Spatio-temporal Event Forecasting and Precursor Identification

Yue Ning[1], Liang Zhao[2], Feng Chen[3], Huzefa Rangwala[2]

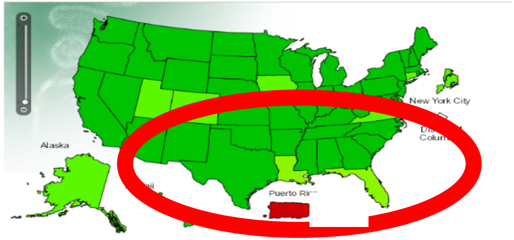
[1]Stevens Institute of Technology
[2]George Mason University
[3]University of Texas at Dallas

Anchorage, Alaska
August 4, 2019

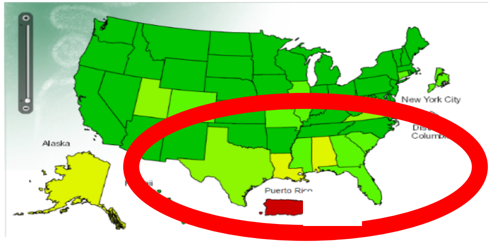
Roadmap

- Introduction and motivation
- Part 1: Precursor Identification
- Part 2: Temporal Event Forecasting
- Part 3: Spatio-temporal Event Forecasting
- Conclusion and Future Directions

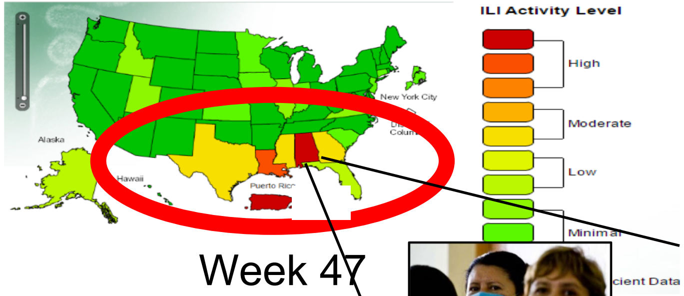
What are societal events?



Week 45



Week 46



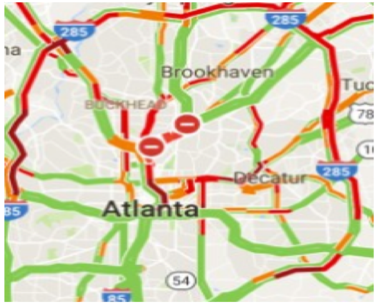
Week 47

Epidemics outbreak on Week 47 ending Nov 22, 2014 in southern region

influenza

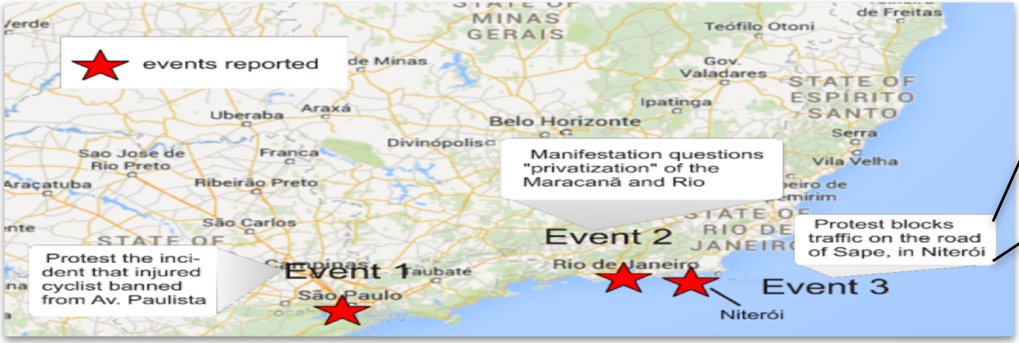
Terrorism events in Africa

Terrorist attacks



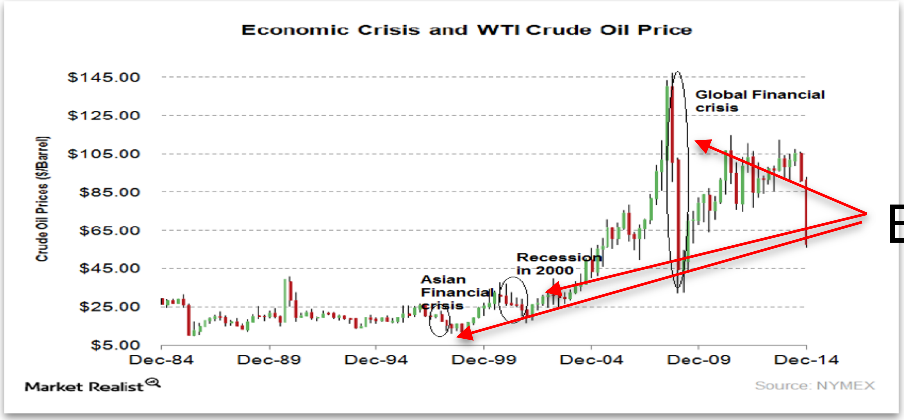
Traffic congestion

What are societal events?

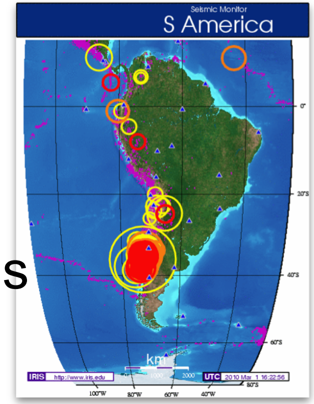


Protests

Civil unrest events on Mar 17, 2013 in Brazil



Economics crisis



Earthquake events

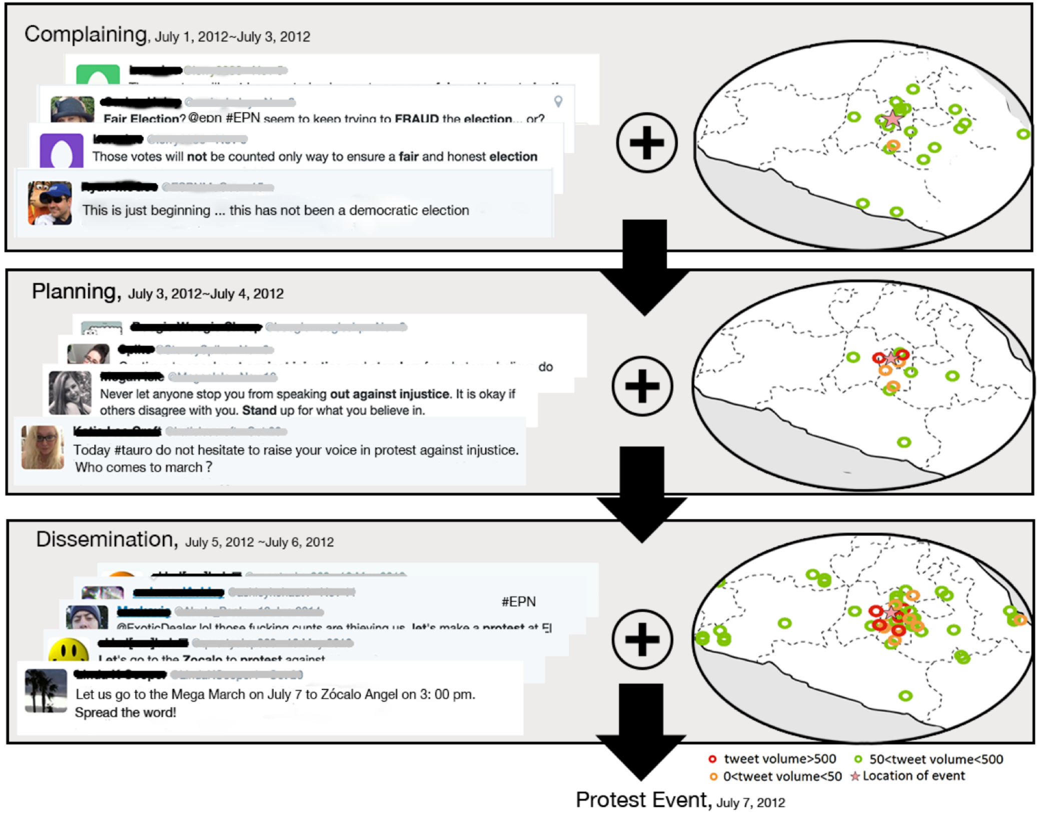
https://www.google.com/search?q=economic+crisis&source=lnms&tbm=isch&sa=X&ved=0ahUKEwiBlsa-uYnfAhVKtlkKHVU7BKoQ_AUIDygC&biw=1050&bih=1114&dpr=1.75#imgrc=Jo1_FhlQuv_7NM:

https://ds.iris.edu/ds/newsletter/vol12/no1/63/february-2010-m88-chile-quake/

Societal Events

Societal Events are Forecastable

Civil unrest

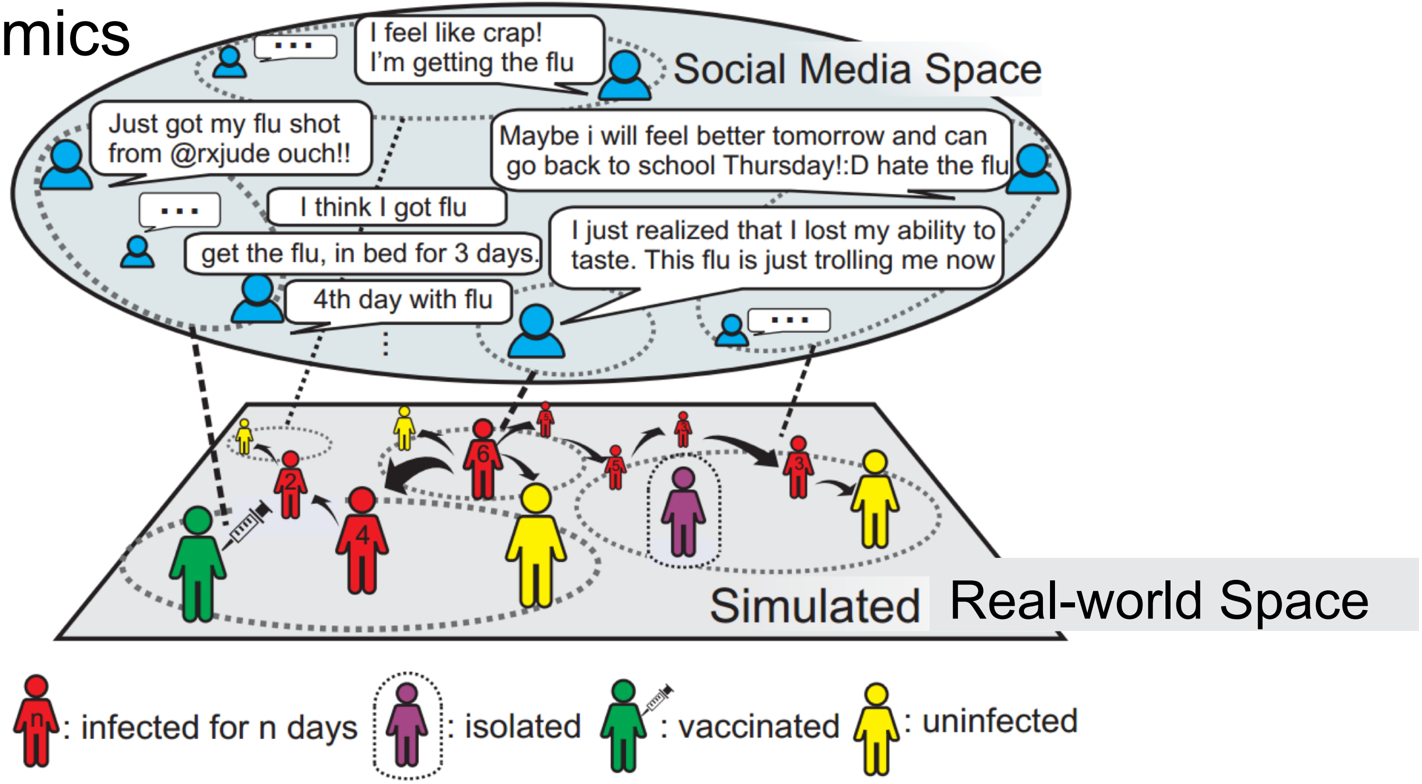


Societal Events are Forecastable

- Transportation congestion

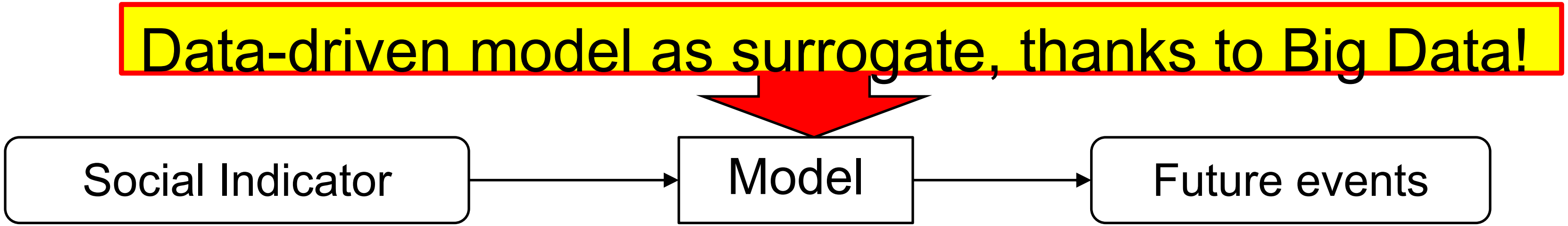
Societal Events are Forecastable

Epidemics



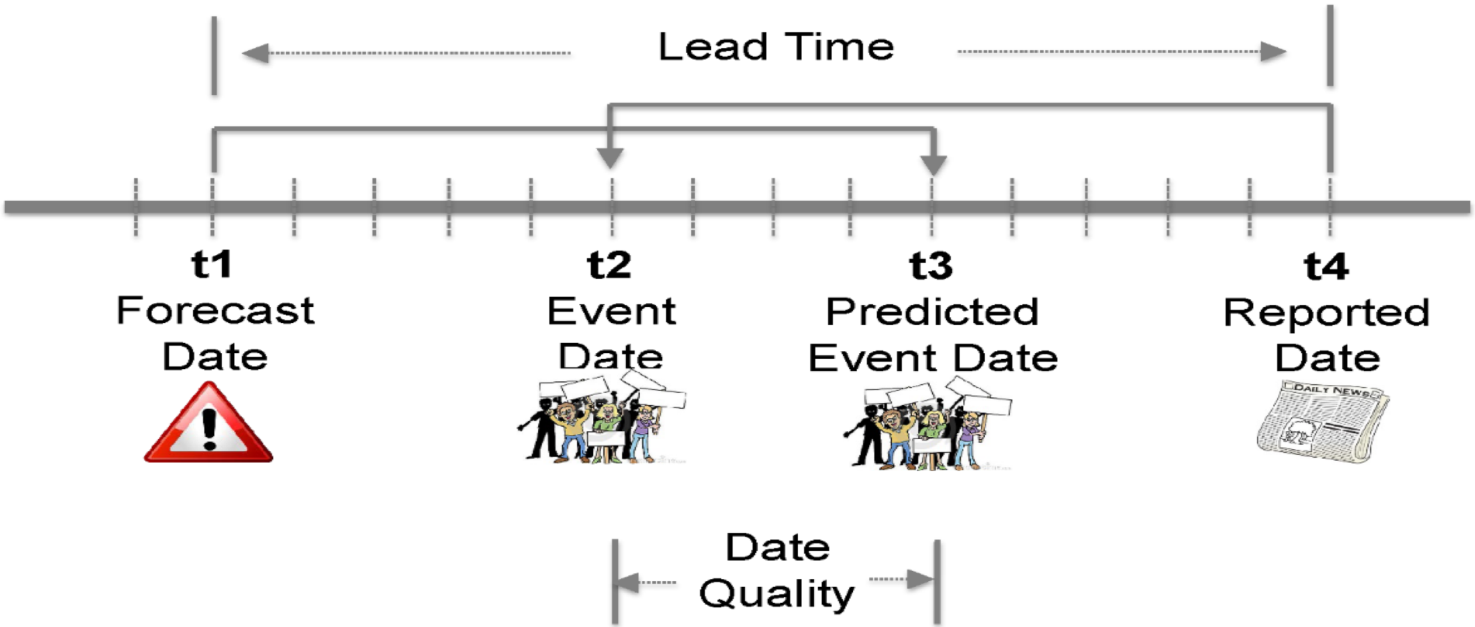
Societal Event Forecasting

- Given some indicators, the task of societal event forecasting is to predict the time, location, and topic of a thing occurring in the future with significant social impact.
- Underlying mechanism of societal events
  - Complex
  - Hard to comprehensively model
  - Largely unknown

