

## APPENDIX A — LIST OF MOST FREQUENT TERMS FOR ANXIETY AND DEPRESSION DISORDERS (DSM-5)

The tables A.1 and A.2 list the most frequent terms extracted from the description for Anxiety and Depression Disorders, respectively, according to the DSM-5 manual. The tables highlight in orange the 7 common in both disorders. In yellow, 4 terms of the same meaning are highlighted, but with different syntaxes.

Table A.1: Anxiety Disorders: List of Terms as described in DSM-5.

List of most frequent terms					
anxious	agoraphobia	apprehensive	arousal	attacks	avoid
avoidance	blank	clinging	control	crazy	crying
danger	depersonalization	difficult	discomfort	distress	escape
excessive	failure	falling	fear	feared	fearful
feelings	freezing	future	headache	help	humiliating
impairment	incapacitating	intense	interactions	irritability	losing
mind	mutism	nervios	neuroticism	palpitations	panic
paresthesias	persistent	phobia	restlessness	risk	screaming
shrinking	sleep	soreness	stress	tantrums	tension
thoughts	trembling	uncontrollable	unexpected	worry	

Table A.2: Depressive Disorders: List of Terms as described in DSM-5.

List of most frequent terms					
aggression	angry	antidepressant	anxiety	appetite	behavioral
bereavement	bereavementrelated	capacity	careful	change	concentration
concerns	death	depressive	difficulty	disgust	displeasure
distress	disturbance	dyscontrol	dysphoric	dysregulation	dysthymia
empty	energy	explosive	extreme	facilitated	failure
fatigue	feelings	frustration	grief	guilt	hopelessness
hypersomnia	impairment	insomnia	interest	irritable	loss
mood	outbursts	overeating	pain	persistent	pleasure
retardation	sad	selfesteem	sleep	suffering	suicide
thoughts	timing	vulnerabilities	worse	worthlessness	

## APPENDIX B — PERFORMANCE ANALYSIS: EXPERT CLASSIFIERS

In the experiment for the formation of specialist classifiers, the performance analysis compared the performance of these models when developed using the LSTM, CNN and Hybrid architectures by target condition. This experiment used a basic settings for each architecture, where all models use the same pre-trained GloVe 6B *embedding*. The results in terms of Average (AVG) and Standard Deviation (SD) for the repetitions of these experiments are presented in Table B.1. It should be noted that after this statistical analysis, for each target condition, new experiments were carried out to fine-tune the explored architectures resulting in the final performance presented in this article.

Table B.1 allows us to verify that, in general, the models maintain a regular performance through the repetitions, presenting a low standard deviation for the analyzed metrics. Among the models explored, it is noted that, in general, the CNN architecture presents the least variation, in terms of standard deviation, considering the metrics analyzed for all target conditions. An inverse behavior is observed for hybrid models, in all target conditions.

Table B.1: Expert Classifiers: Average (AVG) and e Standard Deviation (SD).

Mental Disorder	Architecture	Precision		Recall		F1	
		AVG	SD	AVG	SD	AVG	SD
Anxiety	LSTM	0.77	0.03	0.73	0.05	0.75	0.02
	CNN	0.71	0.00	0.80	0.00	0.75	0.00
	Hybrid	0.67	0.06	0.69	0.05	0.68	0.03
Depression	LSTM	0.77	0.02	0.72	0.04	0.74	0.01
	CNN	0.82	0.04	0.68	0.07	0.74	0.02
	Hybrid	0.68	0.08	0.77	0.06	0.72	0.04
Comorbidity	LSTM	0.73	0.02	0.66	0.02	0.69	0.01
	CNN	0.82	0.00	0.70	0.01	0.76	0.01
	Hybrid	0.58	0.08	0.73	0.05	0.64	0.04

Table B.2, presents the result of the statistical test in terms of *textit p-value* for each metric analyzed. The paired two-tailed Student T Test (*alpha* = 0.5) was adopted to establish the performance comparison between two architectures, considering each target condition. It is worth remembering that, as a null hypothesis, there is no difference between two architectures compared for the task of classifying a given target condition. Thus, Table B.2 presents the *p-value* resulting from this comparison for the metrics Precision, Revocation and measure F (F1) between architecture combinations, by target condition.

The results presented in Table B.1 show that all performance comparisons performed between the architectures registered statistically significant differences for the task

Table B.2: Comparisons between Architectures: Result of the Student T-Test.

