



# Emotion & Personality Detection

*F·R·I·E·N·D·S*

Workshop on NLP for Conversational AI (co-located with ACL 2023): FETA challenge

Kelly Chiu, Wally Li, Yian Wang, Carrie Yuan

# Task descriptions

```
{  
  "utterance_id": "s01_e02_c01_u002",  
  "speakers": ["Joey Tribbiani"],  
  "transcript": "Yeah, right!.....Y'serious?",  
  "tokens": [  
    ["Yeah", ",", "right", "!"],  
    ["....."],  
    ["Y'serious", "?"]  
  ],  
  "emotion": "Neutral"  
},
```

Emotion labels



Primary task: Emotion Detection

```
{  
  scene_id: s01_e01_c05  
  character: Paul  
  Text: <b>Monica Geller</b>: Oh my  
        God!<br><br>  
        <b>Paul</b>: I know, I know,  
        I'm such an idiot. I guess I ...  
        <br><br>  
  AGR: 1  
  CON: 0  
  EXT: 1  
  OPN: 1  
  NEU: 0  
}
```

Personality labels

Agreeable:
Conscientious:
Extraverted:
Open to experience:
Emotionally Stable:

Adaptation task: Personality Detection

# Work Process

## D2 (Primary Task)

- Baseline (probability BoW)
- BERT

## D3 (Primary Task)

- SVM(BoW)
- SVM(Tfidf)
- BERT(fine-tuning)

⇒ Analysis: Confusion Matrix on Emotion Recognition + thinking D4

## D4

- Ensemble
- Adaptation: Personality Detection

# A Little Experiment

I'm serious. You're amazing. You know when to spritz, when to lay back.

Really? You don't know what that means to me.

I'm gonna go to the bathroom, maybe I'll see you there in a bit?

Kay!

Wife: ...I think you're (a third person) absolutely delicious.

Husband: Excuse me! I am standing right here!

[Ross is having a rough time at his own wedding]

Ross: Actually, do you guys mind staying here for a while?

Monica: Ugh, y'know, umm we gotta get up early and catch that plane for New York.

Chandler: Yeah, it's a very large plane.

Ross: That's cool.



# A Little Experiment

I'm serious. You're amazing. You know when to spritz, when to lay back.

Really? You don't know what that means to me.



Joyful

I'm gonna go to the bathroom, maybe I'll see you there in a bit?

Kay!



Peace Neutral

Wife: ...I think you're (a third person) absolutely delicious.

Husband: Excuse me! I am standing right here!



Mad

[Ross is having a rough time at his own wedding]

Ross: Actually, do you guys mind staying here for a while?

Monica: Ugh, y'know, umm we gotta get up early and catch that plane for New York.

Chandler: Yeah, it's a very large plane.



Joyful

Ross: That's cool.

