LREC-COLING 2024

LocalTweets to LocalHealth

A mental health surveillance framework based on Twitter data

Vijeta Deshpande[1], Minhwa Lee[2], Zonghai Yao[3], Zihao Zhang[3], Jason Brian Gibbons[4], Hong Yu[1,3,5]

1: University of Massachusetts Lowell, 2: University of Minnesota, 3: University of Massachusetts Amherst

4: University of Colorado Anschutz Medical Campus,

5: University of Massachusetts, Chan Medical School

LocalTweets

NEED OF SUPPLEMENTARY SYSTEMS

- Health surveillance systems
  - Collect data on population health status.
  - Report estimates of population health.
- Disadvantages of the conventional surveillance systems
  - Reporting of the health outcome estimates takes a long time.
  - Skilled labor is needed to design surveys and collect data.
- High engagement social-media platforms
  - Data sources that can facilitate quicker and cheaper data collection avenues.

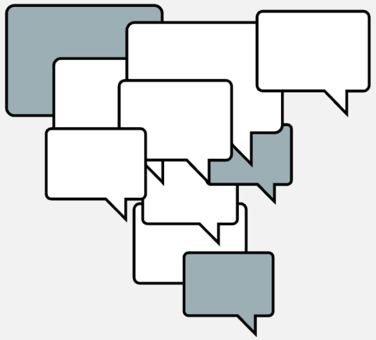
PREVIOUS RESEARCH

- Tweet-level surveillance system
  - Widely published approach.
  - Conceptually limited in representing population.
- Population-level surveillance system
  - Larger population (State- or County-level systems).
  - Limited scrutiny with only the correlation analysis.
  - Text information is not sufficiently leveraged.

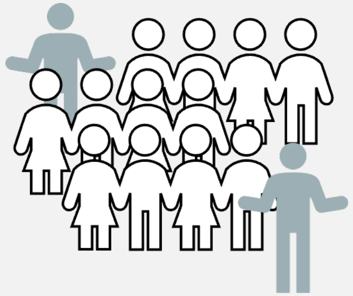
OUR CONTRIBUTIONS

- Focus on (much) smaller geographical units.
  - Population ~ 600 to 3,000
- Efficiently leverage the text information and LLMs.
  - From proprietary LLMs to efficient general and domain adapted models.
- A direct outcome prediction model.
  - Assessment beyond correlation.

BRIEF OVERVIEW



$$\longrightarrow \quad f(\cdot) \quad \longrightarrow$$



Tweets collected from a neighborhood

Parametric model trained on neighborhood-level data

CDC-reported mental health outcome

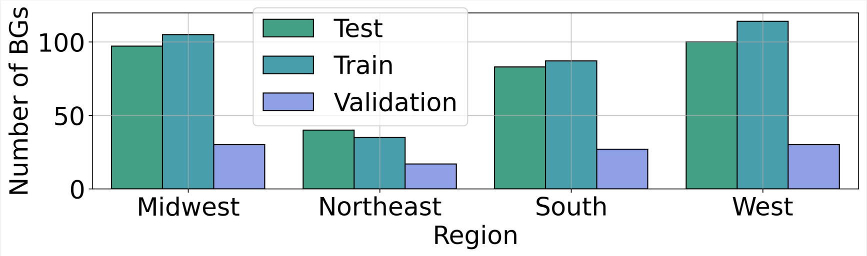
DATA

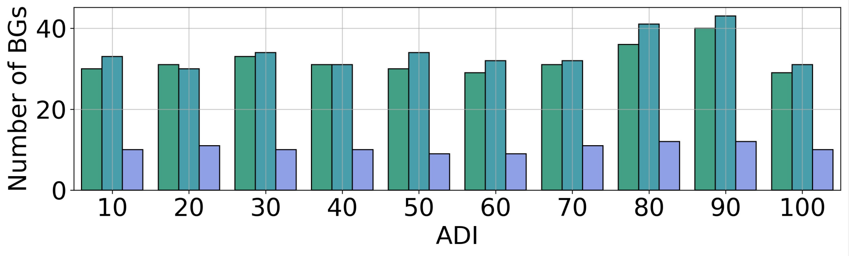
- Sampling neighborhoods (Census Block-Group: BG)
  - Sampling of 1,000 BGs from ~250,000 in contiguous USA.
  - Balancing representation along regional and socio-economic axes.
    - Regional: North-east, South, Midwest, West.
    - Socio-economic: Area Deprivation Index (ADI).
- Collecting tweets
  - Keyword-based collection (mental health and food insecurity) and general tweets.
- Collecting CDC-reported outcome
  - Definition of variable[1]: Percentage of respondents aged ≥18 years who report 14 or more days during the past 30 days during which their mental health was not good.

