A. Prediction Classes

Table 1. Multiclass Predictions.

	0	1	2	3	4	5	6
P1	% is in [0,25)	% is in [25,50)	% is in [50,75)	% is in [75,100]			
P2	% is in [0,25)	% is in [25,50)	% is in [50,75)	% is in [75,100]			
P3	virtue=vice	virtue>vice	virtue<vice				
P4	neutral	positive	negative				
P5	neutral	anger	disgust	fear	joy	sad.	surp.
P6	pos.=neg.	pos.>neg.	pos.<neg.				

Table 2. Binary Predictions.

	0	1
P1	% is in [0,50)	% is in [50,100]
P2	% is in [0,50)	% is in [50,100]
P3	virtue interval $\geq$ vice interval	virtue interval $<$ vice interval
P4	positive, neutral	negative
P5	neutral, joy, surprise	fear, anger, disgust, sadness
P6	pos. interval $\geq$ neg. interval	pos. interval $<$ neg. interval

B. Machine Learning Models and Parameters

Following is a list of all models and the values for the hyper-parameters search:

- Logistic Regression: no hyper-parameters, but solver was set to "lbfgs" and max_iter to 1000.
- Decision Tree: no hyper-parameters.
- Random Forest:
  - ∗ min_samples_leaf: 1, 5, 10, and 20.
  - ∗ n_estimators: 10, 50, and 100.
- Multinomial Naive Bayes: with parameter force_alpha set to True and hyper-parameter search done on:
  - ∗ alpha: 1e-9, 0.05, 0.1, 0.15, 0.2, 1, and 100 for binary; 0.00001, 0.0001, 0.001, 0.1, 1, 10, 100, and 1000 for multiclass.
  - ∗ fit_prior: True and False.
- Complement Naive Bayes: with parameter force_alpha set to True and hyper-parameter search done on:
  - ∗ alpha: 1e-9, 0.05, 0.1, 0.15, 0.2, 1, and 100.
  - ∗ fit_prior: True and False.
- Neural Network: with parameter max_iter set to 500 and hyper-parameter search done on:
  - ∗ hidden_layer_sizes: one hidden layer with 10, 15, 20, and 30 nodes, and two hidden layers with 20, 25, 30, and 35 nodes each.
  - ∗ batch_size: 1, 10, 50, 100 and 200.

The dataset is split and used 90% for training and 10% for testing.

C. Machine Learning Results

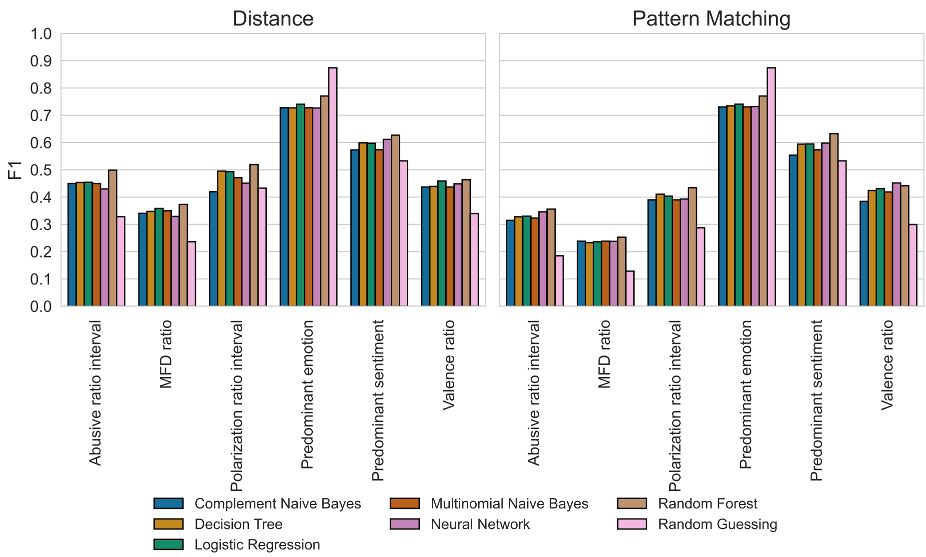


Figure 1. F1 scores pattern-matching and distance of words.

