

# CS445 Group #1: Milestone Report

Serhan Yilmaz, Sadiq Qara, Furkan Eris

November 25, 2024

## 1 Introduction

Stance detection is a critical task in natural language processing (NLP) that focuses on identifying whether a piece of text expresses an agree, disagree, or neutral stance toward a specific topic. This task is crucial for various applications, including opinion mining, social media monitoring, and political sentiment analysis. Through stance detection, we can better understand public opinion, improve content moderation systems, and address misinformation effectively.

In our project, we aim to develop a system that utilizes advanced NLP methods to classify text into stance categories. We are focusing on zero-shot stance detection, where the model must determine stance for topics not seen during training. So far, we have:

- Selected and obtained the VAST dataset for the stance detection task
- Completed initial preprocessing steps
- Split the dataset into training, validation, and test sets
- Reviewed key research papers to design our methodology

## 2 Dataset Selection

We have chosen the VAST dataset for this project because of its comprehensive coverage and suitability for stance detection. The dataset includes text samples paired with topics and annotated with stance labels (agree, disagree, or neutral). Some key features that make VAST ideal for our task include:

- Rich Annotations: Clear and reliable stance labels
- Topic-Text Pairing: Links specific topics to corresponding text
- Diverse Content: Wide range of topics and writing styles
- Benchmarking: Widely used in research, enabling comparisons

Table 1: VAST Dataset Statistics

Split	# Examples	# Unique Comments	# Topics	
			Few-shot	Zero-shot
Train	13,477	1,845	638	4,003
Dev	2,062	682	114	383
Test	3,006	786	159	600

The dataset has been split following standard practices to ensure robust evaluation. As shown in Table 1, we maintain a substantial training set while reserving sufficient data for development and testing. The split also preserves both few-shot and zero-shot topics across all partitions, enabling comprehensive evaluation of the model’s generalization capabilities.

### 3 Approach Plan

Our approach is inspired by recent advances in NLP, particularly the "Zero-Shot Stance Detection" paper by Allaway and McKeown (2020). We plan to implement a Topic-Grouped Attention Network (TGANet) with the following components:

#### 3.1 Model Architecture

- Generalized Topic Representations (GTR)
  - Using BERT embeddings for topic clustering
  - Capturing relationships between topics through unsupervised clustering
  - Enabling generalization to unseen topics
- Contextualized Encoding
  - BERT-based joint encoding of text and topic
  - Self-attention mechanisms for contextual understanding
  - Pre-trained transformer architecture for robust feature extraction
- Topic-Grouped Attention
  - Multi-head attention for topic-specific features
  - Scaled dot-product attention following Vaswani et al. (2017)
  - Dynamic weighting of text components based on topic relevance

## 4 Next Steps

Our development plan consists of three main phases:

### 4.1 Baseline Implementation

First, we will develop a Naive Bayes baseline classifier for stance detection. This baseline system will:

- Preprocess text using standard NLP techniques (tokenization, stopwords removal)
- Extract TF-IDF features for both the text and topics
- Train a Naive Bayes classifier on the training split
- Evaluate performance using macro-averaged F1 score, precision, and recall metrics
- Generate confusion matrices to understand error patterns

### 4.2 TGANet Implementation

Following the baseline, we will implement our main approach:

- Set up the BERT-based encoder for joint text-topic representation
- Implement the Generalized Topic Representations (GTR) module
  - Create topic clustering pipeline
  - Develop centroid-based topic representation generation
- Develop the Topic-Grouped Attention mechanism
- Create the classification layer with proper regularization
- Set up the training pipeline with early stopping

### 4.3 Evaluation and Analysis

Finally, we will conduct comprehensive evaluation:

- Compare TGANet performance against the Naive Bayes baseline
- Generate precision-recall curves for both systems
- Analyze performance on zero-shot vs few-shot topics
- Investigate model behavior on challenging cases
- Document limitations and potential improvements

## 5 References

### References

- [1] Allaway, E., & McKeown, K. (2020). Zero-Shot Stance Detection: A Dataset and Model using Generalized Topic Representations. In *Proceedings of EMNLP 2020*.
- [2] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems*.
- [3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL 2019*.
- [4] Mohammad, S., Kiritchenko, S., Sobhani, P., Zhu, X., & Cherry, C. (2016). SemEval-2016 Task 6: Detecting Stance in Tweets. In *Proceedings of SemEval 2016*.
- [5] Augenstein, I., Rocktäschel, T., Vlachos, A., & Bontcheva, K. (2016). Stance Detection with Bidirectional Conditional Encoding. In *Proceedings of EMNLP 2016*.