

# 3M-Health: Multimodal Multi-Teacher Knowledge Distillation for Mental Health Detection\*

## Supplementary Material

Rina Carines Cabral  
The University of Sydney  
Sydney, NSW, Australia  
rcab5321@uni.sydney.edu.au

Josiah Poon  
The University of Sydney  
Sydney, NSW, Australia  
josiah.poon@sydney.edu.au

Siwen Luo  
The University of Western Australia  
Perth, WA, Australia  
siwen.luo@uwa.edu.au

Soyeon Caren Han<sup>†</sup>  
The University of Melbourne  
Melbourne, VIC, Australia  
caren.han@unimelb.edu.au

### A HYPERPARAMETER SEARCH

We tune each individual model for our experiments using Optuna<sup>1</sup> for 50 trials and optimise the weighted F1 score. Final model construction is done separately from model tuning using a 90:10 validation split. Table A1 details the parameters and the search space used. Table A2 and A3 enumerates the best-found hyperparameters for the final teacher models and student models, respectively.

**Table A1: Hyperparameter Search Space**

Parameter	Search Space
<b>Text-based Teachers</b>	
Dropout	{0.01, 0.05, 0.1, 0.5}
Hidden Layers	{2, 4, 6, 8, 10, 12}
Attention Heads	{2, 4, 6, 8, 12}
Learning Rate	{1e-04, 1e-05, 2e-05, 3e-05, 4e-05, 5e-05}
Weight Decay	{0, 0.01, 0.1}
Epochs	[2-5]
<b>Emotion-based Teachers</b>	
Dropout	{0.01, 0.05, 0.1, 0.5}
Hidden Layers	[2-5]
Hidden Dimension	{100, 200, 300, 400, 500}
Learning Rate	{1e-03, 1e-04, 1e-05}
Weight Decay	{0, 0.01, 0.1}
<b>Audio-based Teachers</b>	
Dropout	{0.01, 0.05, 0.1, 0.5}
Hidden Layers	{2, 4, 6, 8, 10, 12}
Attention Heads	{2, 4, 6, 8, 12}
Learning Rate	{1e-03, 1e-04, 1e-05, 5e-05}
Scheduler Patience	[2-5]
Scheduler Factor	{0.1, 0.5}
<b>Student</b>	
Dropout	{0.01, 0.05, 0.1, 0.5}
Learning Rate	{1e-04, 1e-05, 2e-05, 3e-05, 4e-05, 5e-05}
Weight Decay	{0, 0.01, 0.1}
Hidden Layers	{2, 4, 6, 8, 10, 12}
Attention Heads	{2, 4, 6, 8, 12}
Activation	{“relu”, “gelu”}
Epochs	[3-5]

**Table A2: Best found teacher hyperparameters.**

Parameters	TwitSuicide	DEPTWEET	IdenDep	SDCNL
<b>Text-based Teachers</b>				
Language Model	BERT	MentalBERT	BERT	BERT
Dropout	0.01	0.01	0.05	0.01
Weight Decay	0	0	0	0
Learning Rate	4e-05	4e-05	4e-05	5e-05
Epochs	4	3	5	4
Hidden Layers	10	8	2	4
Attention Heads	8	8	6	12
Batch Size	128	64	64	64
<b>Emotion-based Teachers</b>				
Dropout	0.1	0.1	0.05	0.01
Weight Decay	0	0.01	0.1	0
Learning Rate	1e-05	1e-04	1e-04	1e-05
Hidden Layers	2	4	3	5
Hidden Dim	400	400	100	100
Batch Size	128	64	64	64
<b>Audio-based Teachers</b>				
Dropout	0.1	0.1	0.01	0.01
Learning Rate	5e-05	1e-05	1e-05	5e-05
Hidden Layers	8	6	6	8
Attention Heads	4	6	8	8
Scheduler Patience	4	4	3	4
Scheduler Factor	0.5	0.5	0.5	0.1
Batch Size	32	32	32	32

\*Warning: This paper contains examples that show suicide ideation and depression.

<sup>†</sup>Corresponding author.

