Privacy Preserving Detection of Depression from Speech Data

By: Tracy Miles

UROC Mentored by Syed Hafiz

PROBLEM DEFINITION

Depression is a major health concern across the world. In the United States alone depression is the cause of over two-thirds of the 30,000 reported suicides each year.

It also accounts for 10% of all disability due to physical and mental health problems globally. Depression can be life threatening. It is treatable but only if the person that needs help seeks it. This is often not the case. Therefore using speech data to detect depression can be life saving.

- Speech processing technology such as machine learning models could aid mental health assessments
- Speech patterns have been used to effectively diagnose psychiatric disorders
- Machine learning system + Patient speech data from phone call = Mental health evaluation and analysis

MOTIVATION

The current system does not adequately protect the privacy of the patient data. This is because the current system stores the data in plain text instead of encrypting it. This problem needs to be evaluated. The current system evaluates mental health without proper privacy techniques. The patient information and the mental health evaluation both contain sensitive data.

OBJECTIVE

The goal for this research topic is to evaluate the performance of neural network machine learning models on real life application (Detection of Depression from Speech Data) to protect the privacy of internet users. In this issue we use homomorphic encryption to mitigate the privacy leakages.

MATERIALS AND METHODS

Microsoft SEAL

This is used to get better privacy guarantees when using this system. This was used to homomorphically encrypt and secure the symptom detection feature.

Machine Learning Approach: Multiclass Adversarial Discriminative Domain Generalization (MADDoG)

Three phases of approach below:

Phase 1: Feature extraction – Local

- Kaldi speech recognition toolkit

Phase 2: Emotional Classifier – Cloud

- 5 dense NNs (pretrained) (1 per emotion):
  - Guilt, Worry, Hopelessness, Anger at others, Anger at self
- Outputs an emotion score for each emotion

Phase 3: Evaluation standard deviation of Emotion Scores - Local or Cloud

- Calculate standard deviation of emotion scores over time
- For these emotions, suicidal individuals show less variability than non-suicidal.

THE DENSE NN for EMOTION CLASSIFICATION

- 4 fully connected hidden layers
  - 1,024,512,256,256 neurons
- Activation Functions
  - RReLU for hidden layers
  - Sigmoid for output layer
- Trained on Ecological Measurement of Affect, Speech, and Suicide (EMASS) Dataset

RESULTS & CONCLUSION

Microsoft SEAL has allowed the speech data to be encrypted and this allows the patients to be safe. It also allows the speech data to be secured so the symptom detection can evaluate the mental health and give good resources to make sure the patient can get help.

REFERENCE

“Private AI Bootcamp Competition: Team 4.” *Microsoft Research*, 9 Jan. 2020, www.microsoft.com/en-us/research/video/private-ai-bootcamp-competition-teams-4/.

ACKNOWLEDGEMENT

This study was organized by the Luddy School of Informatics, Computing, and Engineering at Indiana University, Bloomington. The program is called the Undergraduate Research Opportunities in Computing.