Data Infrastructures in Context: Harnessing Social Media Data to Improve Mental Health Outcomes

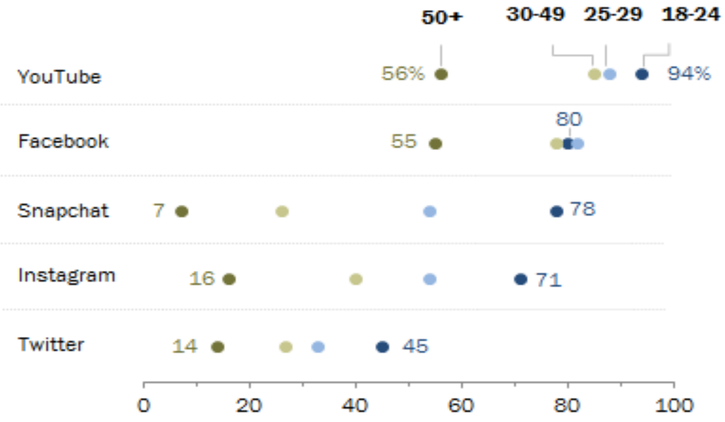
Munmun De Choudhury
munmund@gatech.edu
School of Interactive Computing | Georgia Tech

Georgia Tech College of Computing

A Revolution

Social platforms like Snapchat and Instagram are especially popular among those ages 18 to 24

% of U.S. adults in each age group who say they use ...

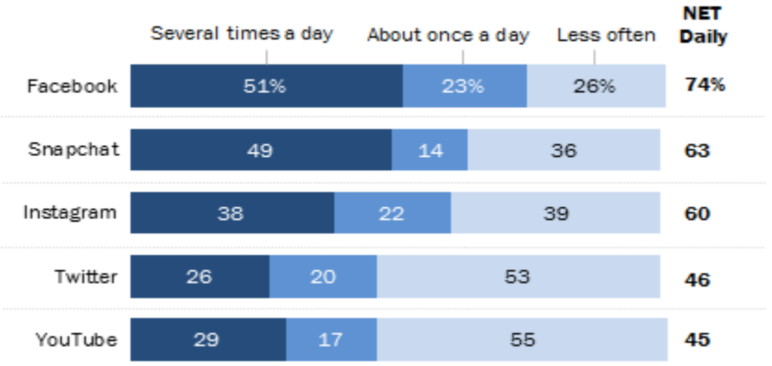


Source: Survey conducted Jan. 3-10, 2018.
"Social Media Use in 2018"

PEW RESEARCH CENTER

A majority of Facebook, Snapchat and Instagram users visit these platforms on a daily basis

Among U.S. adults who say they use __, the % who use each site ...



Note: Respondents who did not give answer are not shown. "Less often" category includes users who visit these sites a few times a week, every few weeks or less often.
Source: Survey conducted Jan. 3-10, 2018.
"Social Media Use in 2018"

PEW RESEARCH CENTER


Le-shaker.fr


Predicting Depression via Social Media
Munmun De Choudhury Michael Gamon Scott Counts Eric Horvitz
Microsoft Research, Redmond WA 98052
Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media
Munmun De Choudhury Emre Kiciman Mark Dredze
Glen Coppersmith Mrinal Kumar
Facebook language predicts depression in medical records
Characterizing and Predicting Postpartum Depression from Shared Facebook Data
Munmun De Choudhury Scott Counts Eric Horvitz Aaron Hoff
JOURNAL OF MEDICAL INTERNET RESEARCH
Original Paper
A Collaborative Approach to Identifying Social Media Markers of Schizophrenia by Employing Machine Learning and Clinical Appraisals
Natural Language Processing of Social Media as Screening for Suicide Risk
Glen Coppersmith, Ryan Leary, Patrick Crutchley and Alex Fine

Data Infrastructures in Context

- Augment existing signals
- Incorporate real-time-ness
- Target underserved populations
- Address issues of construct validity and dataset shift
- Balance theory-driven and data-driven approaches
- Validate algorithms in the real-world setting where they will be used
- Talk to privacy and ethics challenges

Methodological Gaps in Predicting Mental Health States from Social Media: Triangulating Diagnostic Signals

Sindhu Kiranmai Ernala
Georgia Tech
sernala3@gatech.edu

Michael L. Birnbaum
Zucker Hillside Hospital, Psychiatry Research
Mbirnbaum@northwell.edu

Kristin A. Candan
Zucker Hillside Hospital, Psychiatry Research
kcandan@northwell.edu

Asra F. Rizvi
Zucker Hillside Hospital, Psychiatry Research
ARizvi3@northwell.edu

William A. Sterling
Zucker Hillside Hospital, Psychiatry Research
wsterling2@northwell.edu

John M. Kane
Zucker Hillside Hospital, Psychiatry Research
JKane2@northwell.edu

Munmun De Choudhury
Georgia Tech
munmund@gatech.edu

ABSTRACT

A growing body of research is combining social media data with machine learning to predict mental health states of individuals. An implication of this research lies in informing evidence-based diagnosis and treatment. However, obtaining clinically valid diagnostic information from sensitive patient populations is challenging. Consequently, researchers have operationalized characteristic online behaviors as "proxy diagnostic signals" for building these models. This paper posits a challenge in using these diagnostic signals, purported to support clinical decision-making. Focusing on three commonly used proxy diagnostic signals derived from social media, we find that predictive models built on these data, although offer strong internal validity, suffer from poor external validity when tested on mental health patients. A deeper dive reveals issues of population and sampling bias, as well as of uncertainty in construct validity inherent in these proxies. We discuss the methodological and clinical implications of these gaps and provide remedial guidelines for future research.