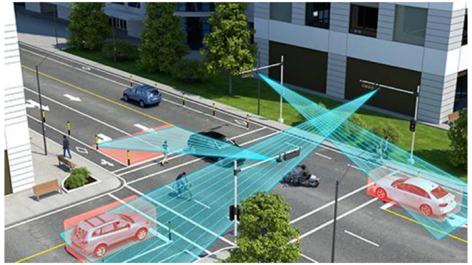


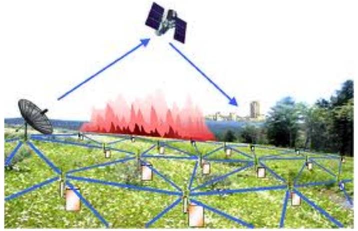
Build the forecaster driven by large historical data

Lead Time


Lead Time
t1
Forecast Date
t2
Event Date
t3
Predicted Event Date
t4
Reported Date
Date Quality

Examples of Social Indicators



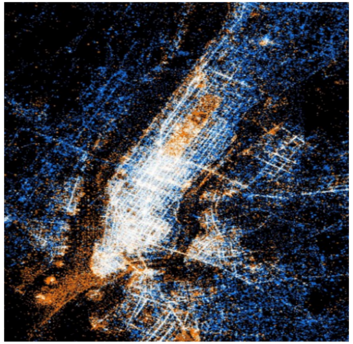


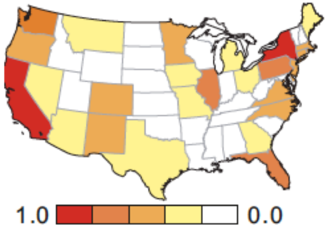
W-ICEWS

The GDELT Project

Characteristics of Social Indicators in Big data Era

- Ubiquitiousness
  - Every user/agent of social media/web/forum is a social sensor.
  - They are everywhere observing the world all the time.
- Timeliness
  - 6,000 tweets every second.
  - 500 million tweets per day.
  - Usually beat the earliest official reports.
- Indicative and predictive signals





Complaints toward Trump on Change.org

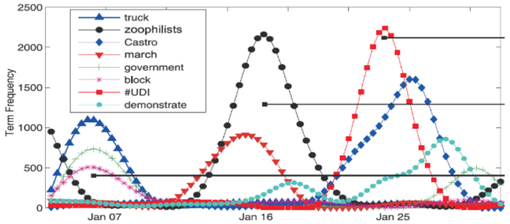
Social Indicators vs. Event Precursors

- Social indicators can be general signals, features, and even distributions in open source data sets

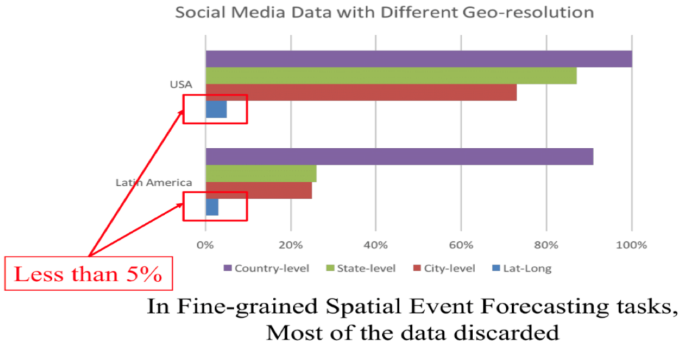
- Event precursors refer to specific examples or instances in the historical data given a prediction

Challenges in Societal Event Forecasting and Precursor Identification

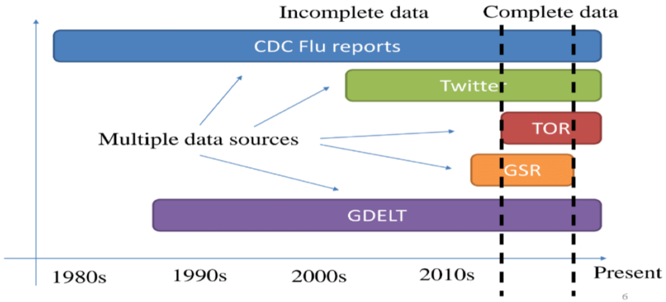
1. Dynamics
new #hashtags, abbreviations, new words



2. Multiple resolution
many messages with country info, few with coordinates



3. Data incompleteness
Reddits enable geo-info this year



4. Big Data Paradox
many data in total, few data for each user

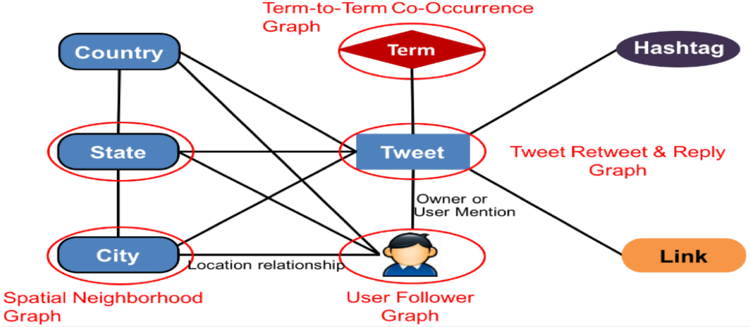
5. Noisy
typos, chit-chat, rumors

Challenges in Societal Event Forecasting and Precursor Identification

7. Multilingual, multi-modal

Dataset	#Tweets	SPA (%)	ENG (%)	POR (%)
Argentina	160,564,890	91.6	7.3	1.1
Brazil	185,286,958	10.1	16.0	73.9
Chile	97,781,414	82.8	16.4	0.8
Colombia	158,332,002	89.8	9.4	0.8
Ecuador	50,289,195	91.1	8.1	0.8
El Salvador	21,992,962	91.5	7.8	0.7
Mexico	197,550,208	83.7	15.4	0.9
Paraguay	30,891,602	92.2	6.4	1.4
Uruguay	10,310,514	89.7	8.8	1.4
Venezuela	167,411,358	92.3	6.9	0.8

8. Heterogeneous network



9. Sparsity in high-dimensional features

Numerous features of vocabulary and profile few are of interest for the research task

Other challenges

- Dependencies among events, e.g., spatial dependencies
- Lack of labeled data, cannot afford to label massive data

- Model interpretability – societal events are influential
- Lack Mechanism Models

Comparisons with Event Detection

Event detection

- Historical or Ongoing events
- Discover anomaly
- Model types
  - Unsupervised learning
- Relevant techniques
  - Anomaly detection
  - Outlier detection
  - Change detection
  - Motif discovery

Event forecasting

- Future events
- Discover the mapping
- Model types
  - Supervised learning
  - Self-supervised learning
  - Semi-supervised learning
- Relevant techniques
  - Autoregressive
  - Markov chain
  - Classification
  - Causal inference

Precursor discovery

- Future events
- Discover the mapping
- Model types
  - Supervised learning
  - Self-supervised learning
  - Semi-supervised learning
- Relevant techniques
  - Multi-instance learning
  - Multi-task learning
  - Classification
  - Deep learning

Comparisons with Spatial Prediction

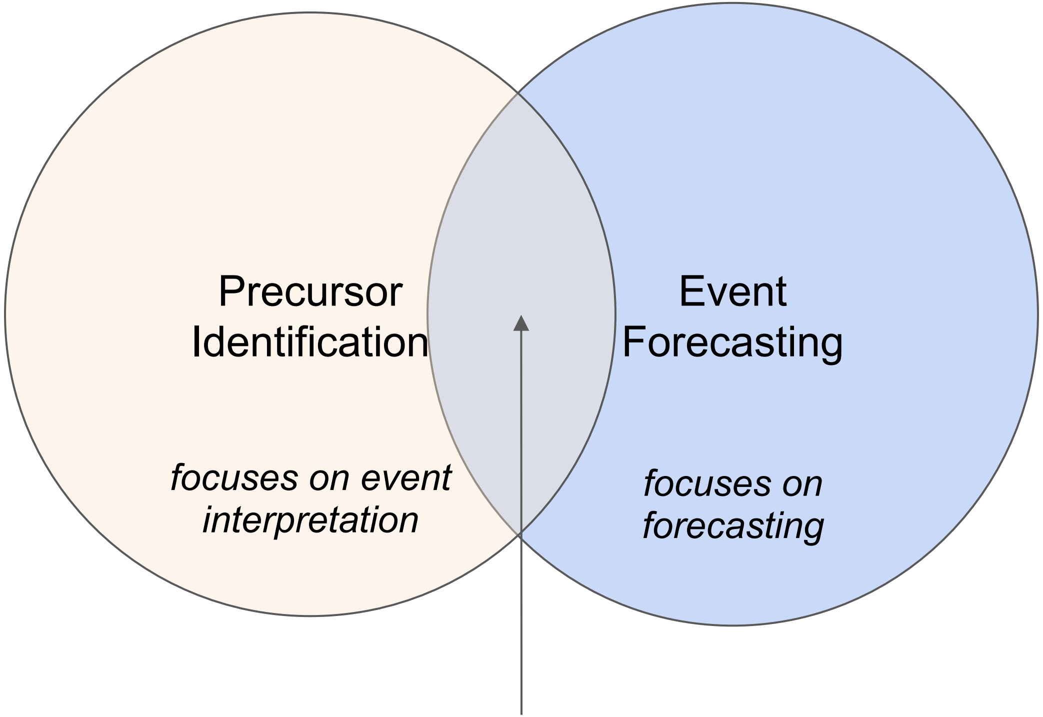
Prediction v.s. Forecasting:

- “Forecasting”: Must be variable in the future.
- “Prediction”: Not necessarily variable in the future.

- Spatial Prediction
  - Dependent variable
    - No need be in the future
    - Usually continuous values –”index”
  - Must have spatial dimension

- **Event Forecasting**
  - **Dependent variable**
    - **Must be in the future**
    - **Usually discrete values – “event”**
  - **No need be in spatial dimension**

Overview



Interpretable Event Forecasting Models

Part 1: Precursor Identification in Spatio-Temporal Event Forecasting

Yue Ning (Stevens Institute of Technology)
Huzefa Rangwala (George Mason University)

2014 Venezuelan National Students Protest

major protests began with student marches led by opposition leaders in 38 cities

Feb. 12

2014 Venezuelan National Students Protest

Opposition Leader, López, called upon students to peacefully protest.

major protests began with student marches led by opposition leaders in 38 cities

Feb. 1

Feb. 12

2014 Venezuelan National Students Protest

López, alongside María Corina Machado launched a campaign to remove Maduro from office.

Jan. 23

Opposition Leader, López, called upon students to peacefully protest.

Feb. 1

major protests began with student marches led by opposition leaders in 38 cities

Feb. 12

2014 Venezuelan National Students Protest

Murder of former Miss Venezuela, Monica Spear.

Former presidential candidate Henrique Capriles shook the hand of President Maduro

Attempted rape of a young student on a university campus in San Cristóbal

The harsh police response to their initial protest

January

López, alongside María Corina Machado launched a campaign to remove Maduro from office.

Jan. 23

Opposition Leader, López, called upon students to peacefully protest.

