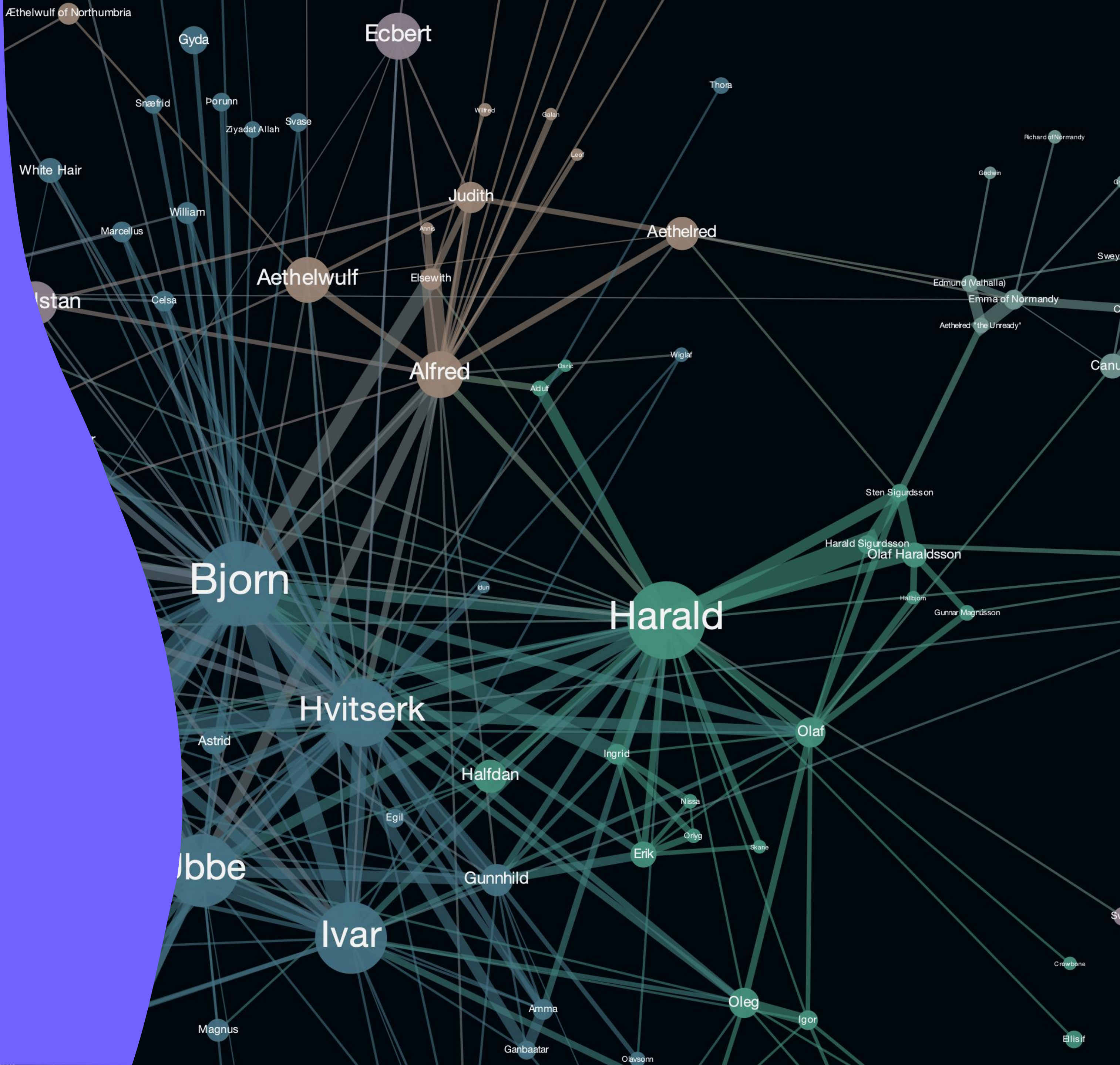


Finding Transformer Circuits with Edge Pruning (NeuRIPS 2024)

By: Group 4

Pitfalls of Other Methods



High Level Overview of Our Method

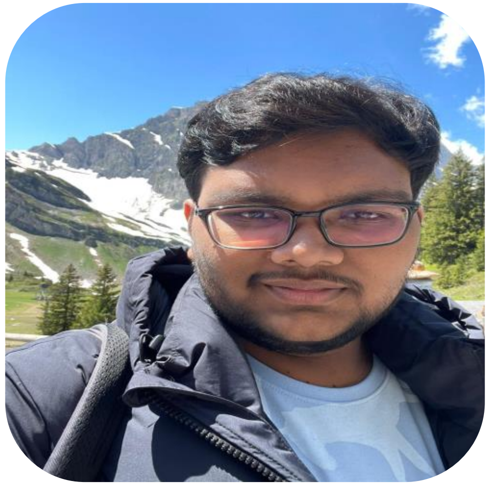
Working of Masking System

Lagrangian Sparsity Enforcement

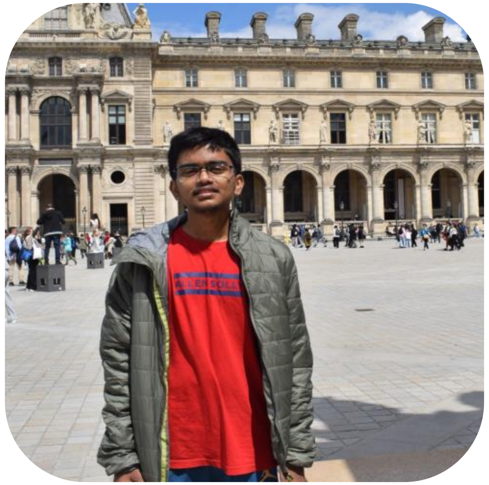
Loss Function

Code Walkthrough

Team



Aryan Lohia
23MA10073



Antariksh Das
23CS10086

Siddhant Singh
23CS10085

THANK YOU