CCS CONCEPTS

• **Computing methodologies → Supervised learning by classification**; **Supervised learning by classification**; • **Human-centered computing → Social media**.

KEYWORDS

mental health; social media; machine learning; validity theory; construct validity; population bias; sampling bias

ACM Reference Format:
Sindhu Kiranmai Ernala, Michael L. Birnbaum, Kristin A. Candan, Asra F. Rizvi, William A. Sterling, John M. Kane, and Munmun De Choudhury. 2019. Methodological Gaps in Predicting Mental Health States from Social Media: Triangulating Diagnostic Signals. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3290605.3300364

1 INTRODUCTION

With rising volumes of data and pervasive use, social media has been widely adopted as a lens to provide insights into behaviors [52], mood [42], psychological traits and states [5, 53], and social interactions of individuals [56]. For mental health, a growing body of work, including that in the human computer interaction (HCI) field, is leveraging naturalistic, unobtrusive data from social media to predict mental health states of individuals [21, 25, 28, 29, 31, 34]. Parallel to HCI, in an emergent field called "digital psychiatry" [100], clinicians are exploring the efficacy of diagnostic predictions from online data for early diagnosis, evidence-based treatment, and deploying timely patient-provider interventions [40, 48].

In this line of research, on the methodological front, supervised machine learning techniques have gained prominence, providing promising predictive outcomes of mental health states [66]. The success of these techniques, however, hinges on access to ample and high-quality gold standard labels for model training. In mental health, gold standard labels often comprise *diagnostic signals of people's clinical mental health states*, for instance, whether *an individual might be suffering from a specific mental illness*, or *at the cusp of experiencing an adverse episode* like a relapse or suicidal thoughts.

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05...$15.00
https://doi.org/10.1145/3290605.3300364

THRIVE: Designing and Deploying Clinical Tools
Powered by Patients' Social Media Data

REALTIME
ENSEMBLE
DATA for
UNDERSTANDING
SUI**C**IDE
EPIDEMIOLOGY

THRIVE: Designing and Deploying Clinical Tools Powered by Patients' Social Media Data

Can we predict relapse with social media?

- Schizophrenia affects about 1% of the world's population
- Up to 80% schizophrenia patients relapse in 5 years

- *Challenge:* Early identification of indicators of relapse for treatment and intervention

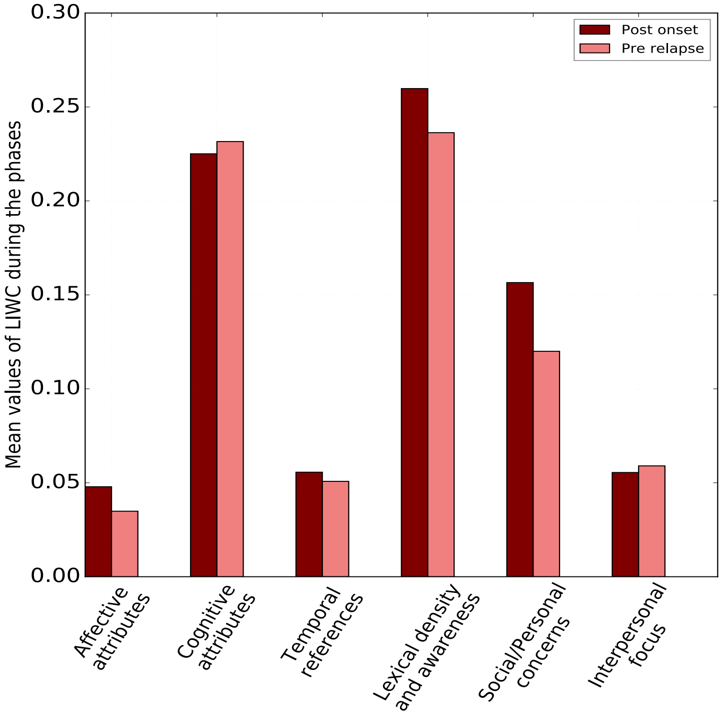
Birnbaum, M. L.*, Ernala, S. K.*, Rizvi, A., Arenare, E., Van Meter, A., De Choudhury, M.** and Kane, J. M.** (2019). *Detecting Relapse in Youth with Psychotic Disorders Utilizing Patient-Generated and Patient-Contributed Digital Data from Facebook*. Nature Partner Journals - Schizophrenia. npj Schizophrenia. * Co-first authors; ** Co-supervising authors

Research supported through a cooperative agreement between Georgia Tech and the Feinstein Institute, the Zucker Early Treatment Program, and NIMH grant R01MH117172 (PI De Choudhury)