Mental Disorder	Architecture	<i>P-Value</i>		
		Precision	Recall	F1
Anxiety	LSTM vs CNN	0.00	0.00	0.28
	LSTM vs Hybrid	0.00	0.17	0.00
	CNN vs Hybrid	0.09	0.00	0.00
Depression	LSTM vs CNN	0.00	0.20	0.87
	LSTM vs Hybrid	0.01	0.04	0.05
	CNN vs Hybrid	0.00	0.01	0.08
Comorbidity	LSTM vs CNN	0.00	0.00	0.00
	LSTM vs Hybrid	0.00	0.00	0.00
	CNN vs Hybrid	0.00	0.12	0.00

of classifying the same mental condition. These results show that there is a difference in performance between the different architectures explored, which may result in a gain in variability for the proposed committee model, as the significant differences observed reflect a *trade-off* between accuracy and recall when considering comparisons between two architectures.

## APPENDIX C — PERFORMANCE ANALYSIS: *DAC STACKING* E VARIATIONS

The quantitative analyzes presented compare the performance of the *DAC Stacking* models in terms of topology (homogeneous vs heterogeneous) and the proposed variations (Base, EXP and DIF). Tables C.1, C.2 and C.3 show the average repetition value resulting from the cross-validation training process, together with the standard deviation values, for the models *DAC Base* and its variations *EXP* and *DIF*, respectively.

To confirm whether the performance differences found were statistically significant, the paired two-tailed Student T-test ( $\alpha = 0.5$ ) was adopted. The results of this test for the performance comparisons between the models *baseline* and the models *DAC Base* and variations *EXP* and *DIF* are presented in Tables C.4, C.5, and C.6, respectively. Tables C.7, C.8, and C.9 present the statistical results for the performance comparisons, considering the topology variations (homogeneous vs heterogeneous) for the models *DAC Base*, *DAC EXP* and *DAC DIF*, respectively.

Finally, the table C.10 presents the performance comparison between the best result, *DAC Base C*, and the models given for each proposed variation *DAC EXP* and *DAC DIF*.

Table C.1: *DAC Base*: Average (AVG) and Standard Deviation (SD).

Model	Multi-label						Control						Anxiety						Depression						Comorbidity					
	EMR			HL			Precision			Recall			F1			Precision			Recall			F1			Precision			Recall		
	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD
<i>DAC Base</i>	0.46	0.00	0.29	0.00	0.67	0.01	0.77	0.01	0.72	0.00	0.62	0.01	0.69	0.03	0.65	0.01	0.57	0.00	0.66	0.02	0.66	0.02	0.61	0.01	0.33	0.00	0.72	0.02	0.45	0.00
L	0.42	0.01	0.28	0.00	0.74	0.02	0.63	0.03	0.68	0.01	0.62	0.02	0.73	0.06	0.67	0.02	0.58	0.02	0.72	0.04	0.64	0.01	0.33	0.01	0.75	0.02	0.46	0.01		
H	0.42	0.02	0.29	0.00	0.75	0.01	0.64	0.05	0.69	0.03	0.60	0.02	0.74	0.01	0.67	0.01	0.57	0.01	0.67	0.01	0.61	0.01	0.33	0.00	0.76	0.00	0.46	0.00		
LCH	0.40	0.02	0.30	0.02	0.75	0.01	0.59	0.07	0.66	0.04	0.58	0.04	0.81	0.04	0.67	0.01	0.56	0.01	0.71	0.06	0.63	0.01	0.32	0.02	0.79	0.05	0.46	0.01		
Baseline L	0.33	0.04	0.37	0.03	0.57	0.05	0.63	0.03	0.60	0.04	0.57	0.01	0.39	0.14	0.45	0.10	0.55	0.01	0.35	0.16	0.41	0.11	0.29	0.01	0.37	0.14	0.32	0.04		
Baseline C	0.28	0.12	0.33	0.03	0.73	0.06	0.48	0.28	0.53	0.24	0.62	0.01	0.36	0.08	0.45	0.06	0.62	0.02	0.37	0.13	0.46	0.09	0.33	0.01	0.36	0.09	0.34	0.04		
Baseline H	0.31	0.04	0.36	0.01	0.60	0.03	0.52	0.13	0.54	0.05	0.54	0.04	0.47	0.16	0.49	0.08	0.55	0.03	0.46	0.09	0.50	0.04	0.30	0.02	0.44	0.11	0.35	0.02		

Table C.2: *DAC EXP*: Average (AVG) and Standard Deviation (SD).