1: Definition taken from CDC website: https://www.cdc.gov/places/measure-definitions/health-status/index.html#mental-health

DATA
(LocalTweets)

- Data instance
  - X: Set of tweets (MH, FI or general)
  - Y: MH outcome
- Size of the data (>22 mil. Tweets)
  - Total 765 data instances
  - From 2015 to 2019
- Regional distribution
  - Less instances in the Northeast
  - Does not affect regional representation
- Socio-economic distribution

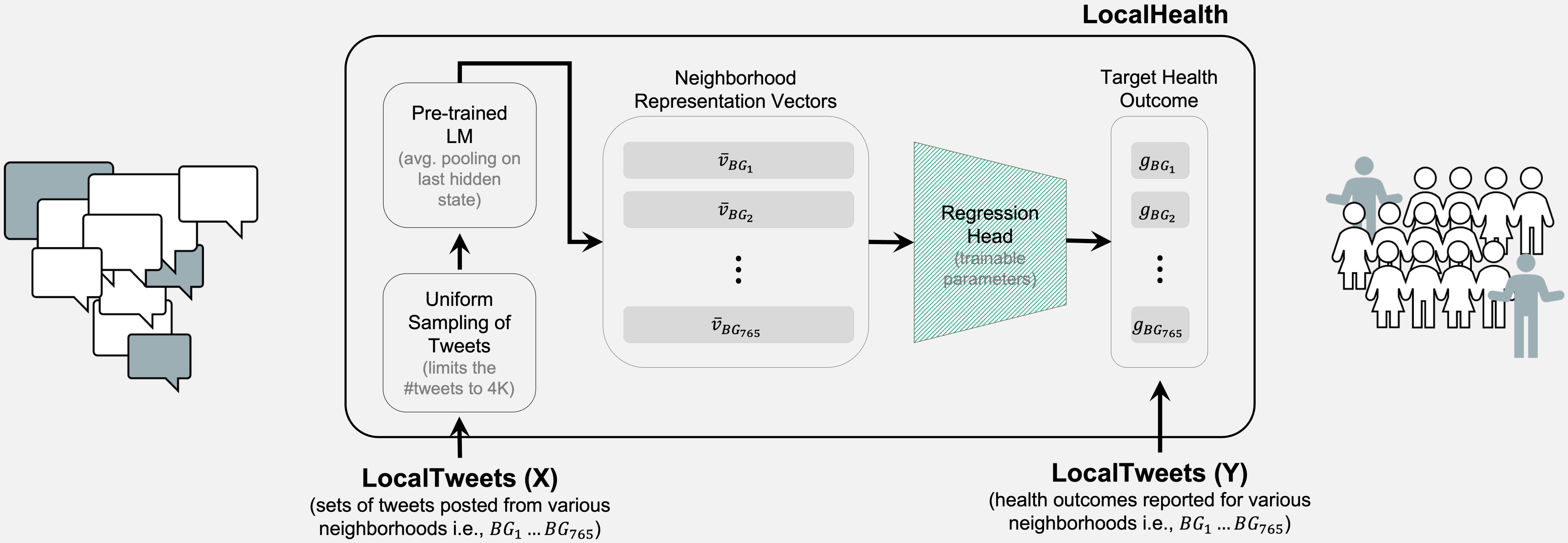




REGRESSION MODEL

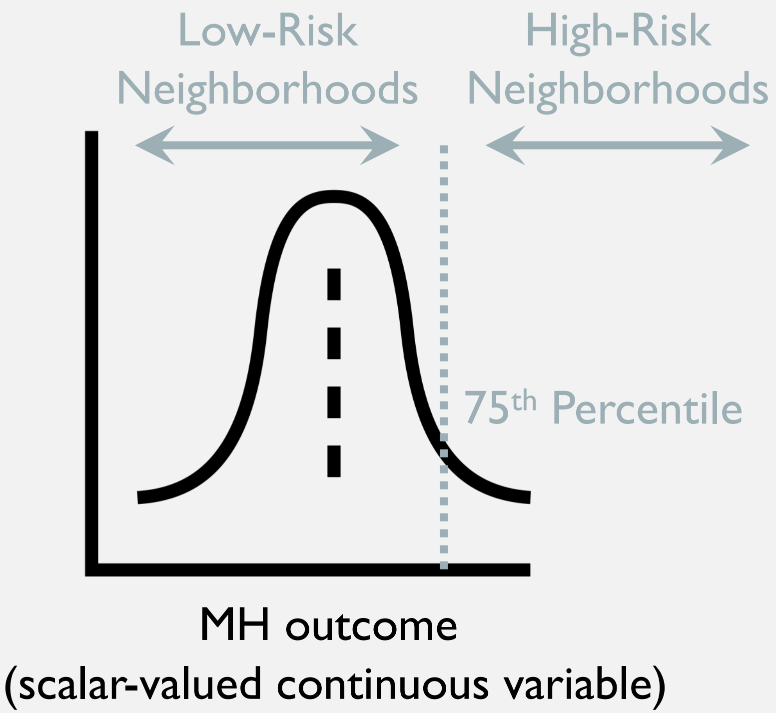
- **Step-1:** Sampling tweets
  - We limit the size of the set of tweets to 4,000 (uniform sampling).
- **Step-2:** Encoding tweets
  - Pre-trained language models: RoBERTa, Twitter-RoBERTa, GPT-3.5 etc. (frozen).
- **Step-3:** Neighborhood representation vector ($\bar{v}$)
  - Mean pooling across the tweet length and number of tweets.
- **Step-4:** Prediction of MH outcome ($\hat{g}$)
  - We use a combination of convolutional and feed-forward NN.