Northwell EP Patient Data Collection

- 110 early psychosis (EP) patients (average age 24 years; 64% female)
  - 51 experienced a relapse
  - Relapse hospitalizations = 124 (Mean = 2.6, Median = 2)
- Full archives of Facebook data
  - 82% of those eligible agreed to share their Facebook data
  - Among those who consented, 100% agreed to share all of their data
  - 96,393 self-posts and self-comments spanning a period of 76 months (52,815 from the 51 relapse patients)

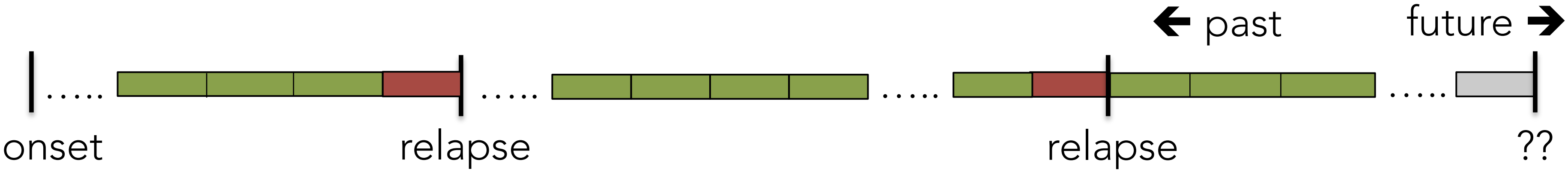
Establishing Feasibility



anxiety medication really struggling
mind racing
coping mechanisms voices really
police called
experience hallucinations
trouble remembering **suicide attempt**
hospitalization believe mental
seek professional
meds work need medication
people watching
visual hallucination voice inside
mentally unstable
new meds cameras **feel lonely**
feel scared anxious fighting

Distinctive n-grams that precede a relapse

Predicting Relapse

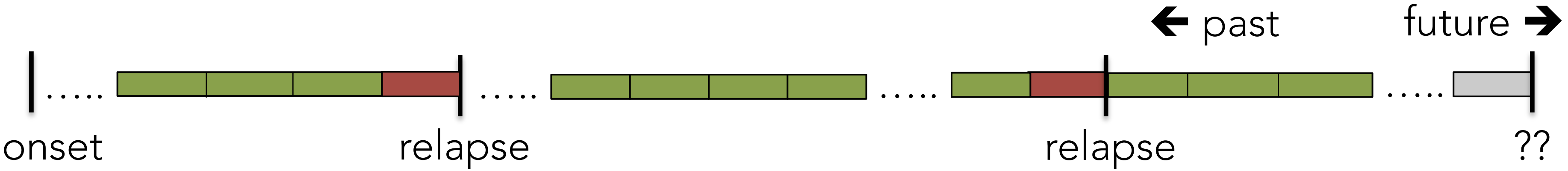


Conceptual framework for personalized relapse prediction

Temporal chunks of Facebook timeline, punctuated by hospitalizations

- Supervised learning approaches unsuitable
  - Relapse is a rare event; sparsity of positive examples
  - Clinical heterogeneity of patients
  - Conceptually no "true" negative examples – anybody can relapse at some point in the future

Predicting Relapse



Conceptual framework for personalized relapse prediction

- Relapse prediction as an anomaly detection problem
  - In a prospective setting, identify is aberrations in behavior, deviating from the baseline
  - Periods of health (1, 2, 3, months) – baseline data
  - Periods of relapse (1 month before hospitalization) – anomalies

Predicting Relapse

All data (healthy:719, relapse:49)

Training + Validation
(90% healthy)

Unseen test data
(10% healthy, 100% relapses)

- Testing on healthy + relapse periods
  - TP, FN, FP, TN = (27, 45, 10, 39)
  - Specificity (relapses predicted as relapse) = TN/TN+FP = 39/49 = 0.79
  - Sensitivity = TP/TP+FN = 27/72 = 0.37

Error analysis: evaluation via clinical chart review

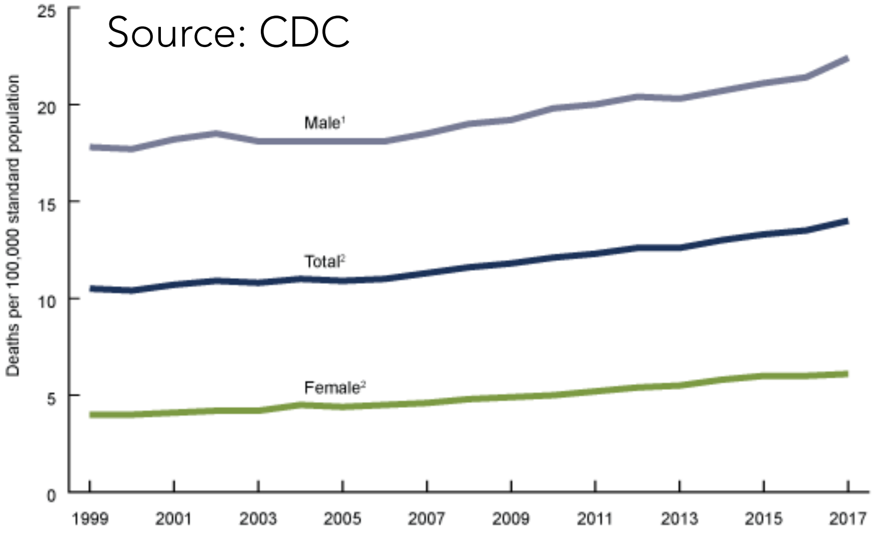
- Analysis of false negatives (periods of relative health wrongly predicted as a relapse)
- For 20 out of the 45 false-negative time periods (44%), data was available from the patient's medical record.
- In 18 of these 20 instances, the presence of psychotic symptoms during periods defined as relative health was documented
  - 6 of these participants had known non-adherence to medication during this time which can contribute to symptomatic exacerbations