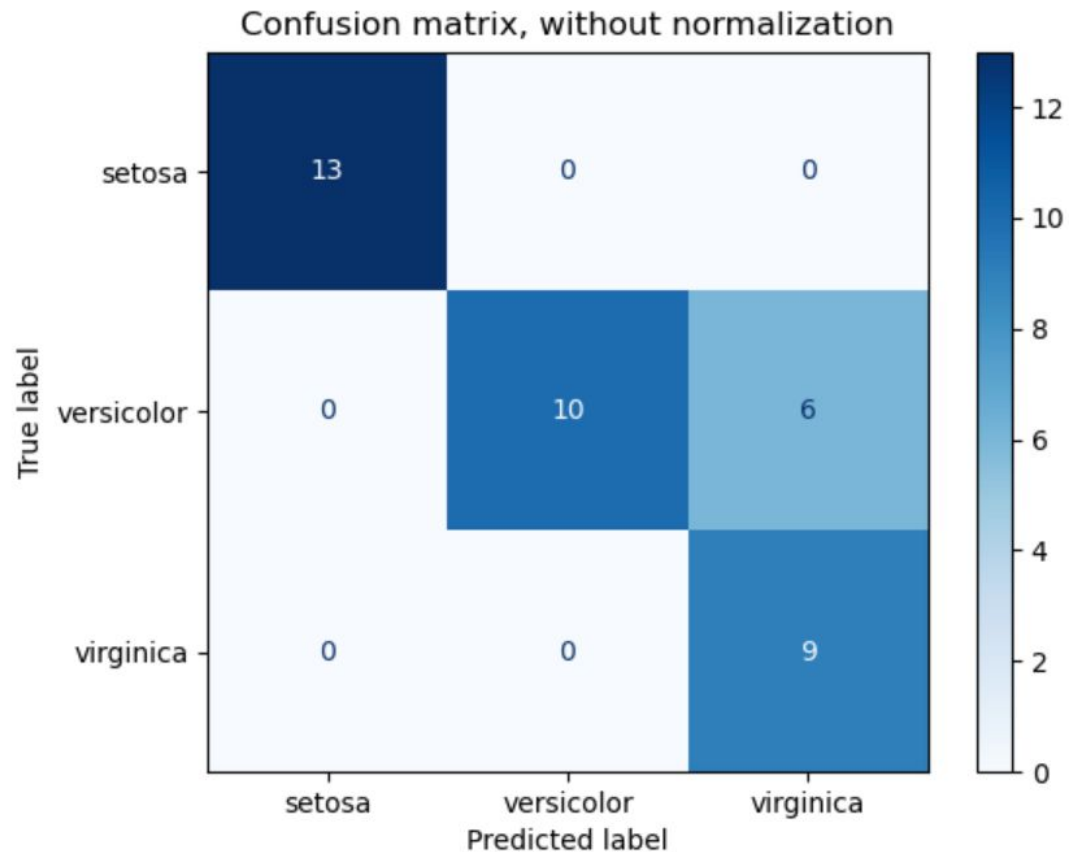


Sad



Joyful Power Peace Neutral Sad Scared Mad

# Confusion Matrix: BERT & Baseline



# Confusion Matrix: BERT & Baseline

	joyful	mad	neutral	peaceful	powerful	sad	scared
joyful	237	10	100	18	4	8	32
mad	49	62	45	3	4	9	36
neutral	92	25	393	37	11	13	65
peaceful	58	10	82	21	4	8	24
powerful	61	10	62	10	6	6	17
sad	12	7	45	11	2	27	32
scared	48	16	115	11	1	12	94

	joyful	mad	neutral	peaceful	powerful	sad	scared
joyful	287	2	109	0	4	0	7
mad	120	8	71	2	1	1	5
neutral	298	13	303	2	2	1	17
peaceful	124	1	76	0	2	0	4
powerful	93	4	65	1	4	1	4
sad	75	0	54	2	1	3	1
scared	149	1	126	2	1	1	17

# Confusion Matrix: SVM BoW & Tfidf

	joyful	mad	neutral	peaceful	powerful	sad	scared
joyful	150	22	172	22	9	9	25
mad	38	39	83	8	10	5	25
neutral	74	29	445	29	10	21	28
peaceful	39	14	104	17	9	7	17
powerful	37	7	78	13	13	7	17
sad	15	12	58	8	6	22	15
scared	39	28	151	19	5	7	48

	joyful	mad	neutral	peaceful	powerful	sad	scared
joyful	193	11	188	1	1	2	13
mad	47	33	110	2	2	0	14
neutral	104	12	488	2	1	13	16
peaceful	60	4	130	0	1	2	10
powerful	43	7	103	2	1	2	14
sad	20	8	83	1	1	12	11
scared	52	12	187	2	1	3	40



# Matrix summary

- Baseline tend to classify input as 'joyful' ( $1146/409 = 280\%$ ;  $287/409=70\%$ )
- SVM (both BoW & tfidf) tend to classify input as 'neutral' ( $1289/636=203\%$ ;  $488/636=77\%$ )
- BERT is good at predicting label 'mad' and 'scared'
- BERT and SVM BoW are comparatively better at predicting label 'peaceful' and 'sad'
- All four models rarely predict the label 'powerful', with an average of only 29 predictions in 2065 data containing 172 'powerful' instances.

Different models are good at predicting different labels → ensemble for improvement

# What is wrong with 'powerful'

I'm serious. You're amazing. You know when to spritz, when to lay back.

Really? You don't know what that means to me.