REGRESSION MODEL
(LocalHealth)

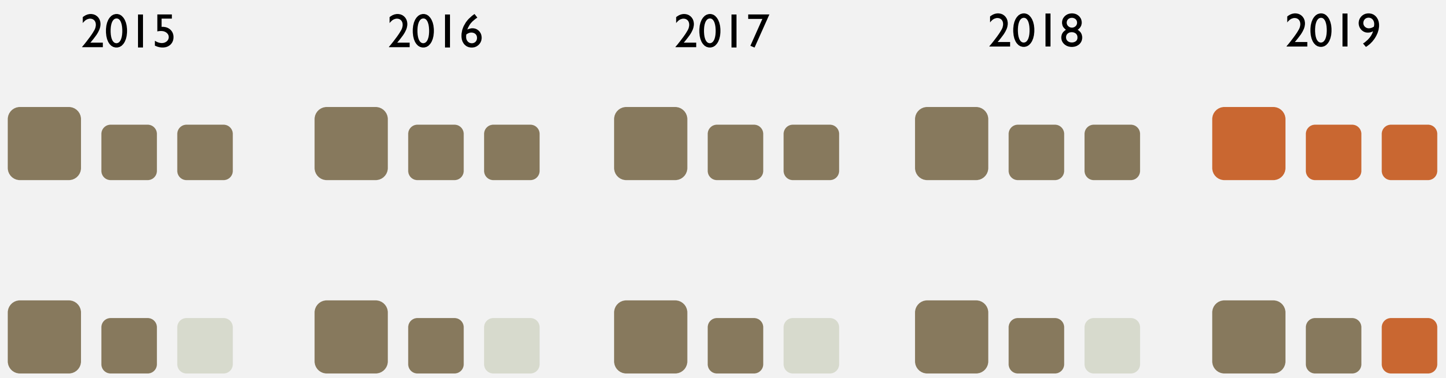


EVALUATION

- We train the model with MSE.
- However, we evaluate the model with categorical metrics.
  - We convert the continuous-valued MH outcome to binary variable, high-risk (1) and low-risk (0).
  - We use predicted and target binary variable to evaluate the model with F1-score and accuracy.
- All reported values are average of 10 random seed runs.



EXPERIMENTS



Forecasting
(Can we predict future MH outcomes?)

Spatial Extrapolation
(Can we predict unreported MH outcomes?)

Training BGs   Removed BGs   Evaluation BGs

EXPERIMENTS

- Forecasting set-up
  - Experiment-1: Which category of tweets is the most beneficial?
  - Experiment-2: Which text encoder achieves the best performance?
  - Experiment-3: What is the effect of data availability?

- Spatial extrapolation setup
  - Experiment-4: Can we predict MH outcomes for the unreported neighborhoods?

EXPERIMENT-1 RESULTS

(Which category of tweets is the most beneficial?)

- General category tweets generalize better.

Input information	F1-score	Acc. (%)
Majority baseline	0.4336	76.56
Text-based (LoR)	0.5224	52.75
Text-based (SVM)	0.5510	**76.73**
ADI-only (LR)	0.6406	72.81
Text-based (LocalHealth) with ADI		
MH only	0.7089	74.52
FI only	0.7117	75.36
MH and FI	0.7085	74.33
General only	**0.7236**	76.48

EXPERIMENT-2 RESULTS

(Which text encoder achieves the best performance?)

- RoBERTa-base achieves competitive performance.
- Minor benefits from using domain adapted models such as Twitter-RoBERTa and PHS-BERT.
- GPT-3.5 achieves the best performance.