Takeaways

- This work allows us to go beyond utilizing social media activity to identify population-based, or group-level characteristics, associated with mental health status—nearly exclusively the only approach employed in prior research.
- With our machine-learning approach, we have demonstrated that personalized methods to longitudinally forecast the likelihood of imminent adverse mental health outcomes, like a relapse, is feasible.

Forecasting Nationwide Suicide Rates

- From 1999 through 2017, the age-adjusted suicide rate increased 33% from 10.5 to 14.0 per 100,000.



REALTIME
ENSEMBLE
DATA for
UNDERSTANDING
SUICIDE
EPIDEMIOLOGY

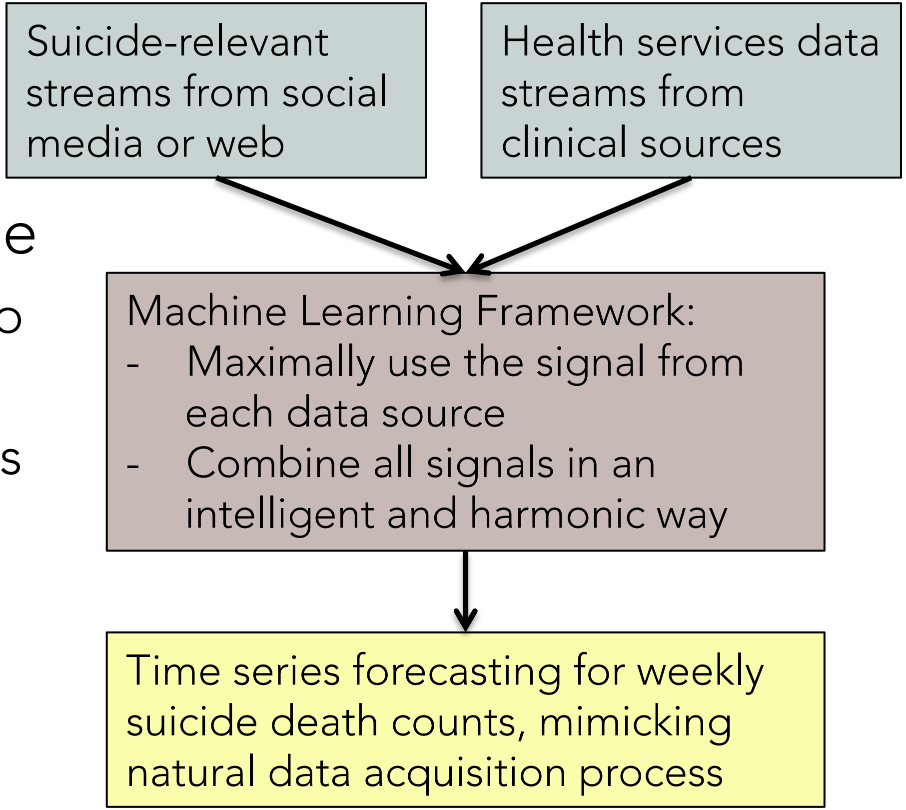
Ongoing work, in collaboration with and funded by the Division of Violence Prevention, CDC through a contractual agreement with Georgia Tech

Forecasting Nationwide Suicide Rates

- In spite of the urgency of this public health problem, there exists a lack of real-time information on suicide fatality trends to guide prevention efforts.
- The NVSS collects information from death certificates that are submitted by the more than 2,000 medical examiner and coroner offices in the U.S.
- Unfortunately, national statistics on suicide rates are delayed by 1-2 years, depending on the time point at which data are queried.

Combining Real-Time Datasets

- Goal: Building a machine learning framework to predict suicide death counts in real-time
  - Time series forecasting problem to predict weekly number of suicide
  - Using multiple time series datasets collected and pre-processed from both social media and clinical sources



Data Collection

- Social media and web

Dataset	Method	Description
Google / Youtube	Keyword searching from Google Trend	Trend scores of 42 keywords
Twitter	Keyword searching by *GetOldTweets*	Number of users who upload at least one tweets retrieved by 38 selected keywords
Reddit	All posts in subreddits from Pushshift.io	Number of posts in the selected 53 subreddits