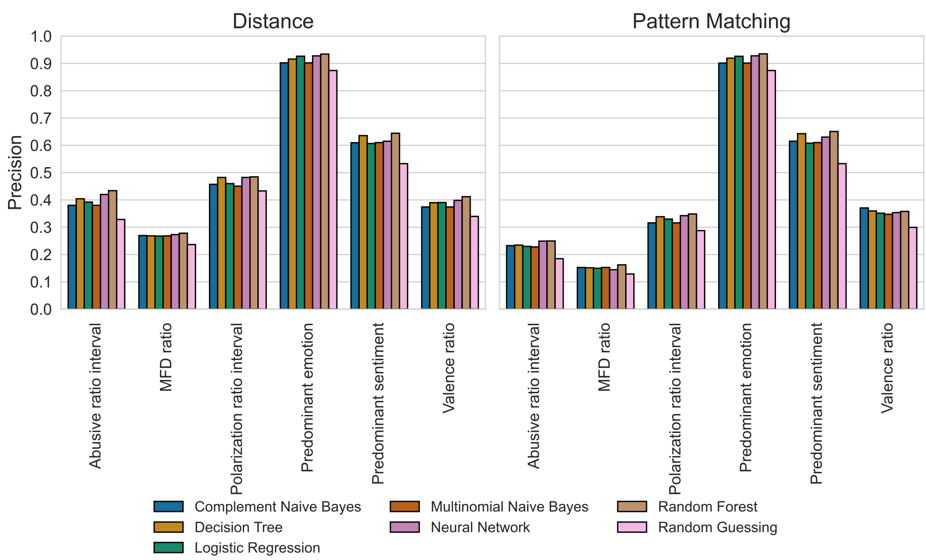


Figure 2. Precision scores pattern-matching and distance of words.

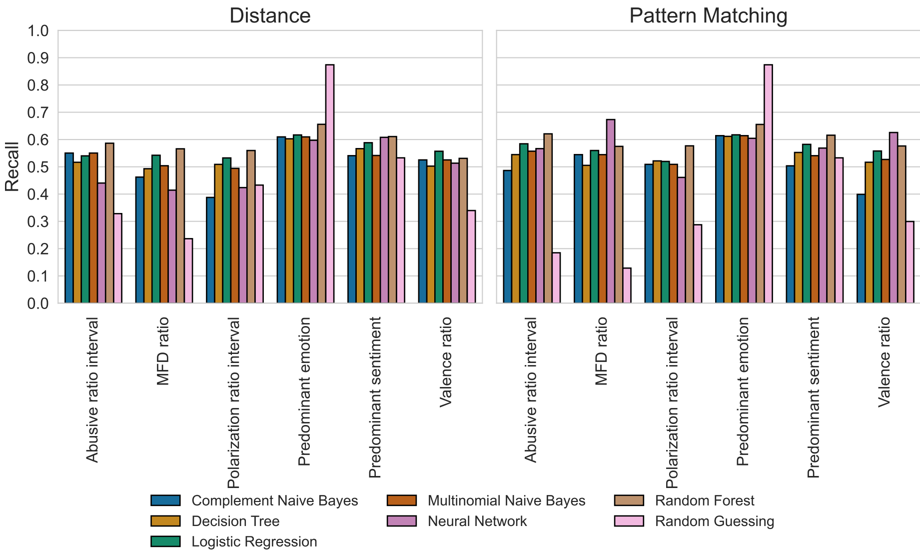


Figure 3. Recall scores pattern-matching and distance of words.

Table 3. F1 score differences with Random Guessing.

Feature	Pattern Matching	Distance
Abusive ratio interval	98.5%	51.94%
MFD ratio	96.64%	57.76%
Polarization ratio interval	51.1%	19.95%
Predominant emotion	-11.83%	-11.83%
Predominant sentiment	18.75%	17.7%
Valence ratio	50.85%	36.57%

D. Prompt structure

<|begin_of_text|><|start_header_id|>system<|end_header_id|>**@username** has the following personality profile:

- Openness: **personality_label** (Score: 0.XX)
- Conscientiousness: **personality_label** (Score: 0.XX)
- Openness: **personality_label** (Score: 0.XX)
- Extraversion: **personality_label** (Score: 0.XX)
- Agreeableness: **personality_label** (Score: 0.XX)
- Communication style: **communication_style_list**
- Personality traits: **personality_trait_category**

@username has engaged in a Twitter conversation. The last tweets from that conversation are:

- "**tweet_text_1**"
- "**tweet_text_2**"
- "**tweet_text_3**"
- "**tweet_text_4**"
- "**tweet_text_5**"

Conversation metrics:

- Abusive words: **metric_label**
- Polarization words: **metric_label**
- Predominant emotion: **predominant_category_label**
- Moral virtue words: **metric_label**
- Moral vice words: **metric_label**
- Valence positive words: **metric_label**
- Valence negative words: **metric_label**
- Predominant sentiment: **predominant_category_label**

<|eot_id|><|start_header_id|>user<|end_header_id|>**PREDICTION_TASK**<|eot_id|>
<|start_header_id|>assistant<|end_header_id|>**GROUND_TRUTH**

Being **PREDICTION_TASK** one of the three following:

In one word, predict if **@username**'s response to the conversation will have a High or Low amount of **type_1_prediction** words.