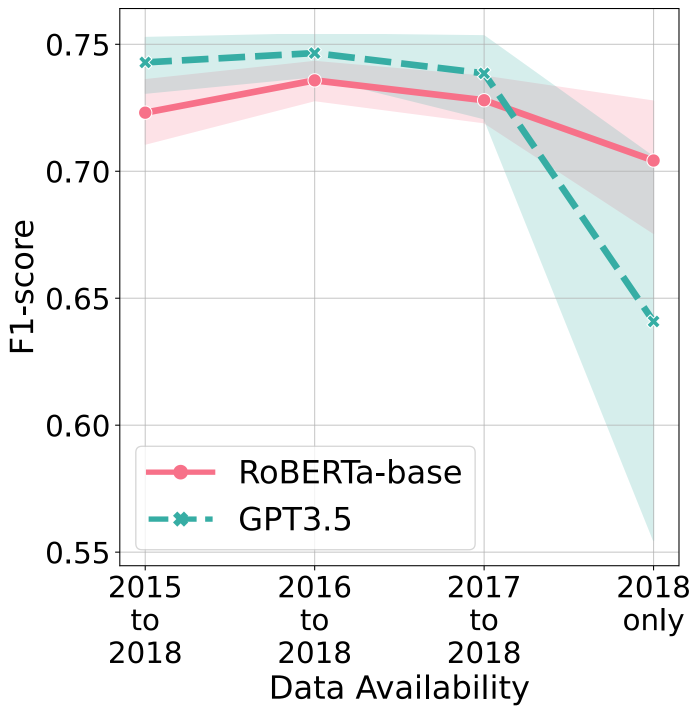
Language Model	Train Par.	F1-score	Acc. (%)
Majority baseline	–	0.4336	76.56
GPT3.5 (0-shot)	0	0.4675	76.21
ADI only	2	0.6406	72.81
RoBERTa-base	210	0.7236	76.48
RoBERTa-large	274	0.7228	76.04
Twitter-RoBERTa-base	210	0.7245	76.44
PHS-BERT	274	0.7301	76.97
GPT3.5	402	**0.7429**	**79.78**

EXPERIMENT-3 RESULTS

(What is the effect of data availability?)

- GPT-3.5 performance reduces more aggressively than RoBERTa-base, with reduction in available data.

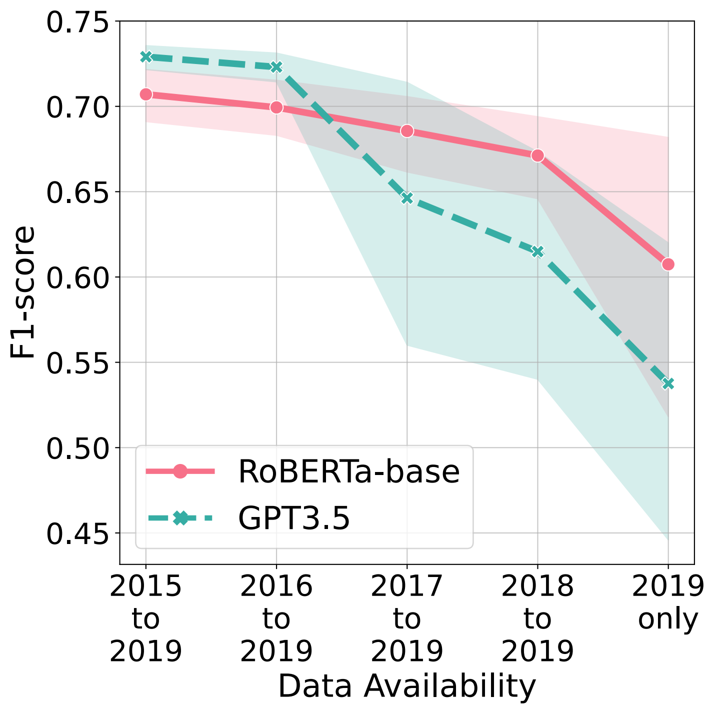
- Slight reduction in the performance after including 2015 data.
  - Possibly due to older data being less relevant for prediction of 2019 MH outcomes.



EXPERIMENT-4 RESULTS

(Can we predict MH outcomes for the unreported neighborhoods?)

- We can predict MH outcomes for the unreported BGs with >0.7 F1-score.
- GPT-3.5 performance reduces more aggressively than RoBERTa-base, with reduction in available data.
- More data monotonically improves performance of both models.



CONCLUSION

- LocalTweets: First neighborhood level dataset for MH surveillance.
- General tweets generalize better than keyword-based tweets.
- Presented approach can be leveraged to predict MH outcomes for neighborhoods not covered in CDC repository.
- Presented data and methods can be utilized directly to identify neighborhoods that can benefit from the establishment of community health programs

PRIVACY AND ETHICAL CONSIDERATIONS

- Accessing tweets from public profiles.
- Preservation of privacy of the individuals posting tweets.
- Bias analysis of LocalTweets.
- Gated sharing of LocalTweets and LocalHealth.

THANK YOU!