Feb. 1

major protests began with student marches led by opposition leaders in 38 cities

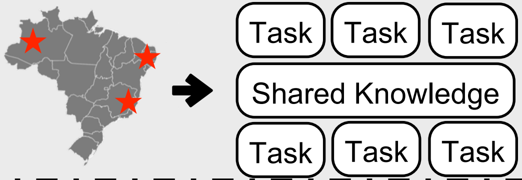
Feb. 12

If social scientists need to do this a lot ……

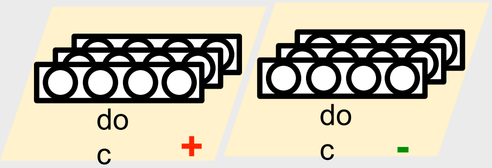
http://phdcomics.com/

The Big Picture

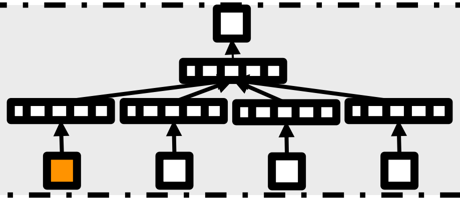
Multi-Task Learning
Relationships between locations;
Spatio-temporal event progression;



Multi-Instance Learning
Label propagation from bag to individual;
Temporal constraints between bags;

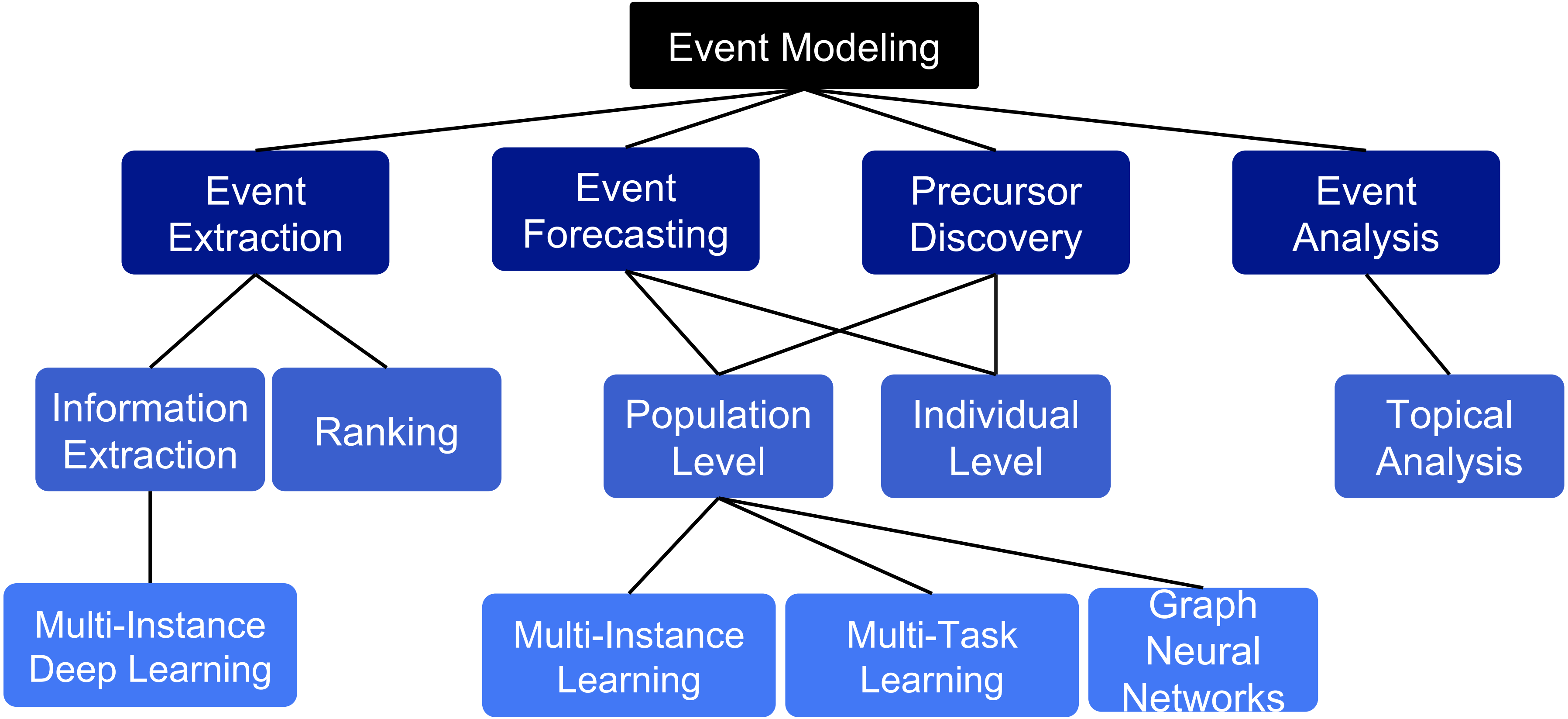


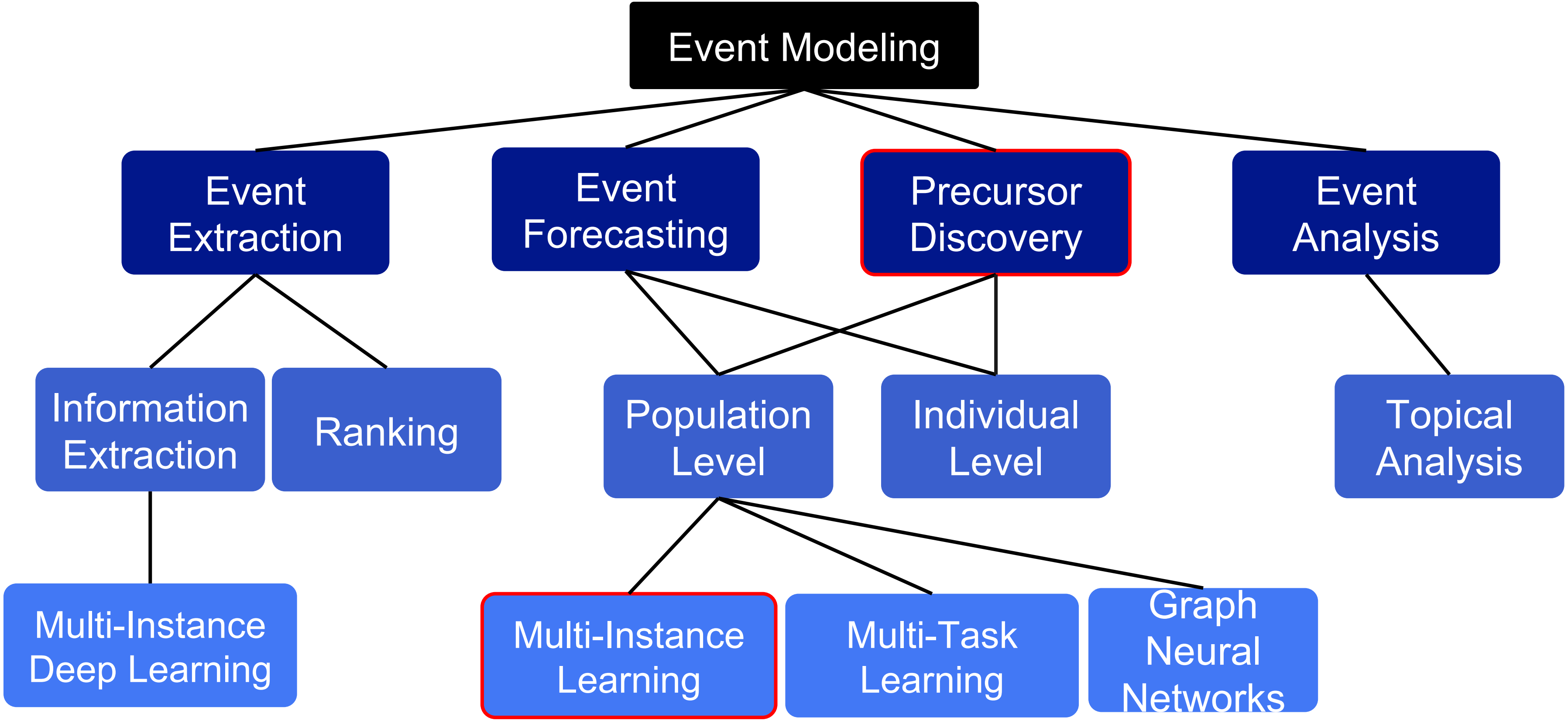
Representation Learning
embeddings; word2vec; doc2vec; etc.



Open Source Indicators
News, blogs, social media, images, videos, time series, etc.

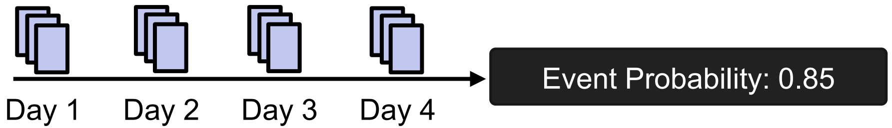
$$(X_1, X_2, \ldots, X_t) \rightarrow Y_{t+1}$$


Event Modeling
Event Extraction
Event Forecasting
Precursor Discovery
Event Analysis
Information Extraction
Ranking
Population Level
Individual Level
Topical Analysis
Multi-Instance Deep Learning
Multi-Instance Learning
Multi-Task Learning
Graph Neural Networks


Event Modeling
Event Extraction
Event Forecasting
Precursor Discovery
Event Analysis
Information Extraction
Ranking
Population Level
Individual Level
Topical Analysis
Multi-Instance Deep Learning
Multi-Instance Learning
Multi-Task Learning
Graph Neural Networks

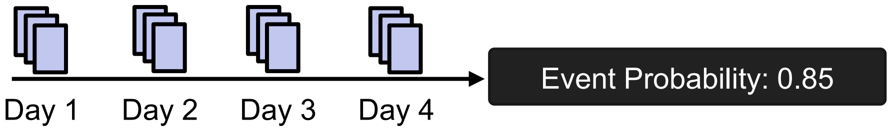
Precursor Discovery

- What is Precursor Discovery in Event Forecasting?
  - Forecast the occurrence of event of interest using historical data

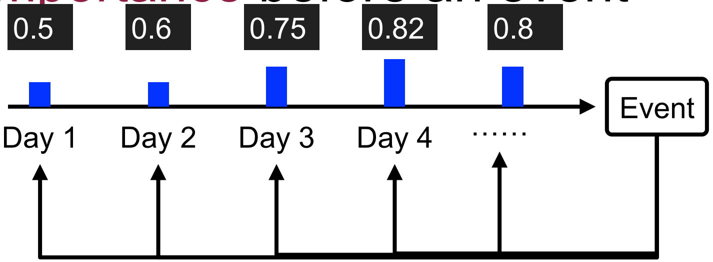


Precursor Discovery

- What is Precursor Discovery in Event Forecasting?
  - Forecast the occurrence of event of interest using historical data

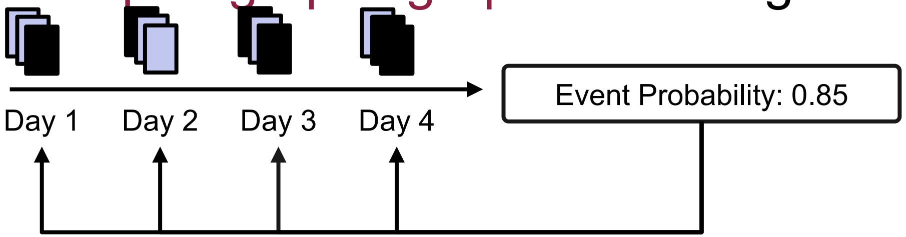


  - Predict days of importance before an event



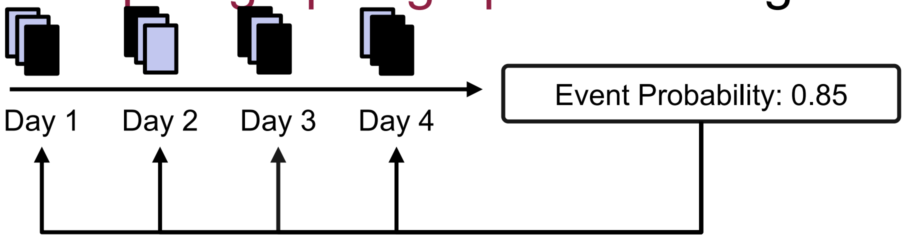
Precursor Discovery

- What is Precursor Discovery in Event Forecasting?
  - Identify key docs/paragraphs/graphs from large-scale input

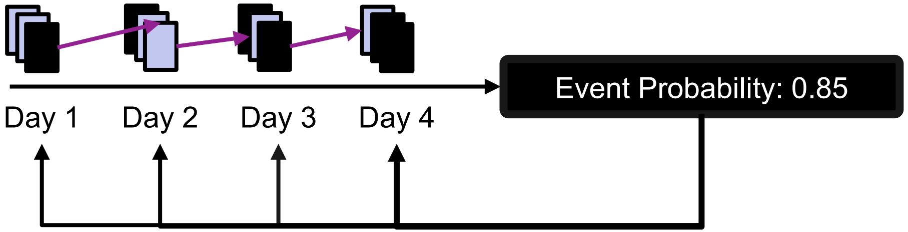


Precursor Discovery

- What is Precursor Discovery in Event Forecasting?
  - Identify key docs/paragraphs/graphs from large-scale input



  - Formalize precursor storylines



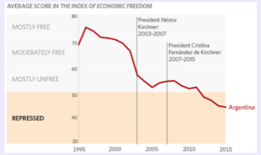
Precursor Storyline

Agenda developed for the migrant perspective

July 28

The Economic Convocation Argentina organized a demonstration

July 29



July 30: Standard of Poverty

Unresolved debts in Argentina

July 31

Aug. 1: International Court of Justice verdict on Argentine debt crisis

Article on government power

Aug. 3

Lear layoffs caused workers to block highway

Aug. 5

Aug. 7: Workers demand for better job opportunities

Existing Methods

- Existing approaches for event forecasting (when), examples:
  - Lasso [Zhao *et al*, TKDE17];
  - Fusion Method [Ramakrishnan *et al*, KDD14];
  - Multi-Task Learning [Zhao *et al*, KDD15];
  - Generative model [Zhao *et al*, SDM15];

  Limitations:
  - Focus on prediction performance, lack of explanation
  - Unable to provide structured evidence

Existing Methods

- Existing approaches for identifying precursors (why), examples:
  - Storytelling [Hossain *et al*, KDD12];
  - Combinational mixed Poisson process [Rong *et al*, KDD15];

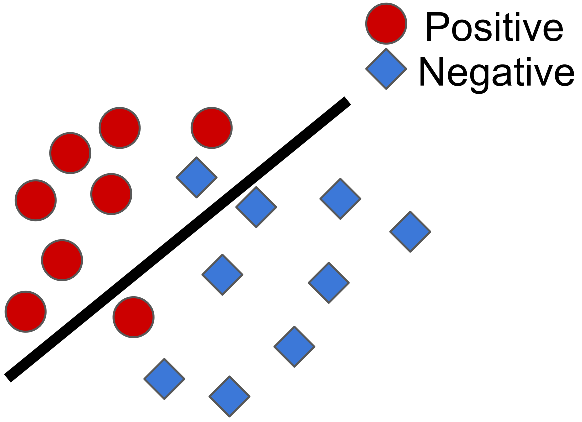
  Limitations:
  - Dependent on observed event sequence (time series, sequential)
  - Lack of predictive value

Modeling Precursors for Event Forecasting via Nested Multi-Instance Learning [Ning et al. KDD16]

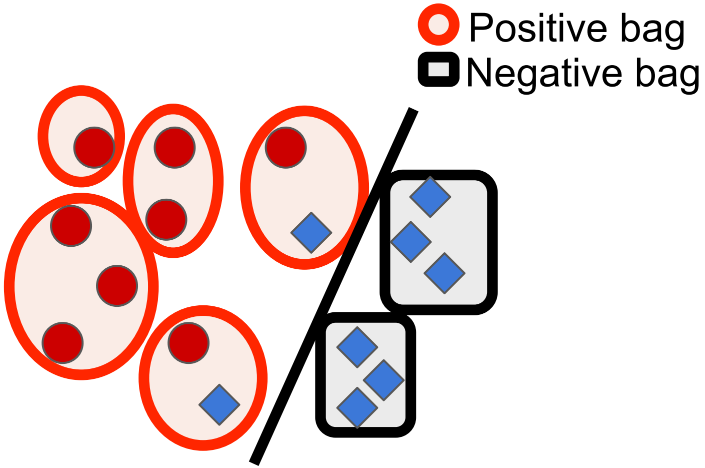
- The proposed method: a nested Multi-Instance Learning framework
  - Solve the above problems together (when & why)
  - Significantly reduce time of manual inspection of specialists/scientists
  - Generate storylines of indicators while predicting events of interest