Power

I'm gonna go to the bathroom, maybe I'll see you there in a bit?

Kay!



Power

Wife: We're very sad that it didn't work out between you and Emily, monkey. But, I think you're (a third person) absolutely delicious.

Husband: Excuse me! I am standing right here!



Power

[Ross is having a rough time at his own wedding]

Ross: Actually, do you guys mind staying here for a while?

Monica: Ugh, y'know, umm we gotta get up early and catch that plane for New York.

Chandler: Yeah, it's a very large plane.



Power

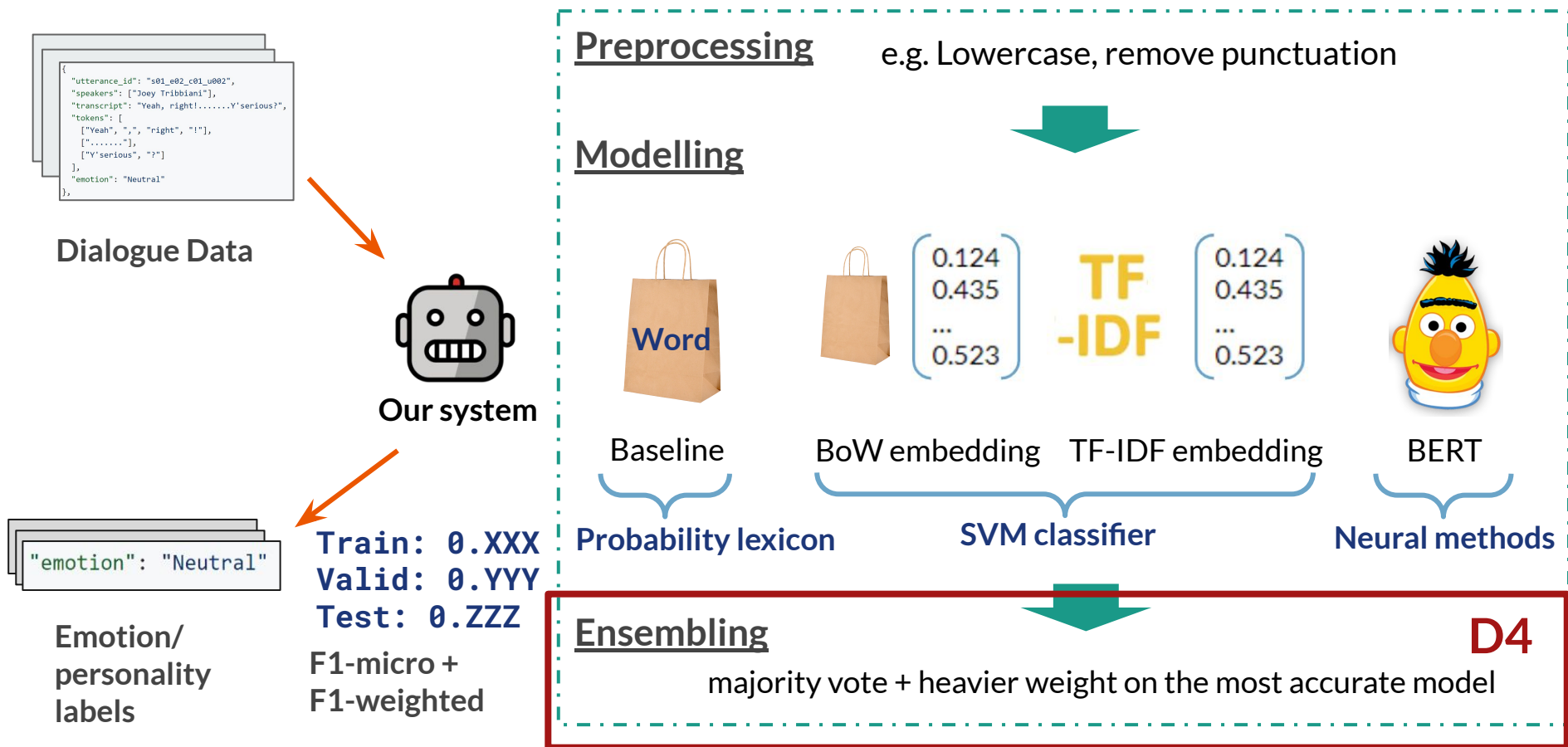
Ross: That's cool.



