

Analyzing Prediction of Depression and Anxiety on Reddit: a Multi-task Learning Approach through GMMTL

Shailik Sarkar, Abdulaziz Alhamadani, Srithi Behal, Lulwah Alkulaib, Chang-Tien Lu

Background & Problem Setting

- Using data different subreddits to redesign the problem of mental health condition detection as multi-task learning framework
- Using Active Learning to tackle data paucity
- **Message Passing and Explainability:** Using Group Lasso to act as a feature explainer and graph-based Message Passing mechanism to model the correlation between tasks

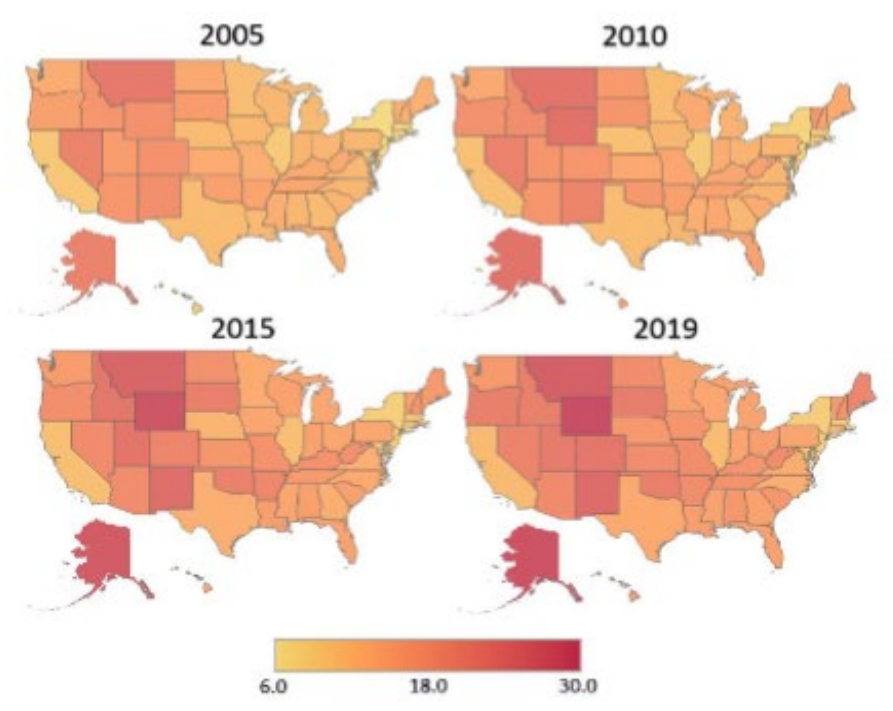
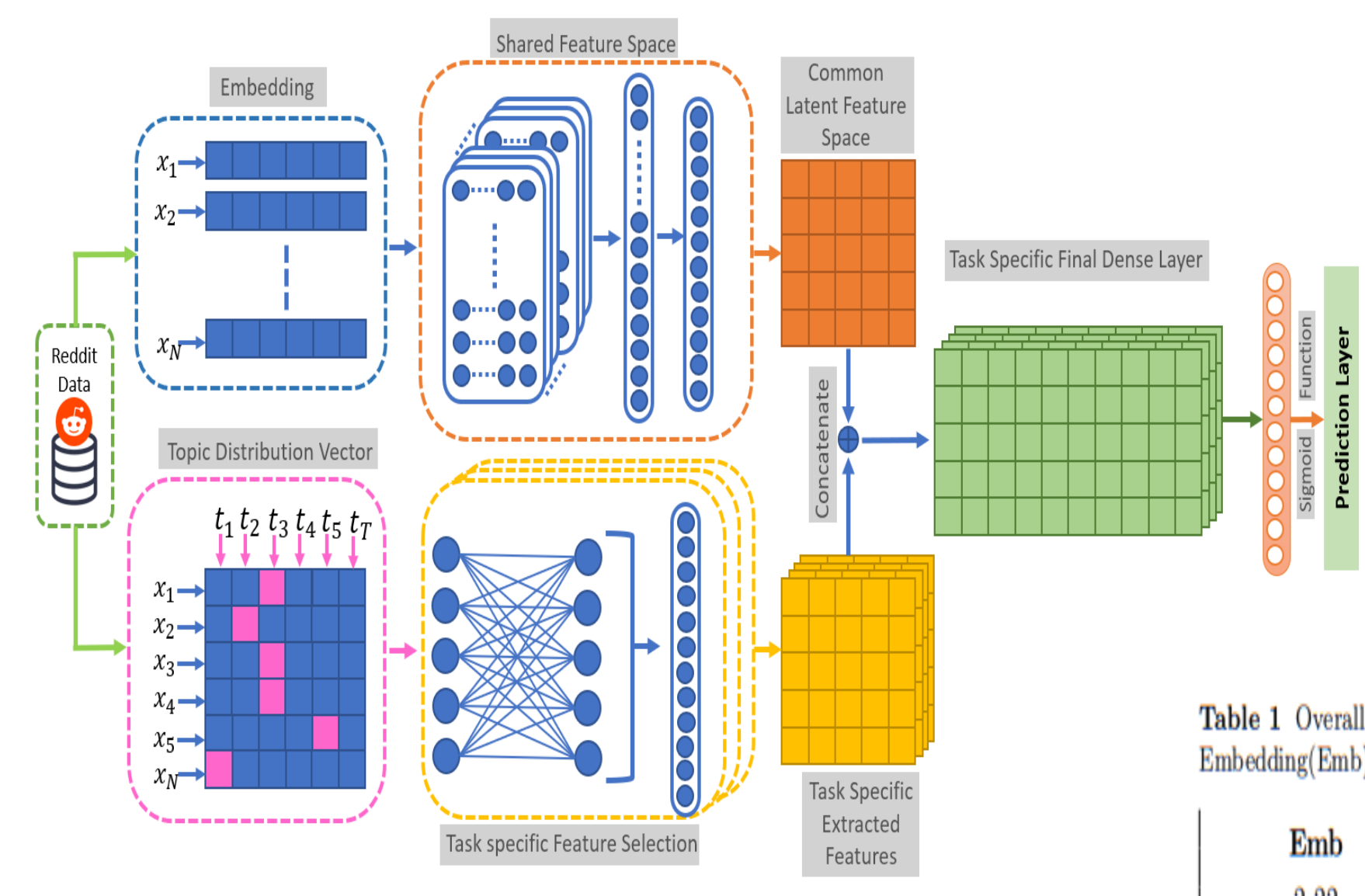