Multi-Instance Learning

Supervised Learning

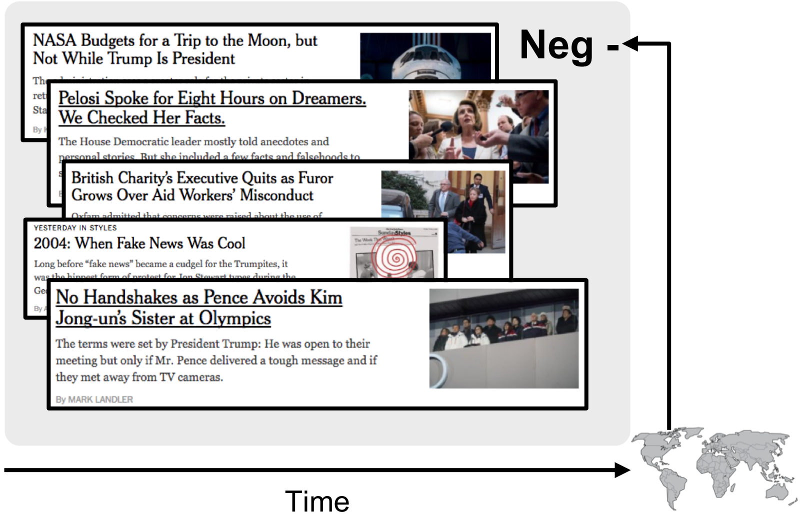


Multi-Instance Learning (MIL)



- Incomplete knowledge about labels in training data
- Propagate bag level supervision to individuals

Event Forecasting in Multi-Instance Learning

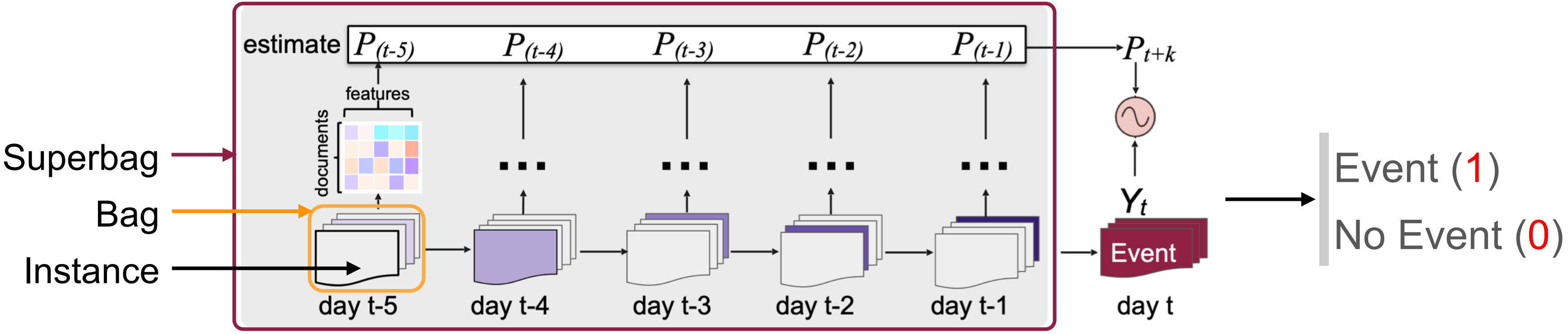


Event: Student Protest
Location: San Paulo
Time: 2014-05-01

Event: None
Location: San Paulo
Time: 2014-05-10

System Overview

- Target Prediction Label, Y



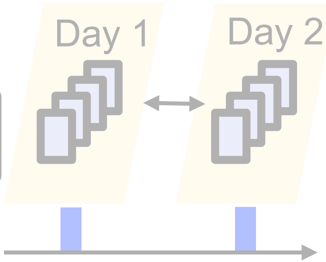
- *Nested Multiple Instance Learning*

- Each news article: *Instance*
- A group of news articles for a day: *Bag*
- A sequential collection of bags: *Super-Bag*
- Label is only associated at the *Super-Bag* Level
- Probabilistic Estimate for every *News Article* (Instance) and *Day* (Bag)

Nested MIL Objective Function

Reduce classification error

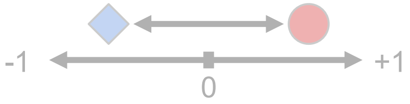
Control the probabilities of consecutive days



$$J(\mathbf{w}) = \frac{\beta}{n}\sum_{\mathbb{S}\in\mathcal{S}} \boxed{f(\mathbb{S}, Y, \mathbf{w})} + \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\\ \mathcal{X}_i,\mathcal{X}_{i-1}\in\mathbb{S}}} \frac{1}{t}\sum_{i=1}^{t} \boxed{g(\mathcal{X}_i, \mathcal{X}_{i-1}, \mathbf{w})}$$

$$+ \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\mathcal{X}_i\in\mathbb{S}\\ \mathbf{x}_{ij}\in\mathcal{X}_i}} \frac{1}{t}\sum_{i=1}^{t} \frac{1}{n_i}\sum_{j=1}^{n_i} \boxed{h(\mathbf{x}_{ij}, \mathbf{w})} + \lambda \boxed{R(\mathbf{w})}$$

Control the margin of instance probabilities

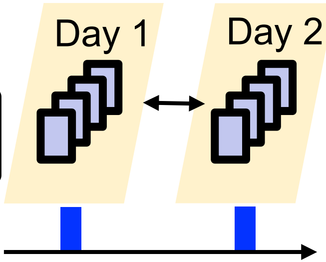


Avoid overfitting

Nested MIL Objective Function

Reduce classification error

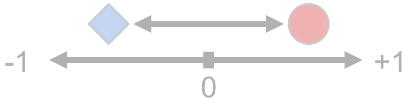
Control the probabilities of consecutive days



$$J(\mathbf{w}) = \frac{\beta}{n}\sum_{\mathbb{S}\in\mathcal{S}} \boxed{f(\mathbb{S}, Y, \mathbf{w})} + \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\\ \mathcal{X}_i,\mathcal{X}_{i-1}\in\mathbb{S}}} \frac{1}{t}\sum_{i=1}^{t} \boxed{g(\mathcal{X}_i, \mathcal{X}_{i-1}, \mathbf{w})}$$

$$+ \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\mathcal{X}_i\in\mathbb{S}\\ \mathbf{x}_{ij}\in\mathcal{X}_i}} \frac{1}{t}\sum_{i=1}^{t} \frac{1}{n_i}\sum_{j=1}^{n_i} \boxed{h(\mathbf{x}_{ij}, \mathbf{w})} + \lambda \boxed{R(\mathbf{w})}$$

Control the margin of instance probabilities



Avoid overfitting

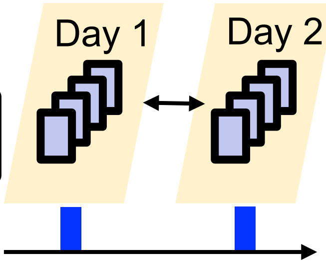
Nested MIL Objective Function

Reduce classification error

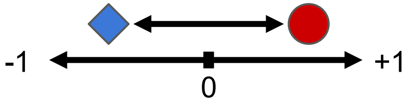
Control the probabilities of consecutive days

$$J(\mathbf{w}) = \frac{\beta}{n}\sum_{\mathbb{S}\in\mathcal{S}} \boxed{f(\mathbb{S}, Y, \mathbf{w})} + \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\\ \mathcal{X}_i,\mathcal{X}_{i-1}\in\mathbb{S}}} \frac{1}{t}\sum_{i=1}^{t} \boxed{g(\mathcal{X}_i, \mathcal{X}_{i-1}, \mathbf{w})}$$

$$+ \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\mathcal{X}_i\in\mathbb{S}\\ \mathbf{x}_{ij}\in\mathcal{X}_i}} \frac{1}{t}\sum_{i=1}^{t} \frac{1}{n_i}\sum_{j=1}^{n_i} \boxed{h(\mathbf{x}_{ij}, \mathbf{w})} + \boxed{\lambda R(\mathbf{w})}$$



Control the margin of instance probabilities

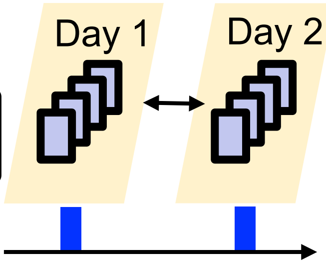


Avoid overfitting

Nested MIL Objective Function

Reduce classification error

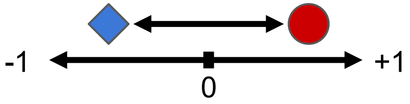
Control the probabilities of consecutive days



$$J(\mathbf{w}) = \frac{\beta}{n}\sum_{\mathbb{S}\in\mathcal{S}} \boxed{f(\mathbb{S}, Y, \mathbf{w})} + \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\\ \mathcal{X}_i,\mathcal{X}_{i-1}\in\mathbb{S}}} \frac{1}{t}\sum_{i=1}^{t} \boxed{g(\mathcal{X}_i, \mathcal{X}_{i-1}, \mathbf{w})}$$

$$+ \frac{1}{n}\sum_{\substack{\mathbb{S}\in\mathcal{S};\mathcal{X}_i\in\mathbb{S}\\ \mathbf{x}_{ij}\in\mathcal{X}_i}} \frac{1}{t}\sum_{i=1}^{t} \frac{1}{n_i}\sum_{j=1}^{n_i} \boxed{h(\mathbf{x}_{ij}, \mathbf{w})} + \lambda \boxed{R(\mathbf{w})}$$

Control the margin of instance probabilities



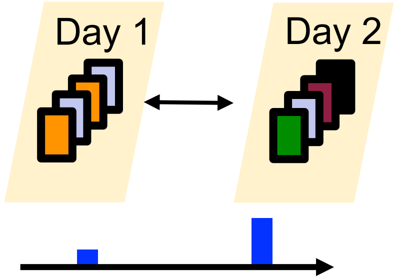
Avoid overfitting

Nested MIL-Delta Objective Function

$$J(\mathbf{w}) = \frac{\beta}{n} \sum_{\mathbb{S} \in \mathcal{S}} f(\mathbb{S}, Y, \mathbf{w}) + \frac{1}{n} \sum_{\substack{\mathbb{S} \in \mathcal{S}; \\ \mathcal{X}_i, \mathcal{X}_{i-1} \in \mathbb{S}}} \frac{1}{t} \sum_{i=1}^{t} \boxed{g(\mathcal{X}_i, \mathcal{X}_{i-1}, \mathbf{w})}$$

$$+ \frac{1}{n} \sum_{\substack{\mathbb{S} \in \mathcal{S}; \mathcal{X}_i \in \mathbb{S} \\ \mathbf{x}_{ij} \in \mathcal{X}_i}} \frac{1}{t} \sum_{i=1}^{t} \frac{1}{n_i} \sum_{j=1}^{n_i} h(\mathbf{x}_{ij}, \mathbf{w}) + \lambda R(\mathbf{w})$$

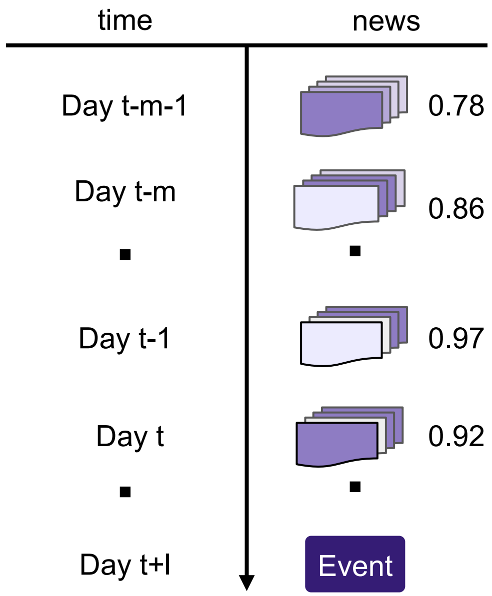
$$g(\mathcal{X}_i, \mathcal{X}_{i-1}, \mathbf{w}) = \underbrace{\Delta(\mathcal{X}_i, \mathcal{X}_{i-1})}_{\text{Cross-bag similarity}} (P_i - P_{i-1})^2$$



Precursor Discovery in Nested MIL

procedure PD-nMIL
  Input: $\mathcal{S} = \{(\mathbb{S}_r, Y_r)\}_{r \in n^+}, \mathcal{M}$
  Output: $\{(ps_r, Y_r)\}_{r \in n^+}$
  for super bag (S_r, Y_r) **do**
    $ps_r = []$
    for t = 1,2,...,h(history days) **do**
      $y_t = []$
      for $\mathbf{x}_{tm} \in \mathcal{X}_t$ **do**
        $\hat{y}_{tm} = \sigma(\hat{\mathbf{w}}\mathbf{x}_{tm})$
        if $\hat{y}_{tm} > \tau$ **then**
          $y_t \leftarrow (m, \hat{y}_{tm})$
      sort(y_t) by $\hat{y}_{tm}$ in descending order
      $ps_r \leftarrow$ m where m in top(y_t)
  return $\{(ps_r, Y_r)\}_{r \in n^+}$

