Argumentative Stance Prediction: An Exploratory Study on Multimodality and Few-Shot Learning

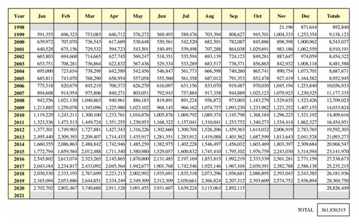
Arushi Sharma*, **Abhibha Gupta***, Maneesh Bilalpur*
School of Computing and Information, University of Pittsburgh

MOTIVATION

To investigate whether multimodality improves argumentative stance prediction for gun control and abortion related tweets.

Example

Gun control: The past year has seen more #NICS checks than any same month. #Gunsense #guncontrol is dead. #SelfDefense is a #HumanRight



Abortion: Turns out that women had to wait much, much longer - until 2017 - to get access to medical abortions. #riseuparchive #womenshistory #feminism #abortion #prochoice

Importance

- Enhances understanding of public opinion dynamics Eg: Stance on controversial topics
- Aids in immediate feedback analysis for policy makers

Challenges

- Analyzing brief informal social media texts
- Integrating text and image data effectively in prediction models

DATASET

- Dataset consists of Twitter texts along with their images from two topics–gun control and abortion.
- Each text-image pair corresponding to a tweet are annotated with a stance (support or oppose)
- The abortion dataset is imbalanced by a 1:3 support:oppose stance ratio.
- Imbalance is addressed through weighted cross-entropy loss, with higher weight for minority category.

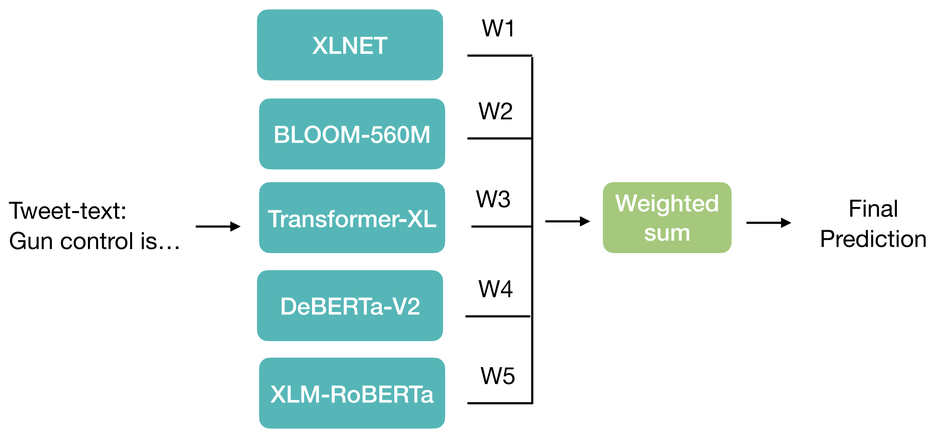
Dataset	Train	Validation	Test
gun-control	920	100	150
abortion	891	100	150

RESEARCH QUESTIONS

1. How well does language as a stand-alone modality perform at argumentative stance prediction?
2. Does incorporating image information improve prediction performance?
3. How do Large-Language Models (LLMs) in few-shot setting compare against fine-tuned unimodal and multimodal models?

