

ESDM: Early Sensing Depression Model in Social Media Streams LREC-COLING 2024

Bichen Wang, Yuzhe Zi, Pengfei Deng, Yanyan Zhao*, Bing Qin Harbin Institute of Technology, China

Introduction

- Early Depression Detection Task
 - Early Depression Detection(ERD) is centered around the early identification of depressive users by our model.
- Key Problem
- Not every step taken by the user is labeled, which results in a mismatch in model requirements between training and testing phases.



• We must strike a balance between accuracy and time. Rapid detection may not always be feasible due to these constraints.

Motivation

- Previous studies train TRD models using full posting histories, which is unsuitable for ERD and previous models lack optimization for both earliness and accuracy.
- Inconsistencies between training and testing phases make the direct application of TRD models to ERD challenging.

Method

• CPI and DCM

•CPI Module's Role: The CPI processes the current partial sequence to generate a prediction denoted as y^t .

•DCM Module's Role: The DCM module is to decide whether the model should accept the decision based on the prediction y^t . Once DCM accept y^t the model will stop and label the user.

• Joint Training CPI and DCM

• Guide model training through early detection loss.

- Early Detection Loss
 - L_{pred} : Serving as the foundation for the CPI module's ability to classify a user with depression.
 - • L_{dec} : Enhancing the decision-making capabilities of DCM and the partial information classification abilities of CPI.
 - • L_{dealy} : Encouraging DCM to make decisions as early as possible.

Experiments

• Main Results

Model	$F1_{erd}(\uparrow)$	$erde_{50}(\downarrow)$	$erde_5(\downarrow)$	$F1_{trd}(\uparrow)$
LR	0.405	0.084	0.137	0.602
Feature-Rich	0.358	0.084	0.131	0.630

Temporal Decision Dynamics







• Attention Scores for ESDM and BiLSTM-Att

ESDM attention scores	

BiLSTM+Att	0.562	0.096	0.124	0.629
Risk window	0.606	0.097	0.130	0.629
SS3	0.497	0.086	0.133	0.546
EARLIST	0.273	0.148	0.164	0.175
HAN-Psych	0.603	0.081	0.107	0.703
ESDM	0.662	0.077	0.109	0.712

Ablation Study

Model	$F1_{erd}(\uparrow)$	$erde50(\downarrow)$	$erde5(\downarrow)$
ESDM	0.662	0.077	0.109
$-L_{pred}$	0.571	0.092	0.117
$-L_{dec}$	0.578	0.098	0.121
$-L_{delay}$	0.653	0.108	0.126