<u>Selection of precursors based on their estimated probabilities</u>

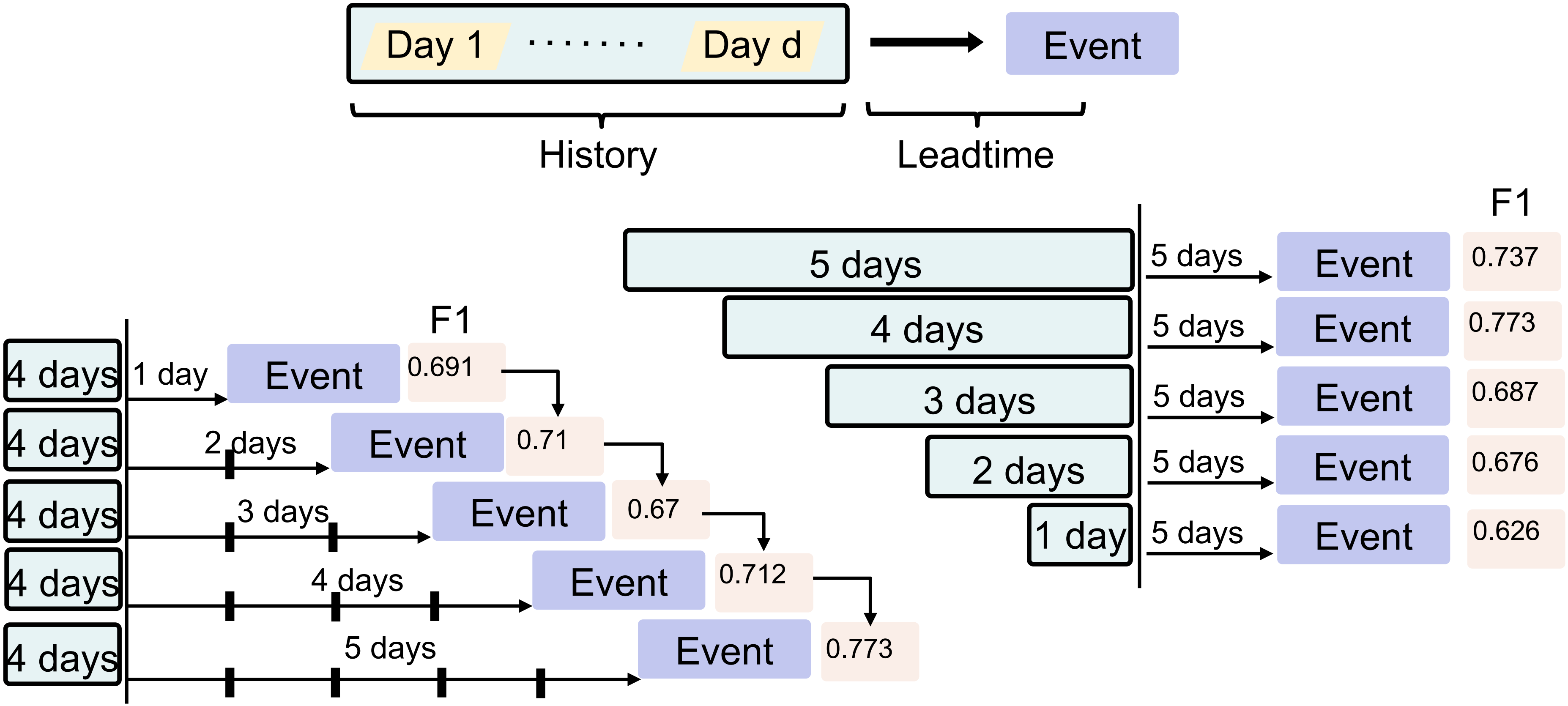


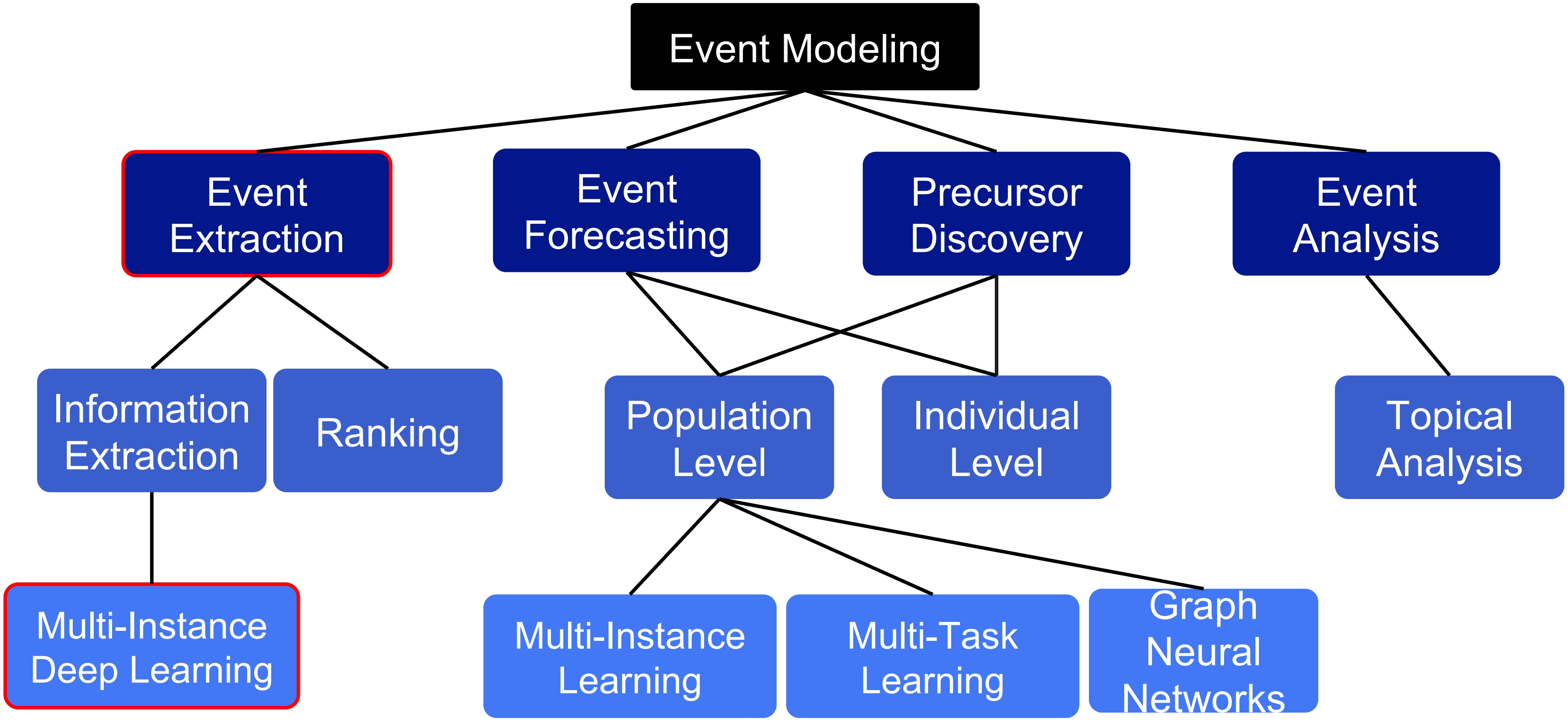
Predictive Performance

	Argentina		Brazil		Mexico	
	Acc	F-1	Acc	F-1	Acc	F-1
SVM	0.611($\pm$0.034)	0.406($\pm$0.072)	0.693($\pm$0.040)	0.598($\pm$0.067)	0.844($\pm$0.062)	0.814($\pm$0.091)
MI-SVM	0.676($\pm$0.026)	0.659($\pm$0.036)	0.693($\pm$0.040)	0.503($\pm$0.087)	0.880($\pm$0.025)	0.853($\pm$0.040)
rMIL-NOR	0.330($\pm$0.040)	0.411($\pm$0.092)	0.505($\pm$0.012)	0.661($\pm$0.018)	0.499($\pm$0.009)	0.655($\pm$0.025)
rMIL-AVG	0.644($\pm$0.032)	0.584 ($\pm$0.055)	0.509($\pm$0.011)	0.513($\pm$0.064)	0.785($\pm$0.038)	0.768($\pm$0.064)
GICF	0.589($\pm$0.058)	0.624($\pm$0.048)	0.650($\pm$0.055)	0.649 ($\pm$0.031)	0.770($\pm$0.041)	0.703($\pm$0.056)
nMIL	**0.709**($\pm$0.036)	0.702($\pm$0.047)	**0.723**($\pm$0.039)	0.686($\pm$0.055)	**0.898**($\pm$0.031)	**0.902**($\pm$0.030)
nMIL-Δ	0.708($\pm$0.039)	**0.714**($\pm$0.034)	0.705($\pm$0.048)	**0.698**($\pm$0.045)	0.861($\pm$0.014)	0.868($\pm$0.014)
nMIL-Ω	0.687($\pm$0.038)	0.680($\pm$0.045)	0.713($\pm$0.028)	0.687($\pm$0.038)	0.871($\pm$0.013)	0.879($\pm$0.014)

1. Nested structure models: nMIL, nMIL-Delta, nMIL-Omega
2. The averaged daily estimates help predict events of interest
3. Effect of time accumulation > a single input

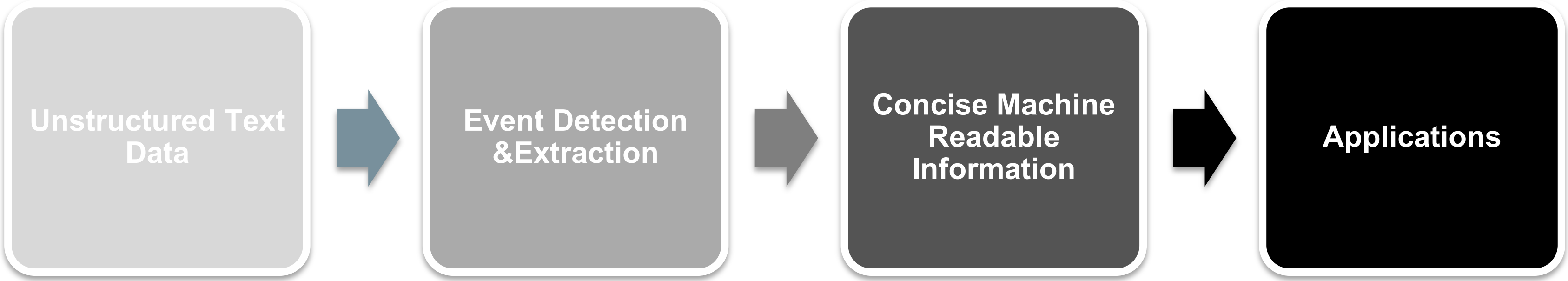
How Early can NMIL Forecast?



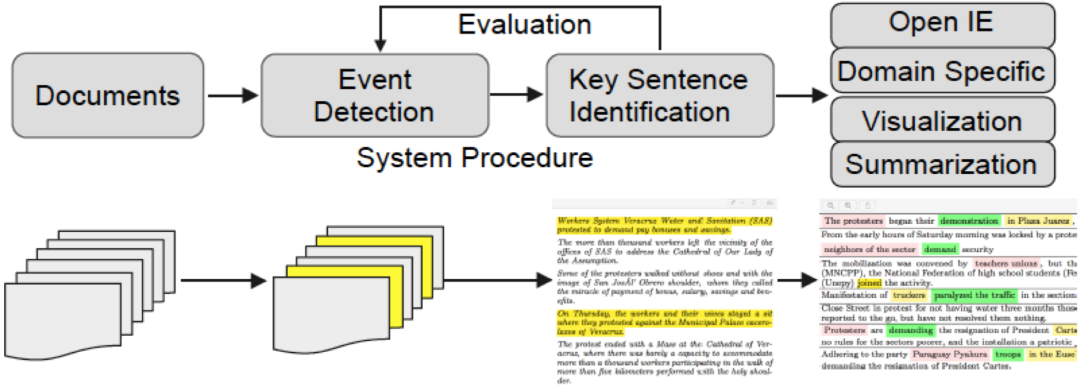

Event Modeling
Event Extraction
Event Forecasting
Precursor Discovery
Event Analysis
Information Extraction
Ranking
Population Level
Individual Level
Topical Analysis
Multi-Instance Deep Learning
Multi-Instance Learning
Multi-Task Learning
Graph Neural Networks

Identifying Key Sentences and Detecting Events
[W. Wang et al. CIKM16]

- Most of the available text data are expressed using natural languages
- Transform the unstructured text data into machine readable format
- Help human analysts ingest broader information with less effort



Problem Formulation & Motivations



- Automatically detect civil unrest events.
- Identify key sentences without ground truth labels.
- Allows for event summarization
- Downstream event encoding
- Visualization and human-in-the-loop

Challenges

Labor Intensive

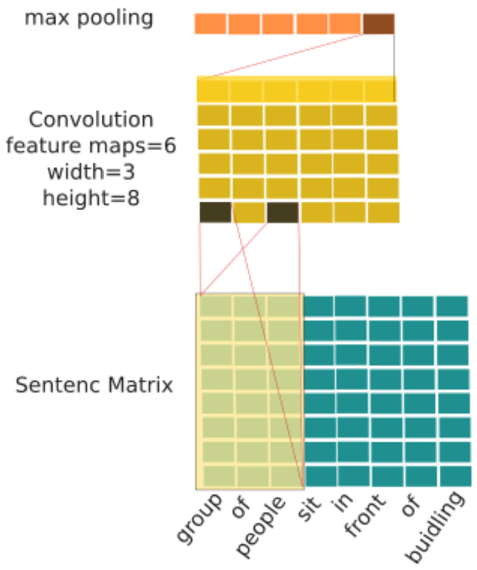
Time Consuming

Hard to Adapt to new Domain

Document label is relatively easy to obtain

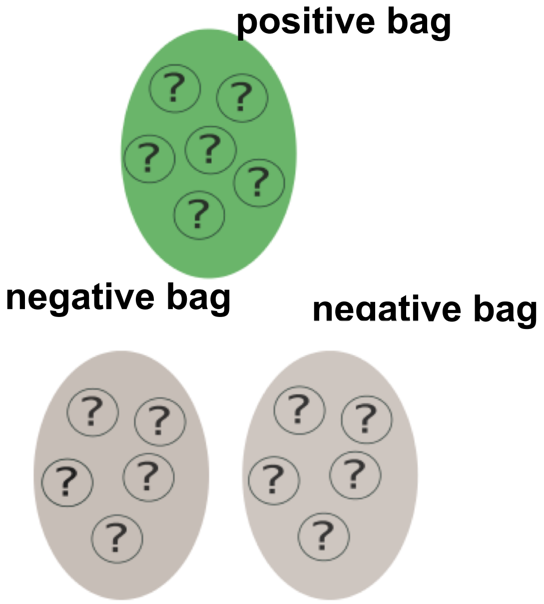
Multi-Instance Learning + Representation Learning

Convolutional Neural Network



Learn distributed representation for instances

Multi-Instance Learning



Transfer bag label to instance label

Standard Supervised Learning

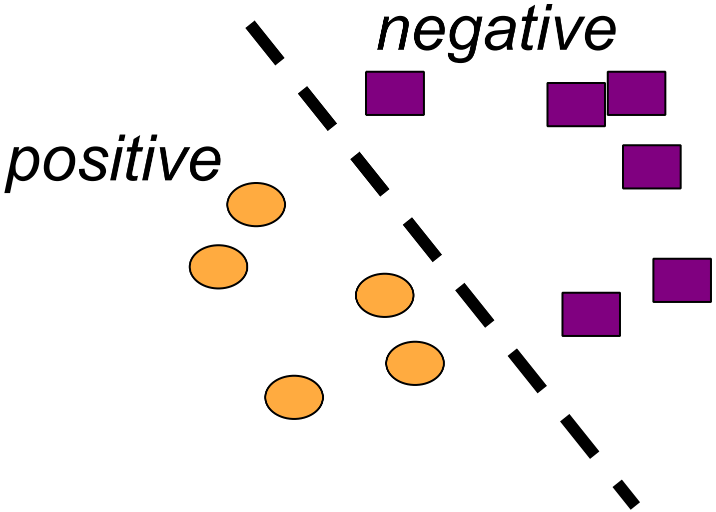
Standard Supervised Learning

Find a function from the input space (X) to the output space (Y)