Model	Multi-label						Control						Anxiety						Depression						Comorbidity					
	EMR			HL			Precision			Recall			F1			Precision			Recall			F1			Precision			Recall		
	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD
<i>DAC EXP</i>	0.46	0.01	0.29	0.00	0.67	0.01	0.78	0.02	0.72	0.00	0.61	0.01	0.69	0.02	0.65	0.00	0.57	0.01	0.64	0.04	0.60	0.02	0.32	0.01	0.71	0.03	0.45	0.01		
L	0.42	0.01	0.28	0.01	0.74	0.01	0.66	0.06	0.69	0.03	0.63	0.02	0.68	0.09	0.65	0.04	0.58	0.02	0.71	0.05	0.64	0.01	0.33	0.01	0.72	0.09	0.45	0.01		
C	0.43	0.02	0.29	0.01	0.72	0.03	0.68	0.06	0.70	0.02	0.60	0.01	0.76	0.04	0.67	0.01	0.56	0.02	0.69	0.05	0.62	0.01	0.32	0.02	0.76	0.04	0.44	0.01		
H	0.41	0.02	0.29	0.02	0.75	0.01	0.61	0.05	0.67	0.03	0.62	0.04	0.73	0.10	0.66	0.02	0.57	0.04	0.69	0.11	0.62	0.02	0.33	0.03	0.75	0.11	0.46	0.01		
LCH	0.33	0.04	0.37	0.03	0.57	0.05	0.63	0.03	0.60	0.04	0.57	0.01	0.39	0.14	0.45	0.10	0.55	0.01	0.35	0.16	0.41	0.11	0.29	0.01	0.37	0.14	0.32	0.04		
Baseline L	0.28	0.12	0.33	0.03	0.73	0.06	0.48	0.28	0.53	0.24	0.62	0.01	0.36	0.08	0.45	0.06	0.62	0.02	0.37	0.13	0.46	0.09	0.33	0.01	0.36	0.09	0.34	0.04		
Baseline C	0.31	0.04	0.36	0.01	0.60	0.03	0.52	0.13	0.54	0.05	0.54	0.04	0.47	0.16	0.49	0.08	0.55	0.03	0.46	0.09	0.50	0.04	0.30	0.02	0.44	0.11	0.35	0.02		

Table C.3: *DAC DIF*: Average (AVG) and Standard Deviation (SD).

Model	Multi-label						Control						Anxiety						Depression						Comorbidity					
	EMR			HL			Precision			Recall			F1			Precision			Recall			F1			Precision			Recall		
	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD	AVG	SD
<i>DAC DIF</i>	0.46	0.00	0.29	0.00	0.68	0.01	0.77	0.01	0.72	0.00	0.61	0.01	0.71	0.02	0.66	0.01	0.57	0.00	0.66	0.01	0.61	0.00	0.32	0.00	0.73	0.02	0.45	0.00		
L	0.42	0.02	0.28	0.01	0.73	0.01	0.65	0.05	0.69	0.03	0.62	0.02	0.72	0.07	0.67	0.02	0.57	0.01	0.72	0.03	0.63	0.01	0.33	0.01	0.75	0.05	0.46	0.01		
C	0.42	0.02	0.29	0.01	0.74	0.00	0.62	0.04	0.67	0.03	0.60	0.02	0.76	0.09	0.66	0.03	0.58	0.05	0.69	0.16	0.62	0.06	0.32	0.01	0.69	0.17	0.43	0.04		
LCH-LC	0.40	0.02	0.30	0.01	0.74	0.01	0.60	0.05	0.66	0.03	0.59	0.04	0.76	0.10	0.66	0.02	0.55	0.01	0.76	0.04	0.64	0.01	0.31	0.01	0.77	0.09	0.44	0.01		
LCH-C	0.33	0.04	0.37	0.03	0.57	0.05	0.63	0.03	0.60	0.04	0.57	0.01	0.39	0.14	0.45	0.10	0.55	0.01	0.35	0.16	0.41	0.11	0.29	0.01	0.37	0.14	0.32	0.04		
Baseline L	0.28	0.12	0.33	0.03	0.73	0.06	0.48	0.28	0.53	0.24	0.62	0.01	0.36	0.08	0.45	0.06	0.62	0.02	0.37	0.13	0.46	0.09	0.33	0.01	0.36	0.09	0.34	0.04		
Baseline C	0.31	0.04	0.36	0.01	0.60	0.03	0.52	0.13	0.54	0.05	0.54	0.04	0.47	0.16	0.49	0.08	0.55	0.03	0.46	0.09	0.50	0.04	0.30	0.02	0.44	0.11	0.35	0.02		

Table C.4: *DAC Base*: Statistical comparison with baselines.