Fig. 1 A geographical heatmap of suicide mortality in different states of USA taken from CDC and social media and news media chatter about mental health crisis

Motivation

- Lack of interpretability in human centered AI
- Lack of high quality annotated data.
- Intersectionality and correlation between discussions about mental-health conditions.
- Using different topic modeling and lexical dictionaries as features

Model Architecture



$$E_{loss} = \sum_{j=1}^{n_{task}} \alpha_j (E_{BCELoss}(\hat{Y}, Y) + \beta L_j)$$

Performance of fine-grained mental health categories

Category	Precision	Recall	F1-Score
Depression	0.913	0.874	0.893
Anxiety	0.872	0.899	0.885
Other	0.860	0.873	0.866

Results compared with baseline methods

Table 1 Overall performance of baseline methods in comparison to our method on 5,000 Reddit submissions for Depression, Anxiety and Rest. Embedding(Emb), Percision (P), Recall (R), and micro-F1 (F1)

Emb	Logistic			KNN			SVM			Random Forest			MLP			GMMTL			
	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	P	R	F1	
2-22																			
TF-IDF	0.732	0.748	0.739	0.708	0.721	0.714	0.781	0.768	0.774	0.749	0.724	0.736	0.794	0.805	0.794	-	-	-	
BERT	0.761	0.752	0.756	0.713	0.738	0.720	0.819	0.801	0.809	0.742	0.726	0.733	0.817	0.841	0.828	-	-	-	
LDA	0.749	0.738	0.743	0.762	0.745	0.753	0.827	0.807	0.816	0.761	0.738	0.749	0.819	0.833	0.825	-	-	-	
BERTopic	0.750	0.739	0.744	0.761	0.740	0.750	0.826	0.815	0.820	0.771	0.752	0.761	0.847	0.826	0.836	-	-	-	
LDA+BERT	0.769	0.751	0.759	0.752	0.763	0.757	0.851	0.839	0.845	0.758	0.773	0.765	0.875	0.861	0.868	0.872	0.889	0.880	
BERTopic+BERT	0.785	0.771	0.778	0.745	0.728	0.736	0.879	0.863	0.869	0.779	0.765	0.772	0.873	0.859	0.866	0.889	0.882	0.885	
EMPATH+BERT	0.744	0.757	0.750	0.732	0.725	0.728	0.811	0.823	0.817	0.759	0.741	0.749	0.825	0.811	0.866	0.851	0.837	0.843	

Case Study-Depression

Most Important Topics for Depression Detection

Topic id	Top Phrases/Words
4	"depression" "feel" "depressed" "feeling" "want" "get" "life" "really" "even"
23	"thoughts" "mind" "things" "intrusive"
9	"want" "to die" "dead" "life" "want dead" "life"
24	"help" "suicidal" "suicidal hotline" "hotline" "need" "want" "talk" "feel"
10	"therapy" "therapist" "appointment" "need" "help" "cbt" "know" "get"

Case Study-Anxiety

Most Important Topics for Anxiety Detection

Topic id	Top Phrases/Words
8	"feel" "know" "anxiety" "depression" "want" "get" "life" "really" "time"
3	"side effect" "zoloft" "lexapro" "anyone" "meds" "take"
23	"thoughts" "mind" "things" "intrusive"
14	"work" "job" "home" "go" "day" "covid"
6	"anxious" "feeling" "calm" "often" "lot" "also" "always" "worrying"