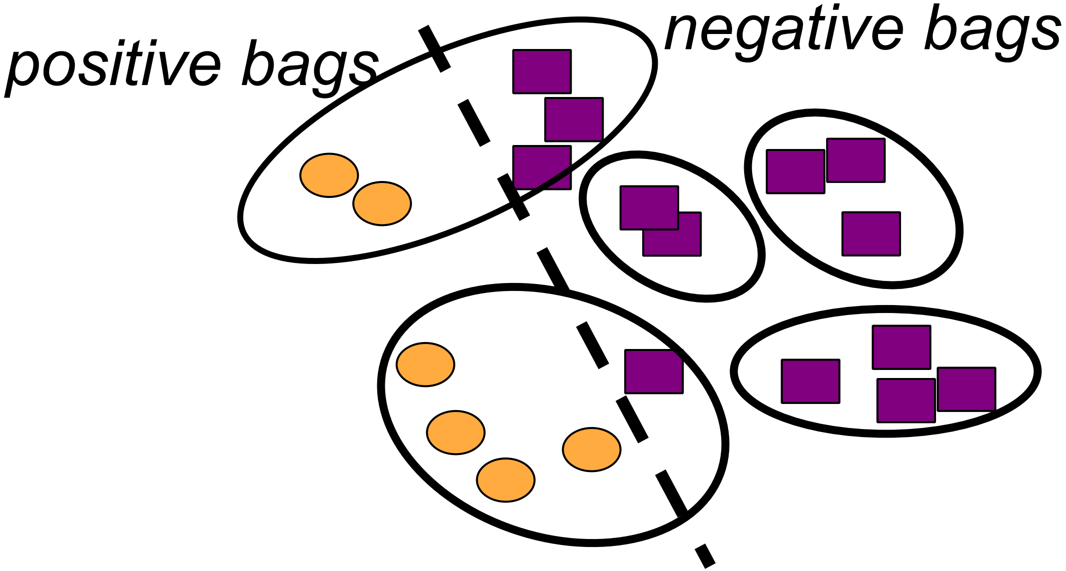
$$f : X \rightarrow Y$$

such that prediction error is low on **unseen examples**

Multiple Instance Learning



Traditional supervised learning



Multiple-instance learning

[Dietterich et al. 1997]

Typical MIL Assumptions

- no positive instances in negative bag
- at least one positive instances in positive bag
- at least k% of positive instances in positive bag
- instances are independently drawn from distribution

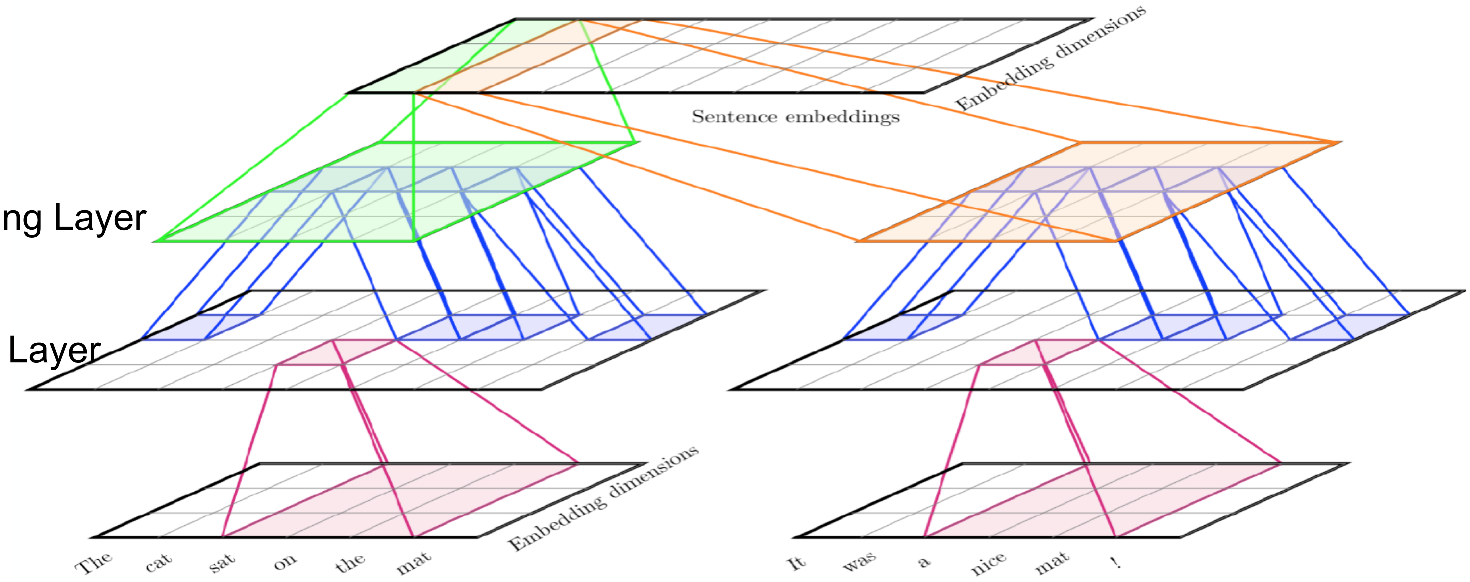
Key Instance Detection

- The task of classic MIL is to train a classifier that labels new bags
- Sometimes positive instances are expected to be identified
  - Protest event detail
  - Customer review
- It is obviously desirable if we can label instances, which will explicitly recognize positive instances

Convolutional Neural Network [Denil et. al. 2014]

K-max Pooling Layer

Convolution Layer



Distributed Representation of Sentences

Local and Context Information

Chile student protests point to deep discontent

By Gideon Long
BBC News, Santiago

11 August 2011 | Latin America & Caribbean

Share

Chile is usually regarded as one of the most orderly and stable countries in South America, so the images that have come out of the capital, Santiago, in recent days have been especially shocking.

Students are calling for free and equal schooling

Thousands of high school and university students have marched through the capital's streets, as well as those of other major cities, demanding a radical overhaul of the education system.

Invariably the demonstrations have ended in violent clashes between masked youths and police officers armed with tear gas and water cannon.

Shops and offices on Santiago's main thoroughfare, the Alameda, have been looted and destroyed.

Model Overview

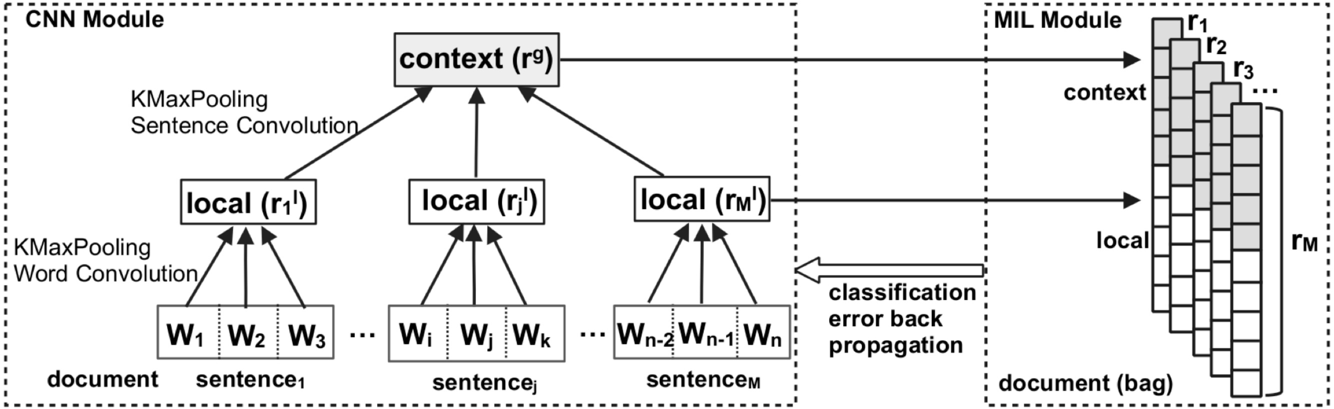


Figure: MI-CNN Model Overview

- We consider each **document** as a **bag** and each **sentence** as an **instance**
- Two layers of Convolutional Layer to construct the Local and Context representation for instances
- Classification information from MIL module is used to fine tuned the Instance Representation

Loss Function

$$L(x,y;\theta,W,F,b) = \underbrace{\frac{1}{N}\sum_{n}^{N}(1-y_n)\log P_n + y_n \log(1-P_n)}_{\text{bag level loss}}$$ → **Cross Entropy Loss**

$$+ \underbrace{\frac{\alpha}{N}\sum_{n}^{N} y_n \max(0, |K_n| - Q_n) + (1-y_n)Q_n}_{\text{instance ratio control loss}}$$ → **Control num of key instances**

$$+ \underbrace{\frac{\beta}{N}\sum_{n}^{N}\frac{1}{M_n}\sum_{m}^{M_n}\max(0, m_0 - \mathrm{sgn}(p_m^n - p_0)\theta^T r_m^n)}_{\text{The instance-level loss}}$$ → **Control probability margin**

$$+ \underbrace{\frac{\gamma}{(\sum_n M_n)^2}\sum_{n}^{N}\sum_{i}^{N}\sum_{m}^{M_n}\sum_{j}^{M_i}(p_m^n - p_j^i)^2 e^{(-\|r_m^n - r_j^i\|_2^2)}}_{\text{instance-level manifold propagation}}$$ → **Control sentences similarity**

$$Q_n = \sum_m 1(p_m^n > 0.5)$$

$$k = |x_i| \times \eta$$

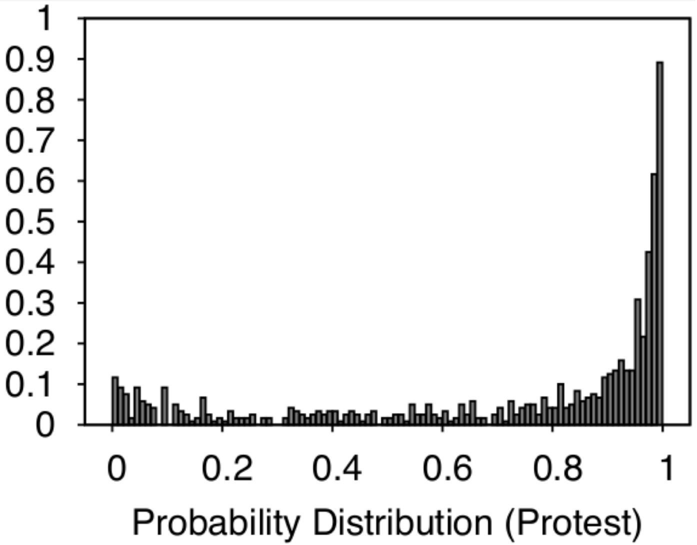
Train the model with Back Propagation

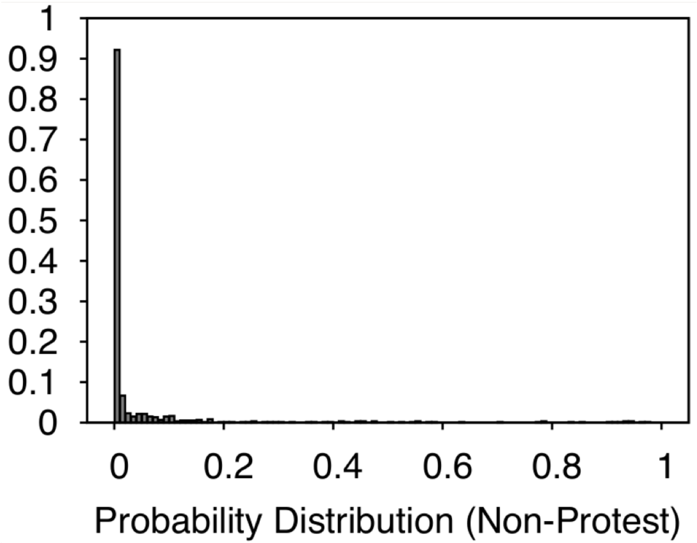
Experiments Performance

	Precision	Recall	F1
SVM	0.818 (0.019)	0.720 (0.008)	0.765 (0.009)
MISVM	0.724 (0.030)	0.584 (0.017)	0.646 (0.018)
CNN Model	0.732 (0.033)	0.783 (0.026)	0.756 (0.007)
GICF	**0.833 (0.019)**	0.421 (0.09)	0.553 (0.086)
MI-CNN (Max)	0.685 (0.030)	0.730 (0.029)	0.706 (0.018)
MI-CNN (Avg)	0.731 (0.069)	0.789 (0.042)	0.759 (0.026)
MI-CNN (Context + Dynamic K)	0.742 (0.036)	**0.813 (0.041)**	**0.775(0.006)**

Table: Experiment Results for Event Detection (Protest or not)

Experiments Performance





The histogram of predicted positive probability for protest and non-protest articles for test set

Compared with Heuristic Methods

Baseline Methods

- Keywords Protest: Select sentences containing protest related words
- Random Sentences: Randomly choose set of sentences
- Start/End Sentences: Select sentences from start and end of articles

	Precision(Std.)	Recall(Std.)	F1(Std.)
Keywords protest	0.755 (0.021)	**0.638 (0.017)**	0.692 (0.018)
Random Sentences	0.681 (0.026)	0.433 (0.019)	0.551 (0.018)
Start/End Sentences	0.751 (0.022)	0.555 (0.026)	0.638 (0.019)
Our model	**0.761 (0.015)**	0.635 (0.024)	**0.693 (0.019)**

Table: SVM classification performance for article label prediction based on sentences selected from different methods

Extracted Key sentences

Positive Sentences	Score	Keywords	Start/End
The protesters began their demonstration in Plaza Juarez, advanced by 16 September to Hidalgo.	0.9992	Yes	No
From the early hours of Saturday morning was locked by a protest the Francisco Fajardo highway from Caricuao, neighbors of the sector demand security	0.9991	Yes	No
The mobilization was convened by teachers unions, but the national March of public colleges and private (MNCPP), the National Federation of high school students (Fenaes) and the Center Union of secondary students (Unepy) joined the activity.	0.9991	No	No
Manifestation of truckers paralyzed the traffic in the section clean-Roque Alonso	0.9991	Yes	Yes
Close Street in protest for not having water three months those who protested pointed out that the problem was reported to the go, but have not resolved them nothing.	0.9991	Yes	Yes
Protesters are demanding the resignation of President Cartes, since they consider that - as they understand - no rules for the sectors poorer, and the installation a patriotic junta in power.	0.9991	Yes	No
Adhering to the party Paraguay Pyahura troops in the Eusebio Ayala Avenue heading to downtown Asuncion, demanding the resignation of President Cartes.	0.9991	No	Yes
From 09:00 hours, tens of inhabitants of the municipal head were concentrated at the entrance of Arcelia and almost 10 o'clock began a March toward the Center, which showed banners against staff of the PF.	0.999	Yes	No
Nurses were stationed opposite the hospital with placards to demand to the authorities of the IPS that their claims are solved immediately.	0.9989	No	No
A group of taxi drivers protested this Monday morning in the central town of el Carrizal municipality, in Miranda State, according to @PorCarrizal the demonstration is due to that, he was denied the circulation to the drivers who benefited from the transport mission.	0.9988	Yes	Yes
Negative Sentences	**Score**	**Keywords**	**Start/End**
Bled some guardians, also protesters, friends and family that went with them.	0.172	Yes	No
The parade by the 195 years of independence of Ambato yesterday (November 12) had a different connotation.	0.0125	Yes	No
This morning, the situation is similar, as already record barricades and demonstrations in the same place, by what police is already around the terminal.	0.0109	Yes	No
The young man asked that they nicely other costume to so participate in the parade.	0.0097	No	No
Employees announced that they will be inside until you cancel them owed assets.	0.0093	No	No
Workers arrived Thursday to the plant where the only person who remained on duty in the place who has not claimed his salary joined the protest.	0.0088	No	No

Table: List of positive and negative sentences selected by our model sorted by score

Sentences Highlighting Cases

Workers System Veracruz Water and Sanitation (SAS) protested to demand pay bonuses and savings.