In one word, predict if **@username**'s response to the conversation will have more **type_2_prediction_category_1**, more **type_2_prediction_category_2**, or Equal amount of **type_2_prediction** words. Your answer should be **type_2_prediction_category_1**, **type_2_prediction_category_2**, or Equal.

In one word, predict which **type_3_prediction** will have **@username**'s response to the conversation. Your answer should be **type_3_prediction_categories**.

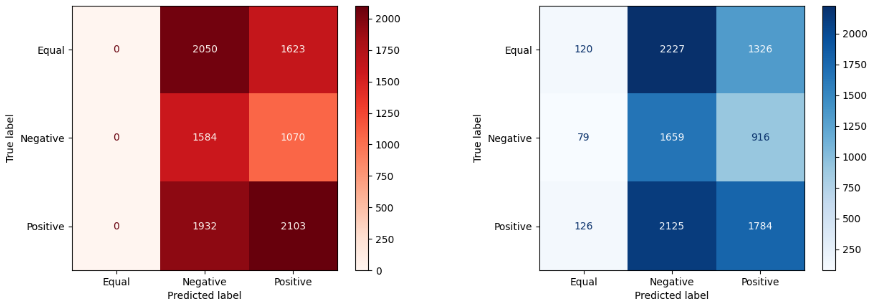
If the prompt corresponds to an example, then **GROUND_TRUTH** is form as the corresponding prediction label as explained on Section **??**, followed by <|eot_id|>. If it is not an example, it is left empty.

Where:

- **username**: Twitter's username of the current user.
- **personality_label**: "High" if the personality score is greater than 0.5, "Low" otherwise.
- **communication_style_list**: list including all categories of Communication Styles and Communication Needs whose scores are greater than 0.5.
- **personality_trait_category**: "Rational" if rational cateogy score is greater than 0.5, "Emotional" otherwise. it can be the case that there are less than five previous tweets.
- **tweet_text_[1:5]**: text of the previous five tweets to the current user's response.
- **metric_label**: "High" or "Low" following what is described on Section **??**.
- **predominant_category_label**: name of the predominant category.
- **type_1_prediction**: "abusive" or "polarization".
- **type_2_prediction**: "moral" or "valence".
- **type_2_prediction_category_1**: "Virtue" or "Positive", which corresponds.
- **type_2_prediction_category_2**: "Vice" or "Negative", which corresponds.
- **type_3_prediction**: "sentiment" or "emotion".
- **type_3_prediction_categories**: a comma-separated list of all sentiment or emotion categories, with ", or" before the final item.

Zero-shot experiment prompts are built following the previous template and using the test dataset, with **GROUND_TRUTH** left empty. For the Few-shot experiments, a prompt is built and attached for each different prediction class using the training dataset and filling **GROUND_TRUTH**; then a prompt like the zero-shot experiments is built for the prediction, and attached at the end.

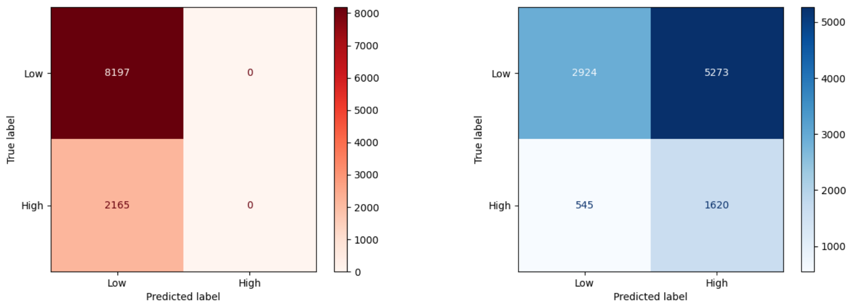
E. Large Language Models Results



(a) Zero-shot.

(b) Few-shot.

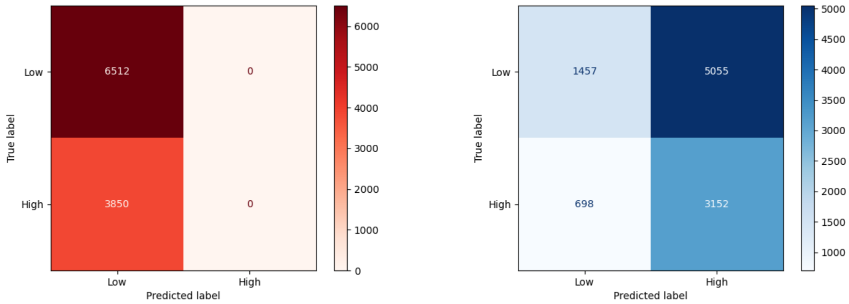
Figure 4. Confusion matrices for Valence.



(a) Zero-shot.

(b) Few-shot.