- Clinical/Health services data (provided by CDC)
  - Essence, Call, Poison

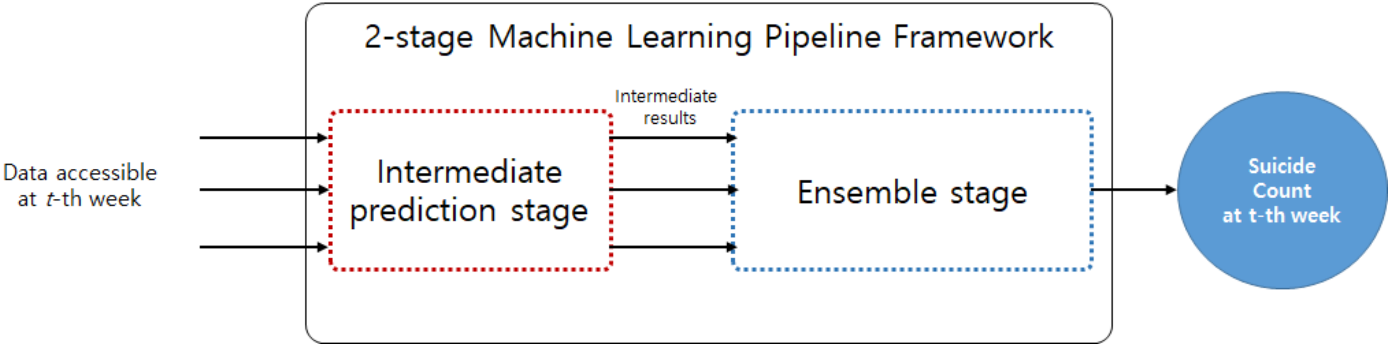
Data Preprocessing and Description

- All datasets encoded as input vectors of weekly granularity
  - For the social media and web datasets, we (1) compute weekly time series from the number of each keywords, then (2) net sum of the time series of all the keywords

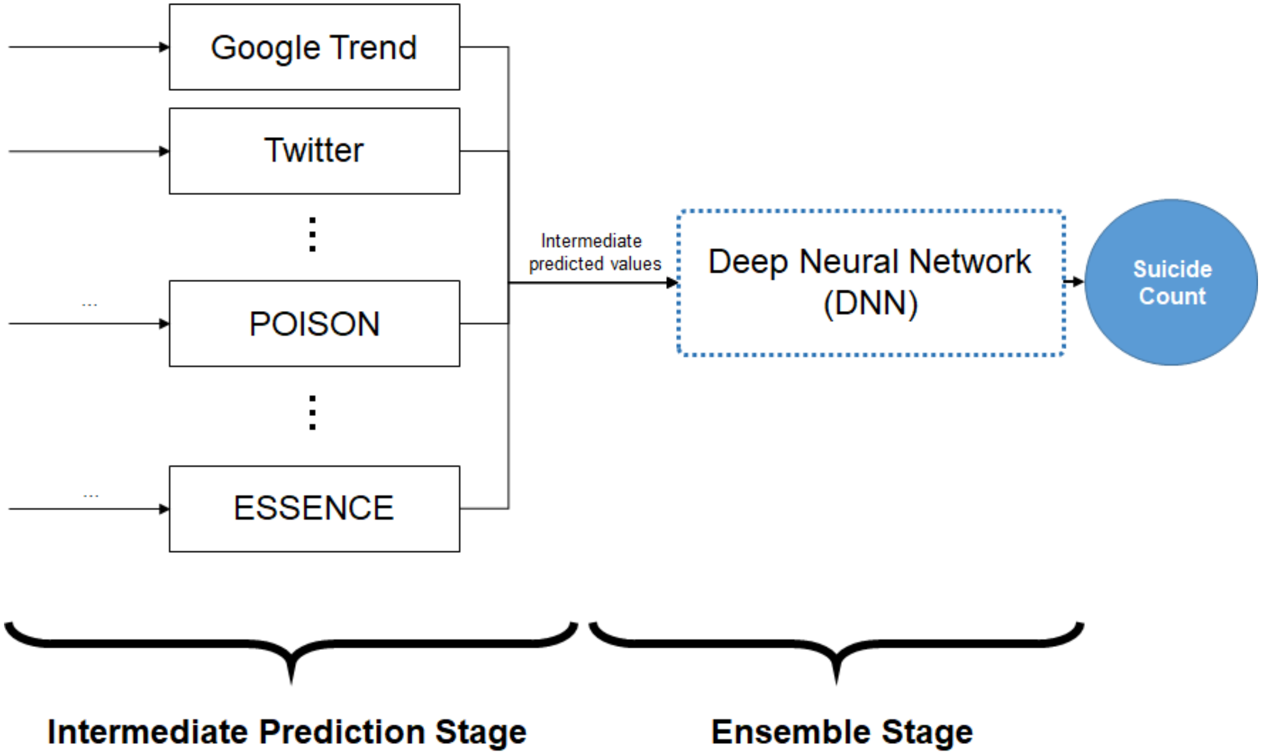
Type	Source	Period	Amount	Used Feature
Public health signals	POISON	2014-2017	-	Death counts over all poison control data
	CALL	2014-2017	-	Number of answered Lifeline calls
	ESSENCE	2015-2017	-	Normalized ESSENCE-REDUCE counts over total ED visits
Social media	Google (Health)	2014-2017	-	Trend scores (already normalized by Google Trends)
	Youtube (Mental health)	2014-2017	-	Trend scores (already normalized by Google Trends)
	Reddit	2014-2017	2,314,533 posts; 638,657 users	Normalized posts (by #posts in all subreddits)
	Twitter	2015-2017	9,327,472 posts; 5,565,341 users	Normalized users (by active #Twitter users)