EXPERIMENTS

Ensemble Stance Prediction

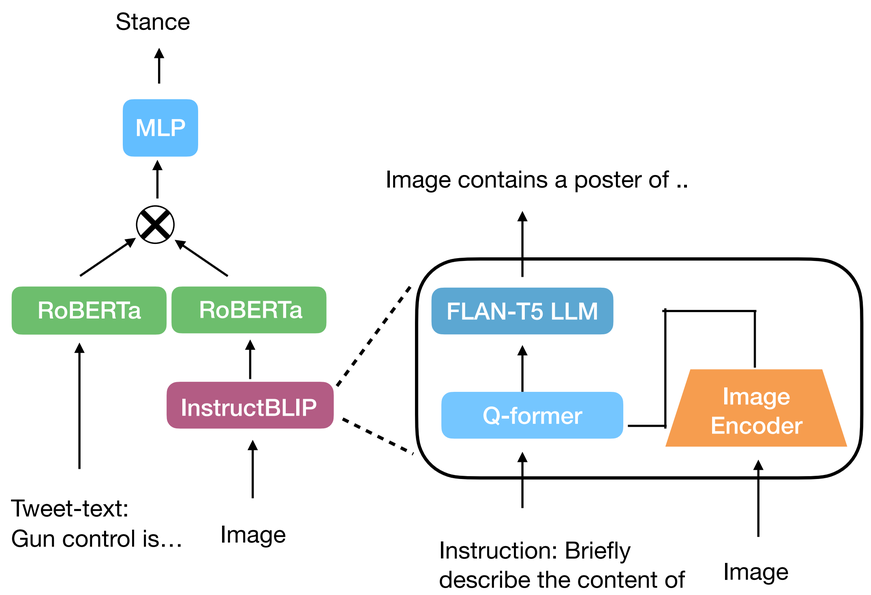


Weights based on F1-score on validation set. XLNet and BLOOM-560M received the predominant weights for attaining the highest F1 score on abortion and gun-control datasets respectively.

Model	Precision	Recall	F1
XLNET	0.619	0.924	0.741
BLOOM-1B	0.760	0.660	0.710
BLOOM-560M	0.707	0.898	0.791
Transformer-XL	0.571	0.881	0.693
DeBERTa-V2	0.560	0.710	0.630
XLM-RoBERTa	0.650	0.880	0.750
Ensemble	**0.743**	**0.906**	**0.817**

Support stance performance using text-based transformer models.

Multimodal Stance Prediction

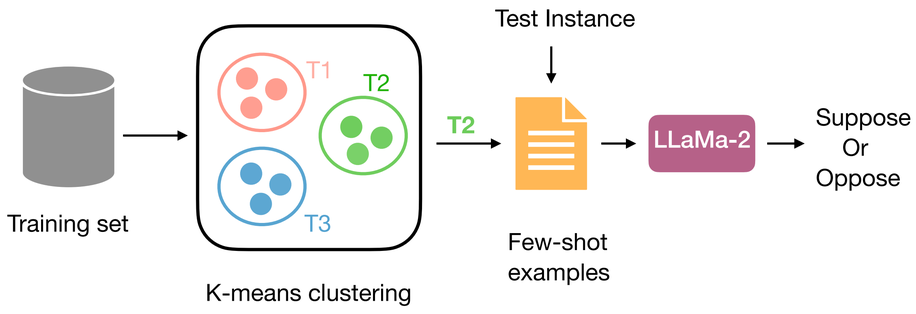


Multimodal RoBERTa configuration. The figure shows the input image summarized as text through instructBLIP and then used to fine-tune the RoBERTa model together with the tweet-text. Shared color between RoBERTa models indicates tied weights.

Model	Precision	Recall	F1
ViLT	0.680	0.432	0.528
FLAVA	0.570	0.650	0.610
Multimodal - RoBERTa	**0.531**	**0.932**	**0.677**

Support stance performance using image-text multimodal transformer models.

Few Shot Stance Prediction using LLM's



Topic T2 for the test instance was selected based on its similarity to the clusters.

Model	Precision	Recall	F1
Baseline (Support-only)	0.395	1.00	0.566
Zero-shot	0.440	0.290	0.350
Four-shot	0.420	0.640	0.500
Four-shot with K-means	**0.450**	**0.700**	**0.550**

Support stance performance using LLaMA-2 based few-shot experiments.

DISCUSSION

1. BLOOM-560M slightly underperforms the ensemble, highlighting a trade-off between the ensemble's performance and the significant computational resources required for a small improvement.
2. Vision-language models, not typically trained to read text in images, face challenges in such tasks. However, when guided with specific instructions, they effectively describe objects and recognize text in images.
3. LLaMA-2 model significantly improves with four-shot prompting highlighting the potential of few-shot learning, especially in capturing nuanced stances with similar themes.

Gun control	Abortion
Gun violence as a mental health problem	Reproductive rights of women
Effects of gun violence on children	#savethebabyhumans hashtag
Pro-gun control politicians	Roe v Wade abortion case
Racism and gun control	Religion and motherhood
Trump and guns	Pro-life
Illegal acquisition of guns	Birth control pills
Supreme Court and gun control	Abortion is murder
Second amendment right	Supreme Court and abortion
	Abortion is evil
	Natural law right to Life

Themes identified using k-means clustering for few-shot examples in gun control and abortion datasets. Same theme(s) captured by multiple clusters resulted in fewer themes than reported clusters.

FUTURE WORK

- Incorporating domain knowledge and alternative prompting methods such as Question Decomposition and Tree-of-Thought, offers a way to not only predict stance but also provide the rationale, addressing LLaMA-2's performance limitation.