Power

1. Somewhat underrepresented in the training data (8.8%)
2. Heavily mislabeled

# System overview



# Change for adaptation task - Preprocessing

```
{  
  scene_id: s01_e01_c05  
  character: Paul  
  
  Text: <b>Monica Geller</b>: Oh my  
        God!<br><br>  
        <b>Paul</b>: I know, I know,  
        I'm such an idiot. I guess I ...  
        <br><br>  
  
  AGR: 1  
  CON: 0  
  EXT: 0  
  OPN: 0  
  NEU: 0  
}
```

Personality  
labels for  
**Paul**

Agreeable:
Conscientious:
Extraverted:
Open to experience:
Emotionally Stable:

Multi-parties  
dialogue

Our lexicon models  
(**non-neural models**) needs to  
extract the speaking lines from  
that certain speaker

# Change for primary task - Ensembling

Primary Task: Emotion (Eval)					
	Bert+Tfidf+ BoW	Bert+Tfidf+ lexicon	Bert+BoW+ lexicon	Bert+BoW+Tfidf+ lexicon	Bert
F1-micro ✓	<b>0.4000</b>	0.3855	0.3956	0.3961	0.3719
F1-weighted	0.3478	0.3324	0.3494	0.3332	<b>0.3684</b>

# Change for adaptation task - Preprocessing

```
{  
  scene_id: s01_e01_c05  
  character: Paul  
  
  Text: <b>Monica Geller</b>: Oh my  
        God!<br><br>  
        <b>Paul</b>: I know, I know,  
        I'm such an idiot. I guess I ...  
        <br><br>  
  
  AGR: 1  
  CON: 0  
  EXT: 0  
  OPN: 0  
  NEU: 0  
}
```

Personality  
labels for  
**Paul**

Agreeable:
Conscientious:
Extraverted:
Open to experience:
Emotionally Stable:

Multi-parties  
dialogue

Ideally, the new model needs to focus on the speaking lines from **that certain speaker**.  
But, that is not the case in our experiment.

# Change for adaptation task - Preprocessing for SVM

```
{  
  scene_id: s01_e01_c05  
  character: Paul  
  
  Text: <b>Monica Geller</b>: Oh my  
        God!<br><br>  
        <b>Paul</b>: I know, I know,  
        I'm such an idiot. I guess I ...  
        <br><br>  
  AGR: 1  
  CON: 0  
  EXT: 0  
  OPN: 0  
  NEU: 0  
}
```

Personality  
labels for  
**Paul**

Agreeable:

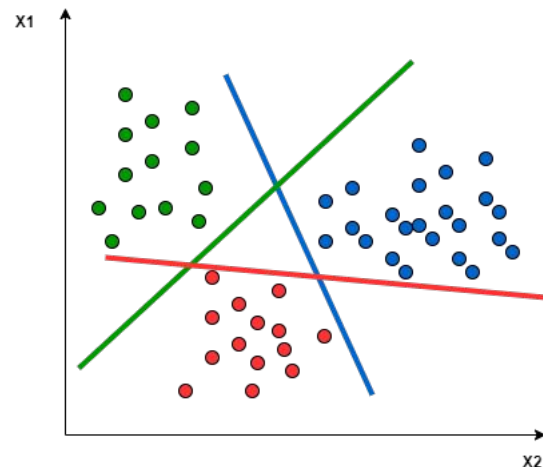
Conscientious:

Extraverted:

Open to experience:

Emotionally Stable:

Multi-parties  
dialogue



Instead of classifying the data points into  $2^5 = 32$  classes, use 5 SVMs that classifies the point into 2 classes for each personality.