Overall Architecture

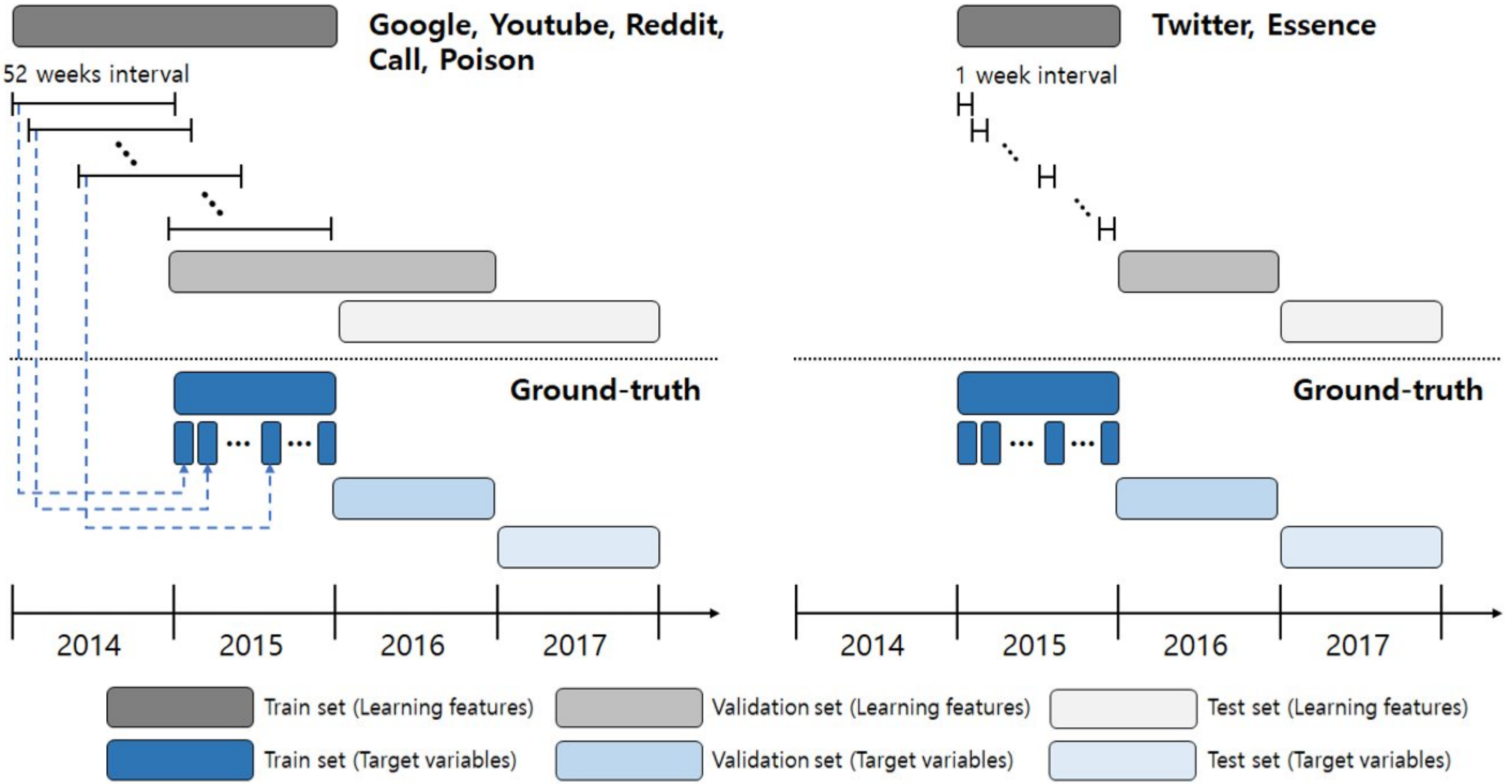
- 2-stage machine learning pipeline framework
- Intermediate prediction stage
  - One ML model for one data stream
  - Output the intermediate result (number of suicide) based on the given data stream
- Ensemble stage