Comparison between Models	Student T-Test: <i>P-Value</i>									
	Multi-label (EMR)		Control		Anxiety		Depression		Comorbidity	
	HL	P	R	F1	P	R	P	R	P	F1
DAC Base L vs Baseline L	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.01	0.00
DAC Base C vs Baseline C	0.05	0.03	0.94	0.27	0.23	0.69	0.00	0.00	0.01	0.23
DAC Base H vs Baseline H	0.01	0.00	0.00	0.06	0.00	0.01	0.02	0.01	0.28	0.00
DAC Base LCH vs Baseline L	0.02	0.02	0.00	0.23	0.06	0.51	0.00	0.01	0.09	0.00
DAC Base LCH vs Baseline C	0.09	0.05	0.51	0.42	0.27	0.08	0.00	0.00	0.01	0.02
DAC Base LCH vs Baseline H	0.02	0.00	0.00	0.25	0.02	0.21	0.01	0.01	0.49	0.00

Table C.5: *DAC EXP*: Statistical comparison with baselines.

Comparison between Models	Student T-Test: <i>P-Value</i>									
	Multi-label (EMR)		Control		Anxiety		Depression		Comorbidity	
	HL	P	R	F1	P	R	P	R	P	F1
DAC EXP L vs Baseline L	0.00	0.01	0.02	0.00	0.00	0.01	0.01	0.01	0.01	0.00
DAC EXP C vs Baseline C	0.05	0.02	0.94	0.19	0.18	0.34	0.00	0.00	0.04	0.00
DAC EXP H vs Baseline H	0.01	0.00	0.01	0.12	0.01	0.03	0.02	0.01	0.75	0.00
DAC EXP LCH vs Baseline L	0.01	0.00	0.00	0.50	0.01	0.05	0.01	0.01	0.14	0.00
DAC EXP LCH vs Baseline C	0.09	0.06	0.63	0.38	0.25	0.90	0.01	0.00	0.06	0.02
DAC EXP LCH vs Baseline H	0.02	0.00	0.00	0.23	0.01	0.02	0.01	0.00	0.22	0.00

Table C.6: *DAC DIF*: Statistical comparison with baselines.

Comparison between Models	Student T-Test: <i>P-Value</i>									
	Multi-label (EMR)		Control		Anxiety		Depression		Comorbidity	
	HL	P	R	F1	P	R	P	R	P	F1
DAC DIF L vs Baseline L	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.02	0.00
DAC DIF C vs Baseline C	0.06	0.03	0.86	0.28	0.23	0.78	0.00	0.00	0.01	0.62
DAC DIF LCH-LC vs Baseline L	0.01	0.00	0.00	0.67	0.03	0.05	0.01	0.01	0.23	0.02
DAC DIF LCH-LC vs Baseline C	0.06	0.09	0.79	0.34	0.25	0.19	0.00	0.00	0.22	0.00
DAC DIF LCH-LC vs Baseline H	0.01	0.00	0.00	0.14	0.01	0.00	0.00	0.00	0.36	0.08
DAC DIF LCH-C vs Baseline L	0.01	0.01	0.00	0.26	0.02	0.22	0.01	0.01	0.18	0.00
DAC DIF LCH-C vs Baseline C	0.09	0.08	0.73	0.42	0.28	0.20	0.00	0.00	0.01	0.01
DAC DIF LCH-C vs Baseline H	0.01	0.00	0.00	0.19	0.01	0.05	0.00	0.00	0.73	0.00

Table C.7: Statistical comparison between variations of *DAC Base*.

Comparison between Models <i>DAC Base</i>	Student T-Test: <i>P-Value</i>													
	Multi-label		Control			Anxiety			Depression			Comorbidity		
	EMR	HL	P	R	F1	P	R	F1	P	R	F1	P	R	F1
L vs C	0.00	0.01	0.00	0.00	0.00	0.50	0.21	0.09	0.47	0.02	0.01	0.21	0.14	0.03
L vs H	0.00	0.43	0.00	0.01	0.08	0.11	0.02	0.11	0.18	0.18	1.00	0.37	0.02	0.00
L vs LCH	0.00	0.36	0.00	0.00	0.03	0.04	0.00	0.02	0.14	0.05	0.00	0.37	0.03	0.21
C vs H	0.38	0.11	0.12	0.65	0.22	0.08	0.77	0.50	0.30	0.03	0.00	0.37	0.37	1.00
C vs LCH	0.06	0.11	0.02	0.18	0.32	0.09	0.08	0.85	0.08	0.54	0.14	0.23	0.11	0.59
H vs LCH	0.21	0.29	1.00	0.12	0.10	0.28	0.02	0.37	0.21	0.14	0.07	0.27	0.23	0.37

Table C.8: Statistical comparison between variations of *DAC EXP*.