# Results: Baseline method - primary & adaptation

Primary Task: Emotion			Adaptation Task: Personality		
	<i>Eval</i>	<i>Test</i>		<i>Eval</i>	<i>Test</i>
F1-micro ✓	0.3021	0.3143	Accuracy	0.5825	0.5636
F1-weighted	0.2304	0.2346			



# Results: BoW-SVM method - primary & adaptation

Primary Task: Emotion			Adaptation Task: Personality		
	<i>Eval</i>	<i>Test</i>		<i>Eval</i>	<i>Test</i>
F1-micro ✓	0.3554	0.3530	Accuracy	0.5825	0.5836
F1-weighted	0.3174	0.3131			

# Results: Tfidf-SVM method - primary & adaptation

Primary Task: Emotion			Adaptation Task: Personality		
	<i>Eval</i>	<i>Test</i>		<i>Eval</i>	<i>Test</i>
F1-micro ✓	0.3714	0.3567	Accuracy	0.5807	0.5655
F1-weighted	0.2990	0.2900			

# Result on BERT

Five repeated experiments on each hyperparameter

Hyperparameter - num of epoch		
<i>num_epoch</i>	<i>F1-micro</i>	<i>F1-weighted</i>
5	0.36975 (0.0148)	<b>0.368825 (0.0094)</b>
10	0.36494 (0.018)	0.36286 (0.0142)
20	<b>0.3735 (0.0131)</b>	0.36665 (0.0092)
Hyperparameter - learning rate		
<i>learning rate</i>	<i>F1-micro</i>	<i>F1-weighted</i>
1.00E-05	<b>0.37205 (0.0103)</b>	0.36185 (0.0103)
3.00E-05	0.36494 (0.018)	<b>0.36286 (0.0142)</b>
5.00E-05	0.363925 (0.018)	0.361825 (0.0138)
Hyperparameter - optimizer		
<i>optimizer</i>	<i>F1-micro</i>	<i>F1-weighted</i>
Adam	0.36494 (0.018)	0.36286 (0.0142)
AdamW	<b>0.3792 (0.0109)</b>	<b>0.3697 (0.0072)</b>
Hyperparameter - effective batch size		
<i>batch size</i>	<i>F1-micro</i>	<i>F1-weighted</i>
20	0.371175 (0.003)	0.361675 (0.0035)
40	<b>0.373375 (0.0149)</b>	<b>0.36295 (0.0152)</b>
60	0.36494 (0.018)	0.36286 (0.0142)



- Number of epochs: 10
- Learning rate: 1e-5
- Optimizer: AdamW
- Effective batch size: 40

Hyperparameter tuning on  
**Emotion task**

Same optimized sets on  
**personality task**

## Adaptation task:

- Baseline (default)
  - Number of epochs: 10
  - Learning rate:  $3e-5$
  - Optimizer: Adam
  - Effective batch size: 60
- Modification (tuning)
  - Number of epochs: 10
  - Learning rate:  $1e-5$
  - Optimizer: AdamW
  - Effective batch size: 40

Validation Accuracy: 0.5596



Validation Accuracy: 0.5754



Also have improvement on Adaptation task!

# Result on BERT

Primary Task: Emotion			Adaptation Task: Personality		
	<i>Eval</i>	<i>Test</i>		<i>Eval</i>	<i>Test</i>
F1-micro ✓	0.3719	0.4111	Accuracy	0.5754	0.5564
F1-weighted	0.3684	0.3822			

# Result Comparison before and after ensembling

Best model:

Primary Task: Emotion			Adaptation Task: Personality		
	<i>Eval</i>	<i>Test</i>		<i>Eval</i>	<i>Test</i>
<b>F1-micro</b> ✓	0.3719	0.4111*	<b>Accuracy</b>	0.5825*	0.5836*
<b>F1-weighted</b>	0.3684	0.3822			

Ensembling:

Primary Task: Emotion			Adaptation Task: Personality		
	<i>Eval</i>	<i>Test</i>		<i>Eval</i>	<i>Test</i>
<b>F1-micro</b> ✓	0.4000*	0.3933	<b>Accuracy</b>	0.5701	0.5688
<b>F1-weighted</b>	0.3478	0.3386			

# Issues and successes

Why personality task is less accurate:

1. The dataset of personality task is **too small**

	Train	Dev	Test
<b>emotion</b>	8629	2065	1912
<b>personality</b>	487	114	110



2. Nature of personality task: **subjective!**

	# of annotators	Inter-annotator agreements (pairwise)
<b>emotion</b>	4	85.09%
<b>personality</b>	3	54.92% 