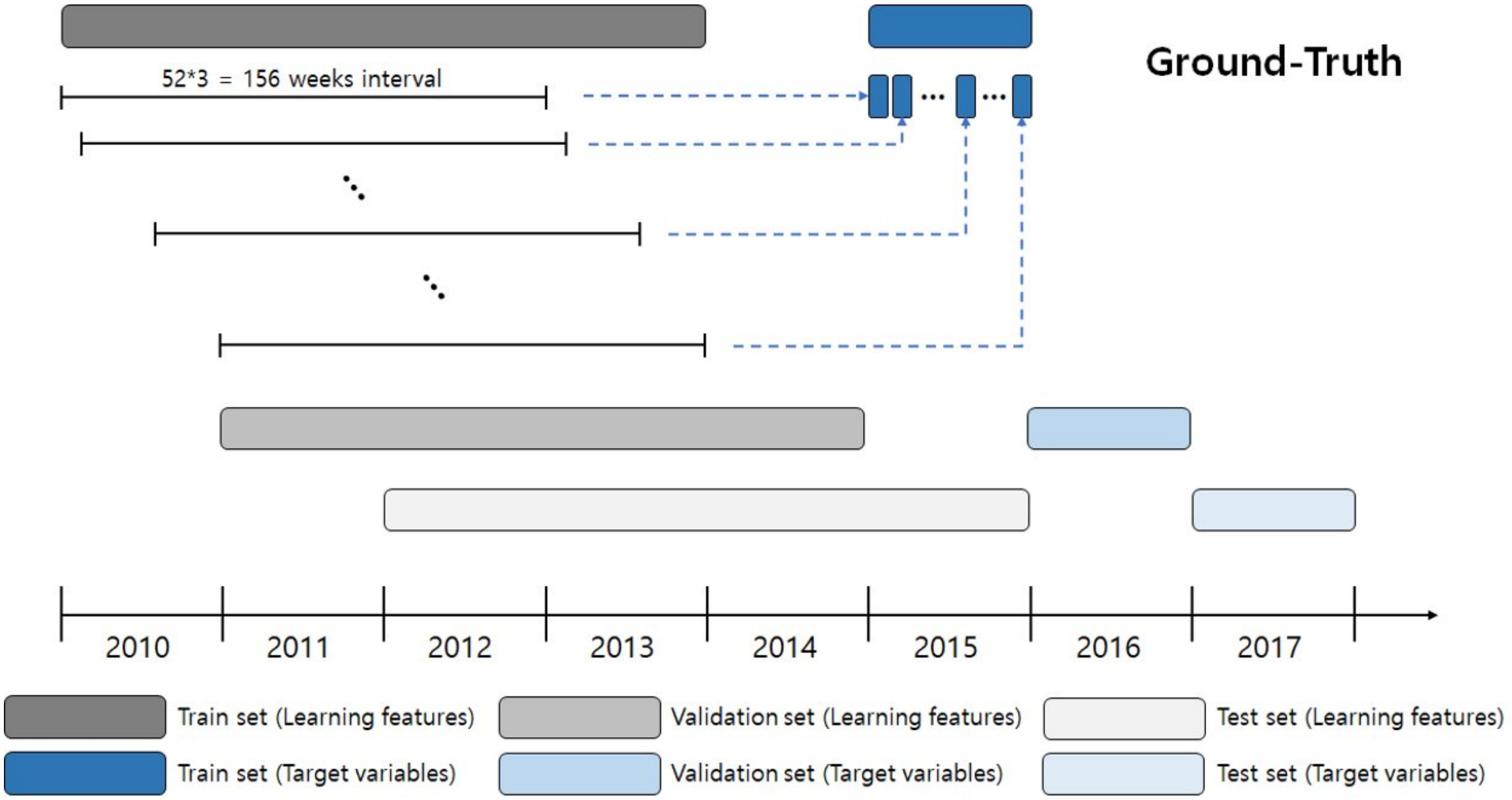
Overall Architecture



Training, Validating, and Test sets



Training, Validating, and Test sets



Prediction Results (Unit Models)