The more than thousand workers left the vicinity of the offices of SAS to address the Cathedral of Our Lady of the Assumption.

Some of the protesters walked without shoes and with the image of San Jos Obrero shoulder, whom they called the miracle of payment of bonus, salary, savings and benefits.

On Thursday, the workers and their wives staged a sit where they protested against the Municipal Palace cacerolazos of Veracruz.

The protest ended with a Mass at the Cathedral of Veracruz, where there was barely a capacity to accommodate more than a thousand workers participating in the walk of more than five kilometers performed with the holy shoulder.

Angelica Navarrete, general secretary of the Union of SAS, insisted on Tuesday that if they do not receive what they owe, they will strike.

During the march, at the height of Zamora Park, a passenger bus of the coastline they were pounced on protesters, upset because he wanted to spend and the march went through, but no injuries.

According to the protesters, the SAS, owed to workers 85 thousand 300 million pesos.

......

Wage and Employment, Labor,
12/19/2015,
[Mexico, Veracruz, Veracruz]

Activists claim the government and Congress of Veracruz to pass legal reform violates international treaties on reproduction. Photo: Roberto Garca Ortiz

Mexico DF. While thousand 647 women still missing in Veracruz since 2010, the government of Javier Duarte de Ochoa puts his effort in punishing those who wish to terminate their pregnancy, because the constitutional reform that protects life from conception merely criminalize, they said activists . At a rally they demanded the governor to veto the amendment which he drove.

Members of different groups demonstrated in front of the representation of the state of Veracruz in Mexico City against the "anti-abortion" reform, local MPs approved last Thursday 21. That delivered a letter to the governor in which he requested to avoid the initiative progresses.

They urged lawmakers not to approve it in a second round in May. Before you reach that round, the town councils of 240 municipalities should discuss. 121 needs to accept it, so the call was also for them.

The right to life, for desparecidas

The amendment to article 4 of the state Constitution is a "smokescreen" to the serious problem of disappearances and increased 500 percent in the number of murders of women in the state, said Adriana Jimenez Patln, of the Network for Sexual Rights and Reproductive in Mexico (ddeser).

......

Government Policies, General Population,
01/27/2016,
[Mexico, Distrito Federal, Ciudad de México]

Event Type and Population specific tokens

$$\text{Score}_c(w) = f_{c,w} \log \frac{N}{n_w}$$

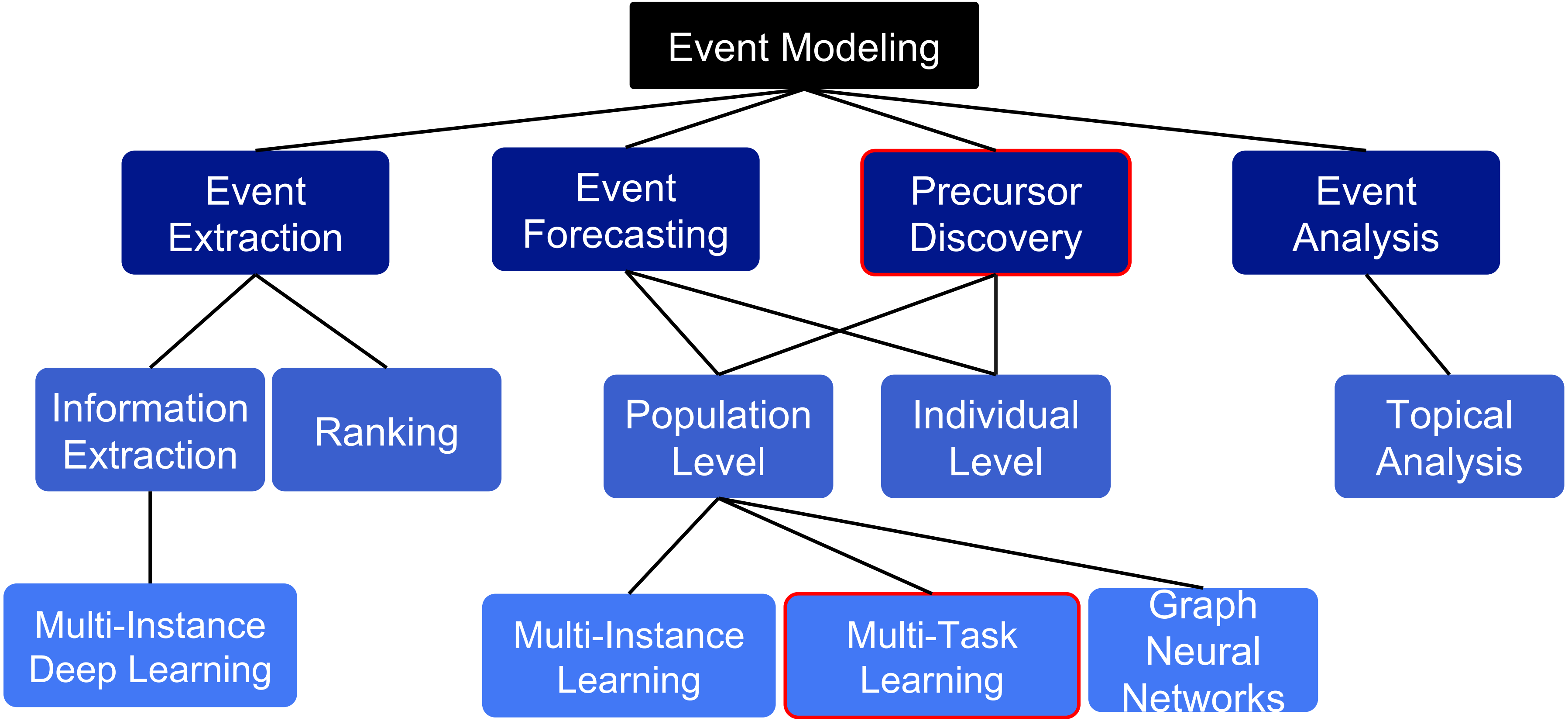
$c \in$ {*Business, Media, Medical, ..., Housing, Energy, Government*}
w : *token in article*
$f_{c,w}$: *frequency of token in category c*
N : *total number of articles*
n_w : *number of articles containing token w*

EventPopluation					**EventType**				
Business	**Media**	**Medical**	**Legal**	**Education**	**Housing**	**Energy**	**Economic**	**Employment**	**Government**
sellers	communicators	health	grant	students	housing	water	producers	worker	national
commercial	journalists	medical	congress	education	neighborhood	energy	mobilization	official	march
drivers	express	hospital	judges	national	service	company	route	drivers	government
strike	agreement	unemployment	specialties	government	terms	sector	budget	payment	demand
transport	exhibited	doctor	reprogramming	teachers	family	neighbors	carriers	wages	square
measure	profession	nursing	budget	college	group	lack	association	unemployment	city
carriers	legislation	clinics	explanation	professor	transfers	supply	ministry	guild	front
public	guards	patients	deny	faculty	place	population	cooperators	employee	hours
municipal	intervened	welfare	approve	school	mutual	authority	peasants	company	demonstration
strength	collaboration	power	exist	dean	bill	organization	PLRA	job	students

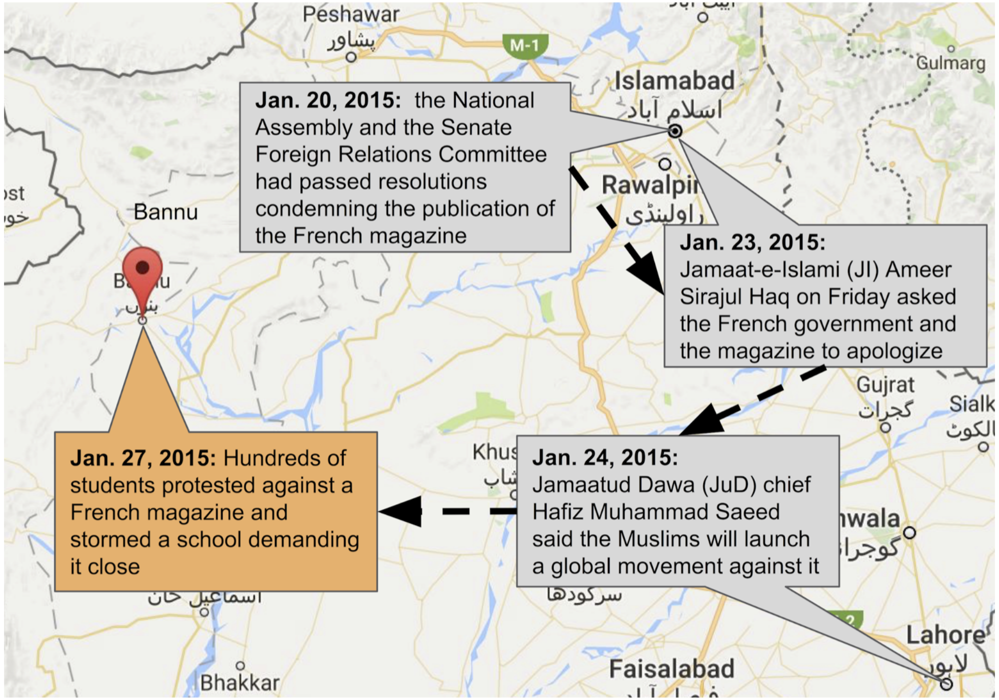
Table: Top scored terms in different categories of event populations and event types. All the articles are represented by the MI-CNN model selected key sentences

Key Takeaways

- Joint Event Detection and Extractions as Multiple Instance Learning.
- Bag Labels Transferred to Instance Labels.
    - Bag to Instance Aggregation Functions
- Distributed Sentence Representation combines local and global context.
    - Updated via back propagation
- Downstream: Visualizer, Event Encoder, Knowledge Graph Construction.

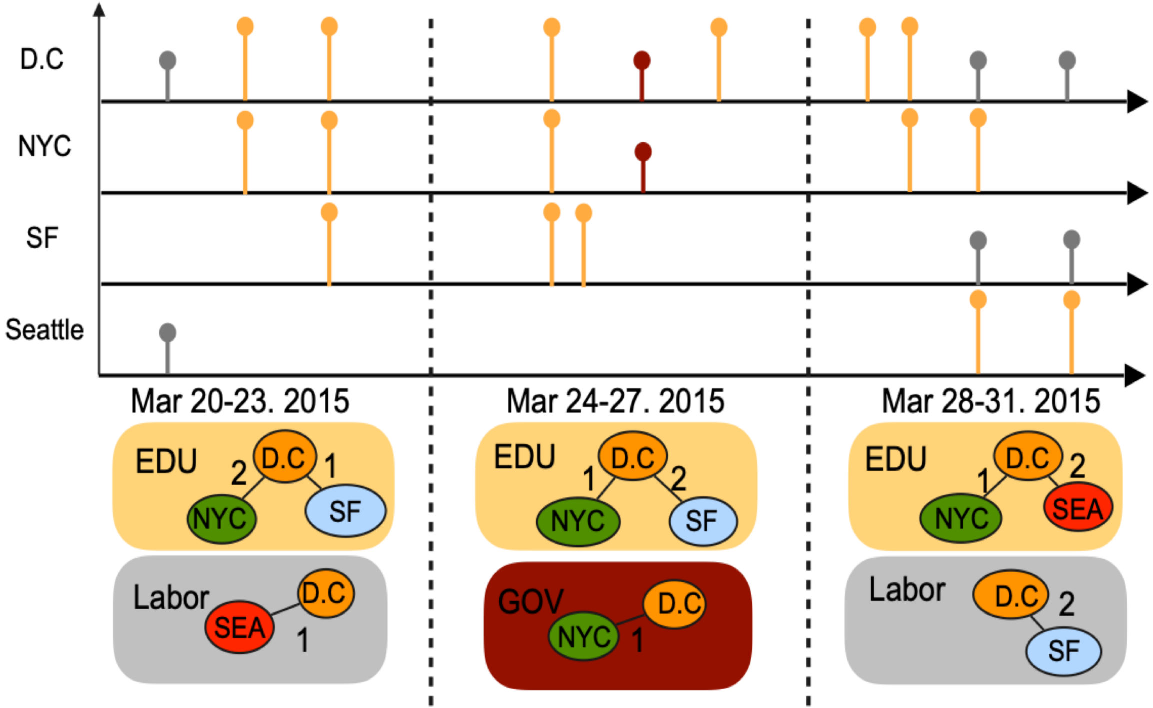

Event Modeling
Event Extraction
Event Forecasting
Precursor Discovery
Event Analysis
Information Extraction
Ranking
Population Level
Individual Level
Topical Analysis
Multi-Instance Deep Learning
Multi-Instance Learning
Multi-Task Learning
Graph Neural Networks

STAPLE: Spatio-Temporal Precursor Learning for Event Forecasting [Ning et al. SDM18]



Event, Geolocation, Time

STAPLE: Spatio-Temporal Precursor Learning for Event Forecasting [Ning et al. SDM18]



Event: Labor Protest
Location: D.C.
Time: 2015-04-01

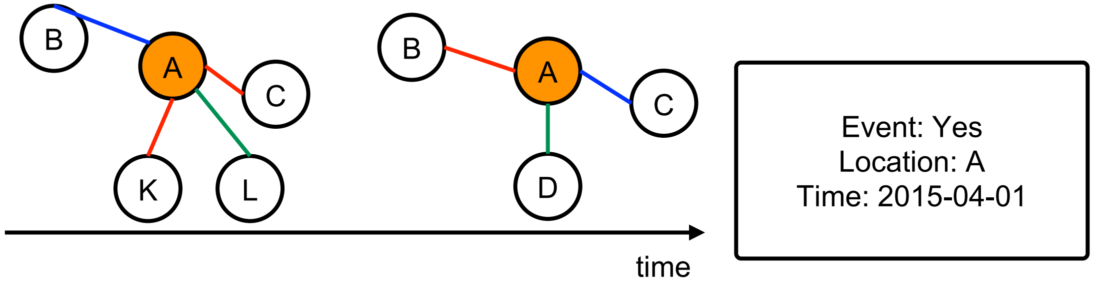
STAPLE: objective function

STAPLE: explicitly enforces pairs of cities with similar event patterns in the past to learn similar model vectors

all tasks share common features

$$\min_{\Theta} \sum_{k \in K} \Big(\underbrace{\frac{N_k}{N} \mathcal{L}(\theta^k)}_{\text{Multi-Instance Loss}} + \underbrace{\boxed{\frac{\lambda_1}{2} \sum_{i}^{N_k} \sum_{l \in \mathcal{G}_t} a_{k,l}^{t_i} \left(\theta^k - \theta^l\right)^2}}_{\text{Spatio-Temporal Constraints}} + \underbrace{\boxed{\frac{\lambda_2}{2} ||\hat{\theta} - \theta^k||_2^2}}_{\text{Global averaging}} + \underbrace{\frac{\lambda_3}{2} ||\theta^k||_2^2}_{L2} \Big)$$