<sup>1</sup><https://optuna.org/>

**Table A3: Best found student hyperparameters for different combinations of teacher modalities.**

Parameters	TwitSuicide	DEPTWEET	IdenDep	SDCNL
<b>Text ✓ Emo ✓ Aud ✓</b>				
Dropout	0.05	0.01	0.1	0.05
Learning Rate	1e-04	3e-05	5e-05	5e-05
Weight Decay	0	0	0	0.01
Hidden Layers	10	6	10	12
Attention Heads	3	12	8	12
Activation	gelu	gelu	gelu	gelu
Epochs	3	5	3	3
<b>Text ✓ Emo ✓ Aud ×</b>				
Dropout	0.1	0.01	0.1	0.05
Learning Rate	1e-04	3e-05	5e-05	4e-05
Weight Decay	0.01	0	0	0
Hidden Layers	10	10	10	12
Attention Heads	12	12	6	3
Activation	gelu	gelu	gelu	gelu
Epochs	4	5	3	5
<b>Text ✓ Emo × Aud ✓</b>				
Dropout	0.05	0.05	0.05	0.05
Learning Rate	1e-04	1e-04	4e-05	5e-05
Weight Decay	0	0	0	0
Hidden Layers	12	4	12	12
Attention Heads	4	8	12	12
Activation	gelu	relu	gelu	gelu
Epochs	4	3	4	5
<b>Text ✓ Emo × Aud ×</b>				
Dropout	0.1	0.01	0.05	0.05
Learning Rate	1e-04	4e-05	1e-4	4e-05
Weight Decay	0	0	0	0
Hidden Layers	12	12	4	10
Attention Heads	8	6	4	12
Activation	relu	gelu	gelu	relu
Epochs	3	4	5	4

## B AUDIO ANALYSIS SAMPLES

**Content Warning:** The following section shows samples of posts indicative of suicide and depression, which may be triggering for some people.

We provide the textual posts associated with each sample presented in Section 5.2 Figure 5 to refer to for detailed comparison and analysis. Table B1 enumerates the post ID, class, and text for the samples from the SDCNL dataset. TwitSuicide, DEPTWEET and IdenDep samples are found in Tables 5 to 7 of the main manuscript. Following the proposed ethical protocols on social media research of [1], usernames and links are masked with special tokens `_USER_` and `_URL_` to protect the identity and privacy of each author. For instance, “@myusername this is the link `http://urlsamp.le`” is masked

as “\_USER\_ this is the link \_URL\_”. Moreover, we mask parts of the text with “\*\*\*” to prevent possible reverse searches.

**Table B1: Samples for the SDCNL audio spectrogram analysis. Each sample has been masked to avoid a reverse search of each post. DEP: Depression; SUI: Suicide**

SDCNL		
ID	Class	Text
<b>0-10 seconds</b>		
39	DEP	*** good qualities in therapist that i can spot early on? ***
777	DEP	I'm *** okay I probably won't, but tonight *** gotta tell myself that. *** until sunrise
914	DEP	Another attempt ***. Hopefully I don't survive
1143	DEP	*** tired of being told it gets better *** never has and never *** will
1278	DEP	How do i know if i have depression *** i just wanna know *** i'm getting close
51	SUI	I'm scared *** alone *** don't know what to do ***
599	SUI	Gave my note to my family *** don't know what to do now
967	SUI	Why do people think they can help *** on ending it all? ***
1459	SUI	*** sleep forever and never want to wakeup again..... ***
1494	SUI	Picking up a gun *** relieved
<b>10-25 seconds</b>		
562	DEP	Songs? *** songs/artist *** feel better/happy when you feel depressed?
786	DEP	*** freaking out constantly *** so lonely Even when I'm around people I'm lonely. I feel crazy. I can't stop looking *** at my awful life, *** such poor judgment
1209	DEP	Anyone want to chat? I'm not in a good place *** Just wondering *** find comfort in each other.
1668	DEP	Any good songs *** to recommend? Any song *** which makes you feel good. Have a good day!
98	SUI	Suicide *** 4 ways I can go... *** overdose *** cop *** bridge *** train
251	SUI	Called off my work , going to end it *** Nice knowing you all I'm finally coming home *** and *** ❤️
757	SUI	*** want to have someone. It's all I've *** wanted *** keeps eating away at me and won't change. *** don't want to be alone anymore.
877	SUI	Wanna kill myself tonight. Having suicidal thoughts *** now. I wanna end it *** slit my wrists. After being on this earth *** I'm done.
889	SUI	*** end it all today No more suffering *** humiliation *** worrying about the future, *** such a coward i wanna die so bad
1864	SUI	Please someone help me I need *** an effective and painless method *** tell me I can't last longer.

## REFERENCES

- [1] Adrian Benton, Glen Coppersmith, and Mark Dredze. 2017. Ethical Research Protocols for Social Media Health Research. In *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, Dirk Hovy, Shannon Spruit, Margaret Mitchell, Emily M. Bender, Michael Strube, and Hanna Wallach (Eds.). Association for Computational Linguistics, Valencia, Spain, 94–102. <https://doi.org/10.18653/v1/W17-1612>