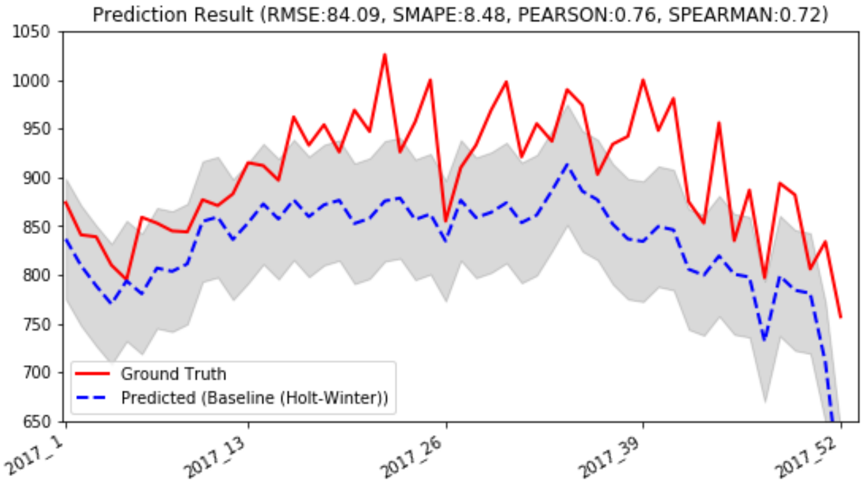
Ground truth: 14.47

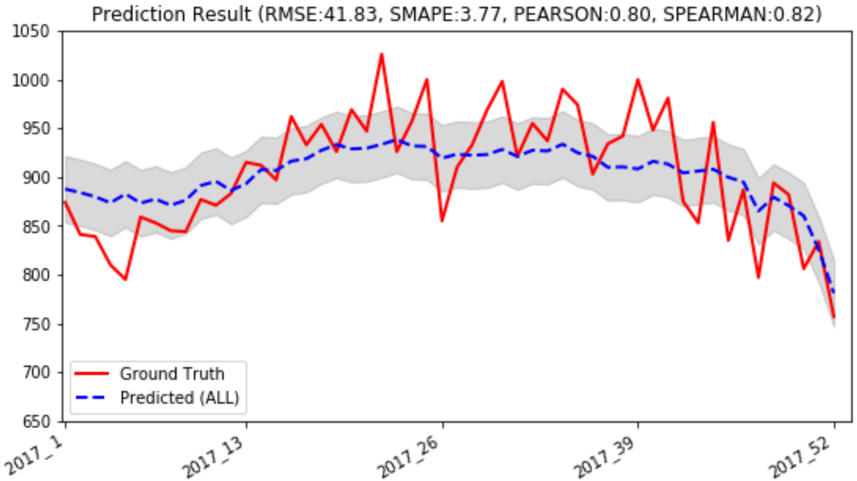
Category	Source	Model	Parameter set	RMSE	SMAPE	Pearson	Predicted Rate	Rate Error (%)
Simple Linear Regression		Linear Regression	-	98.542	10.099	0.704	13.09	9.54
Naïve ML Baseline		Support Vector Regressor	C: 10, epsilon: 0.1, gamma: 0.01, kernel: poly	75.274	7.222	0.750	13.59	6.08
Baseline		Holt-Winter	Seasonal period: 52 Trend: None, Damped: False, Seasonal: multiplicative	84.087	8.477	0.759	13.29	8.15
Clinical Sources	POISON	ElasticNet	alpha: 1.0, l1_ratio: 0.1	177.184	20.891	0.686	11.73	18.94
	CALL	ElasticNet	alpha: 1.0, l1_ratio: 0.1	55.934	4.802	0.496	14.54	0.48
	ESSENCE	Linear Regression	-	54.353	4.882	0.511	14.41	0.41
Social Media	Google (Health)	Random Forest	Number of estimators: 500, Min. Samples of split: 2, Min. Samples of leaf: 4	82.757	7.729	0.588	13.44	7.12
	Youtube (Mental health)	Support Vector Regressor	C: 10, epsilon: 0.1, gamma: 0.1, kernel: poly	87.505	8.160	0.467	13.49	6.77
	Reddit	Support Vector Regressor	C: 100, epsilon: 0.1, gamma: 0.01, kernel: Sigmoid	223.592	27.099	0.564	10.99	24.05
	Twitter	Support Vector Regressor	C: 100, epsilon: 0.1, gamma: 1, kernel: rbf	72.640	6.709	0.389	13.65	5.67

Prediction Results (Ensemble Model) Ground truth: 14.47

Category	Source	Model	Parameter set	RMSE	SMAPE	Pearson	Predicted Rate	Rate Error (%)
Simple Linear Regression		Linear Regression	-	98.542	10.099	0.704	13.09	9.54
Naïve ML Baseline		Support Vector Regressor	C: 10, epsilon: 0.1, gamma: 0.01, kernel: poly	75.274	7.222	0.750	13.59	6.08
Baseline		Holt-Winter	Seasonal period: 52 Trend: None, Damped: False, Seasonal: multiplicative	84.087	8.477	0.759	13.29	8.15
Clinical Source Only				114.818	12.428	0.768	12.76	11.82
Social Media Only				54.582	5.117	0.573	14.44	0.21
Baseline + Clinical Source				42.337	3.752	0.810	14.78	2.14
Baseline + Social Media				81.132	7.635	0.737	13.41	7.33
Clinical Source + Social Media				48.330	4.326	0.790	14.27	1.38
Baseline + Clinical Source + Social Media				41.832	3.766	0.801	14.43	0.28

Comparison with Baseline





Takeaways

- The first comprehensive study to predict suicide mortality in the US, harnessing diverse real-time datasets, including online data.
- Practical use and deployment at CDC
  - Beyond the seasonal component
  - Could be predict when there is likely to be an upturn in suicide fatalities?

Lessons Learned

THRIVE: Designing and Deploying Clinical Tools
Powered by Patients' Social Media Data

REALTIME
ENSEMBLE
DATA for
UNDERSTANDING
SUICIDE
EPIDEMIOLOGY

The path forward…

What improvements to these research infrastructure are needed?

- Algorithmic performance for real world translation
- Trust, interpretability, transparency of the algorithms

What are the future research needs?

- Addressing the gap between analytics and interventions
- Social media is not a source of clinical information information
- Consumer voice
- Negative repercussions

What types of training are most important for this type of research?

- Skill acquisition
  - Burden to public health workers
- Digital navigators
- Building social science in the computing curricula
- Ethics awareness/literacy and training

Thanks!

Questions?

munmund@gatech.edu
http://www.munmund.net
@munmun10

Collaborators: Sindhu Ernala, Daejin Choi, Jordan Taylor, Chaitu Konjeti, Michael Birnbaum, John Kane, Asra Rizvi, Anna Van Meter, Steve Sumner, Kristin Holland, Marissa Zwald, Daniel Bowen, Jing Wang