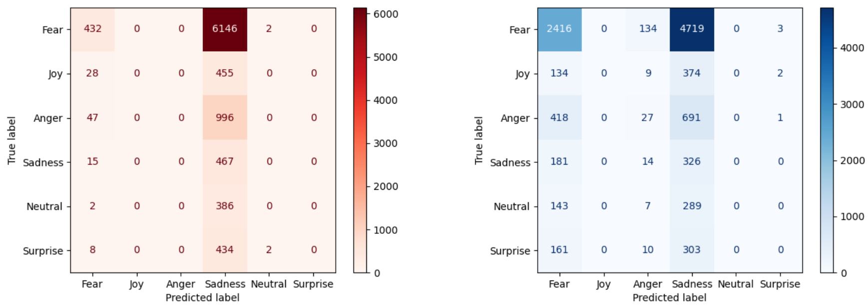
Figure 5. Confusion matrices for Abusive.



(a) Zero-shot.

(b) Few-shot.

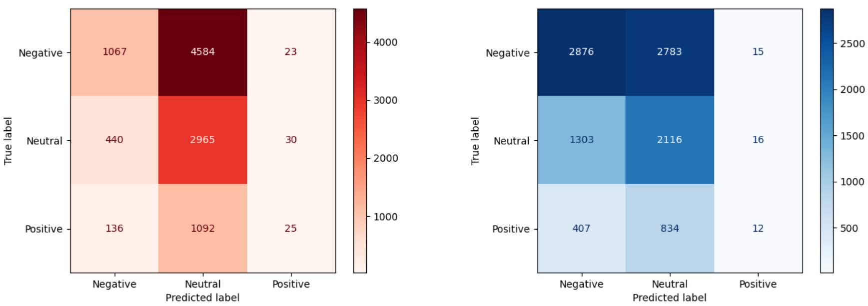
Figure 6. Confusion matrices for Polarization.



(a) Zero-shot.

(b) Few-shot.

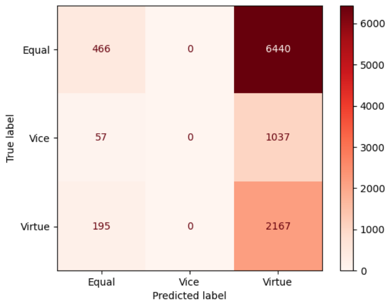
Figure 7. Confusion matrices for Emotion.



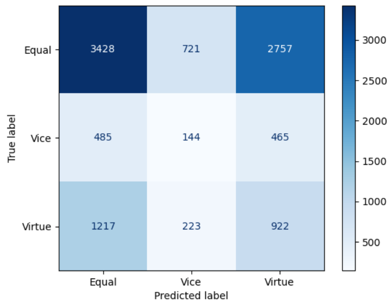
(a) Zero-shot.

(b) Few-shot.

Figure 8. Confusion matrices for Sentiment.



(a) Zero-shot.



(b) Few-shot.

Figure 9. Confusion matrices for MFD.

Table 4. Comparison of Zero Shot and Few Shot performance for each prediction.

Prediction	**Metric**	**Zero Shot**	**Few Shot**
Abusive	Precision	0.0	0.2350
	Recall	0.0	0.7483
	F1	0.0	0.3577
Polarization	Precision	0.0	0.3841
	Recall	0.0	0.8187
	F1	0.0	0.5228
Emotion (Micro)	Precision	0.0868	0.2672
	Recall	0.0868	0.2672
	F1	0.0868	0.2672
Emotion (Macro)	Precision	0.1235	0.1471
	Recall	0.1365	0.1636
	F1	0.0300	0.0969
Emotion (Weighted)	Precision	0.5725	0.5082
	Recall	0.0868	0.2672
	F1	0.0827	0.3252
Sentiment (Micro)	Precision	0.3915	0.4829
	Recall	0.3915	0.4829
	F1	0.3915	0.4829
Sentiment (Macro)	Precision	0.4377	0.4251
	Recall	0.3571	0.3775
	F1	0.2734	0.3469
Sentiment (Weighted)	Precision	0.5081	0.4995
	Recall	0.3915	0.4829
	F1	0.3270	0.4622
MFD (Micro)	Precision	0.2541	0.4337
	Recall	0.2541	0.4337
	F1	0.2541	0.4337
MFD (Macro)	Precision	0.2912	0.3410
	Recall	0.3283	0.3395
	F1	0.1611	0.3283
MFD (Weighted)	Precision	0.4838	0.5100
	Recall	0.2541	0.4337
	F1	0.1638	0.4582
Valence (Micro)	Precision	0.3558	0.3439
	Recall	0.3558	0.3439
	F1	0.3558	0.3439
Valence (Macro)	Precision	0.2410	0.3628
	Recall	0.3727	0.3666
	F1	0.2872	0.2952
Valence (Weighted)	Precision	0.2436	0.3741
	Recall	0.3558	0.3439
	F1	0.2842	0.2917