Comparison between Models <i>DAC EXP</i>	Student T-Test: <i>P-Value</i>													
	Multi-label		Control			Anxiety			Depression			Comorbidity		
	EMR	HL	P	R	F1	P	R	F1	P	R	F1	P	R	F1
L vs C	0.01	0.01	0.00	0.01	0.11	0.02	0.71	0.90	0.34	0.01	0.00	0.19	0.79	0.47
L vs H	0.06	0.82	0.02	0.03	0.13	0.34	0.03	0.01	0.20	0.17	0.24	0.59	0.15	0.70
L vs LCH	0.00	0.57	0.00	0.01	0.02	0.61	0.53	0.28	0.92	0.40	0.29	0.58	0.55	0.23
C vs H	0.45	0.09	0.51	0.69	0.74	0.06	0.17	0.37	0.22	0.59	0.03	0.29	0.45	0.46
C vs LCH	0.27	0.47	0.24	0.40	0.50	0.61	0.56	0.57	0.82	0.76	0.24	1.00	0.75	0.18
H vs LCH	0.00	0.59	0.14	0.05	0.06	0.39	0.57	0.69	0.21	0.96	0.50	0.33	0.78	0.24

Table C.9: Statistical comparison between variations of *DAC DIF*.

Comparison between Models <i>DAC Stacking DIF</i>	Student T-Test: <i>P-Value</i>													
	Multi-label		Control			Anxiety			Depression			Comorbidity		
	EMR	HL	P	R	F1	P	R	F1	P	R	F1	P	R	F1
L vs C	0.01	0.16	0.00	0.00	0.03	0.28	0.84	0.23	1.00	0.00	0.00	0.16	0.35	0.07
L vs LCH-LC	0.02	0.97	0.00	0.00	0.03	0.24	0.28	0.55	0.68	0.69	0.84	0.54	0.61	0.47
L vs LCH-C	0.00	0.20	0.00	0.00	0.01	0.26	0.29	0.51	0.02	0.01	0.01	0.14	0.36	0.18
C vs LCH-LC	0.90	0.40	0.11	0.45	0.60	0.17	0.39	0.87	0.71	0.76	0.58	0.46	0.54	0.19
C vs LCH-C	0.00	0.01	0.03	0.00	0.00	0.03	0.32	0.89	0.05	0.12	0.48	0.01	0.62	0.01
LCH-LC vs LCH-C	0.39	0.36	0.62	0.65	0.66	0.62	0.76	1.00	0.31	0.42	0.50	0.16	0.39	0.76

Table C.10: Statistical comparison between *DAC Base C* and Variations.

Comparison between DAC Base C (C) e Variations (EXP and DIF)	Student T-Test: <i>P-Value</i>													
	Multi-label		Control			Anxiety			Depression			Comorbidity		
	EMR	HL	P	R	F1	P	R	F1	P	R	F1	P	R	F1
C vs EXP L	0.01	0.00	0.00	0.00	0.00	0.02	0.23	0.04	0.62	0.01	0.01	0.03	0.01	0.00
C vs EXP C	0.79	0.47	1.00	0.43	0.35	0.37	0.12	0.12	0.79	0.78	0.75	0.70	0.42	0.29
C vs EXP H	0.46	0.06	0.54	0.27	0.15	0.04	0.44	0.75	0.06	0.07	0.00	0.15	0.69	0.08
C vs EXP LCH	0.18	0.51	0.07	0.49	0.88	0.81	1.00	0.33	0.88	0.54	0.24	0.90	0.94	0.59
C vs DIF L	0.00	0.02	0.00	0.00	0.00	0.14	0.63	0.20	0.62	0.03	0.01	0.02	0.28	0.00
C vs DIF C	0.78	0.72	0.43	0.45	0.39	1.00	0.83	0.63	0.69	0.76	0.37	1.00	0.94	1.00
C vs DIF LCH-LC	0.80	0.11	0.47	0.60	0.89	0.14	0.50	0.48	0.80	0.64	0.44	0.03	0.43	0.20
C vs DIF LCH-C	0.02	0.05	0.18	0.15	0.28	0.19	0.53	0.54	0.00	0.20	1.00	0.03	0.54	0.03