STAPLE: spatio-temporal constraints

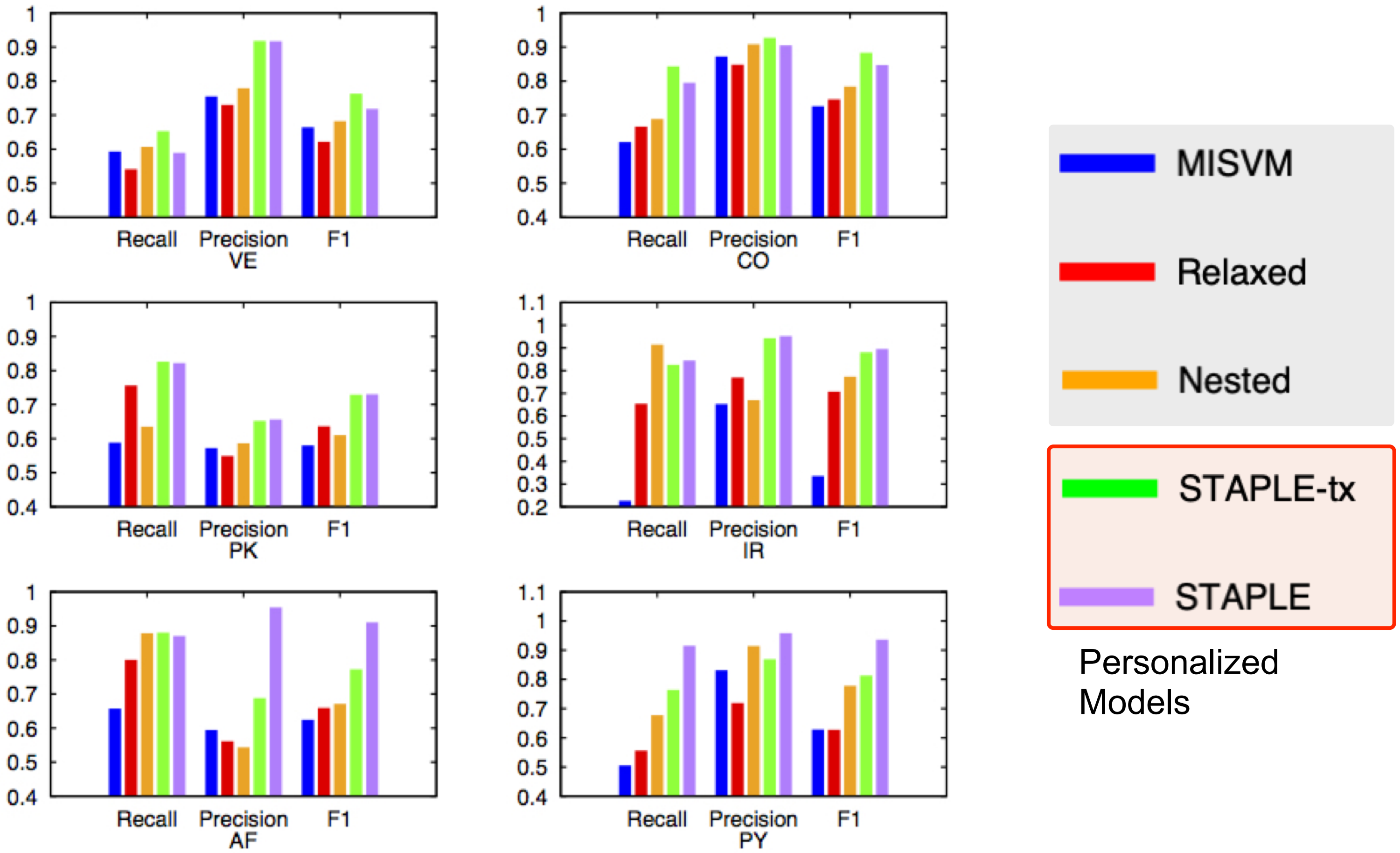


$$a^{t_i}_{k,l} = \left(\sum_c \sum_{t=t_i-H}^{t_i} \boxed{\min(E^k_t(c), E^l_t(c))}\right)' + \left(\boxed{\frac{1}{\mathrm{dist}(k,l)}}\right)'$$

Similar event patterns in the past, similar models

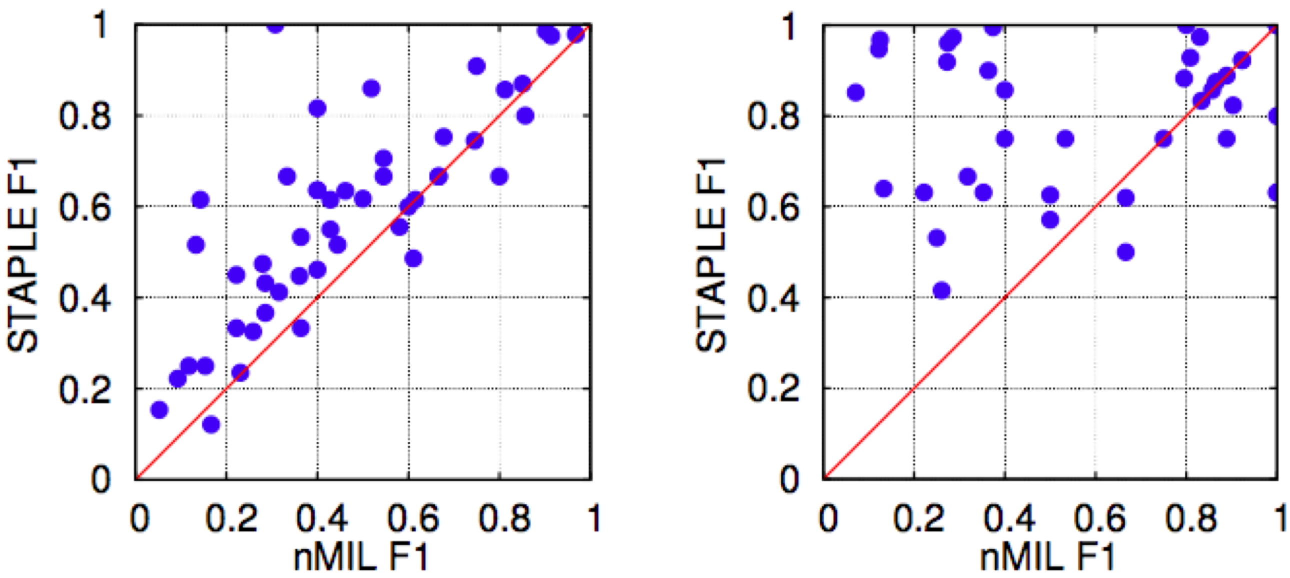
Closer geolocations, similar models

STAPLE:Event Prediction Performance



Personalized Models

City-level Prediction Performance



(a) ICEWS (b) GSR

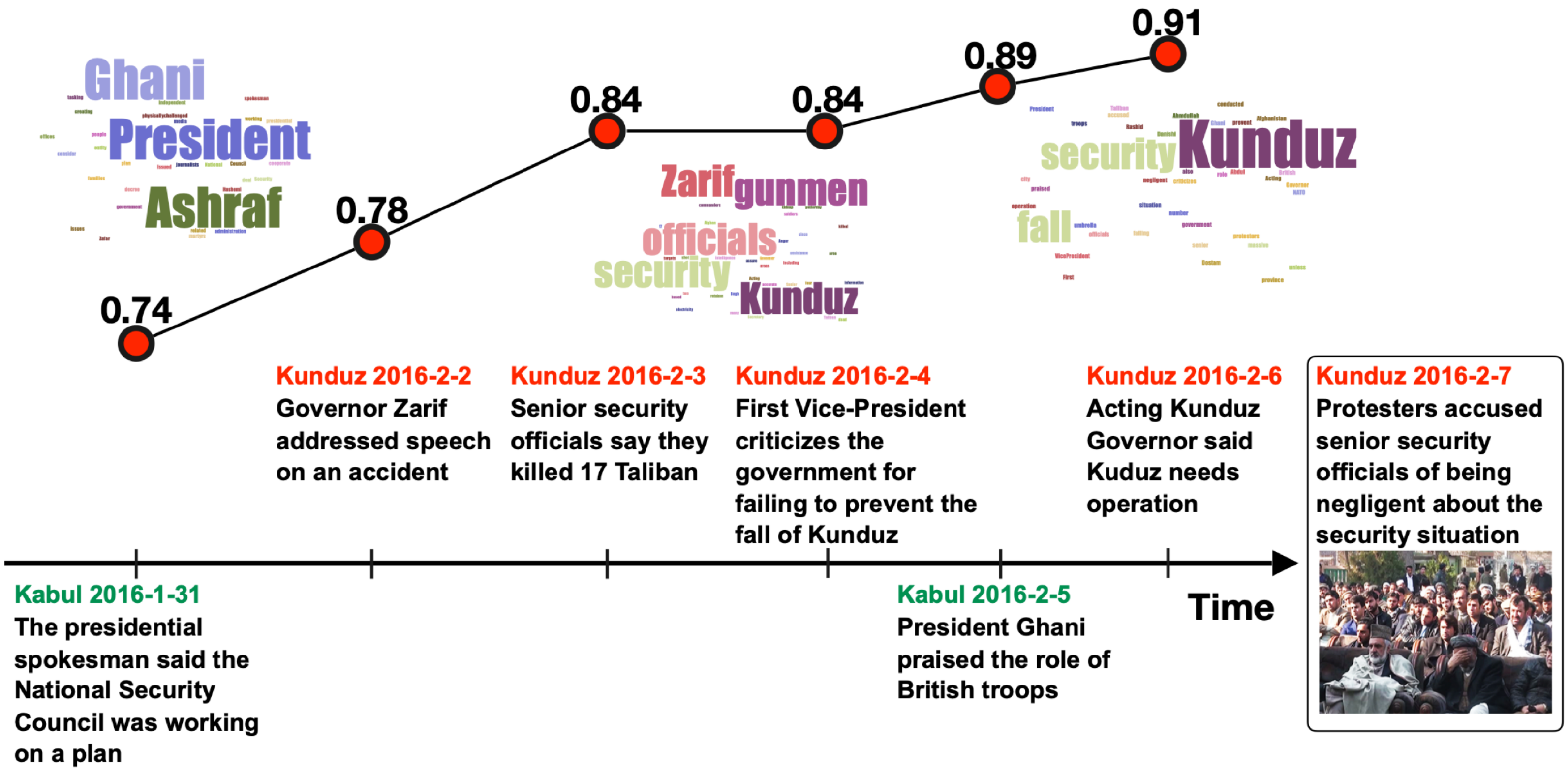
Security-related protest

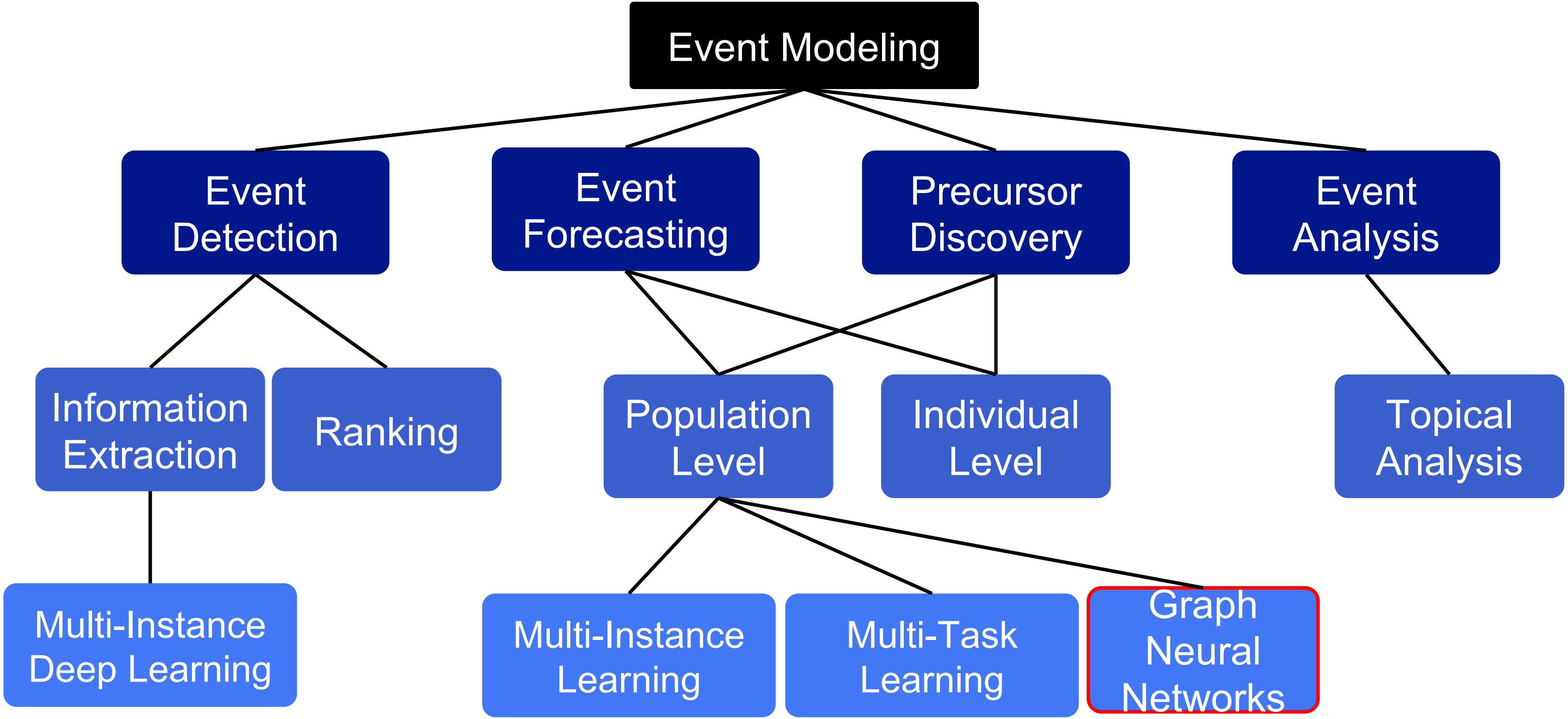
KUNDUZ RESIDENTS STAGE PROTEST AGAINST MOUNTING INSECURITY

February 7, 2016 · Afghanistan · 13 Views

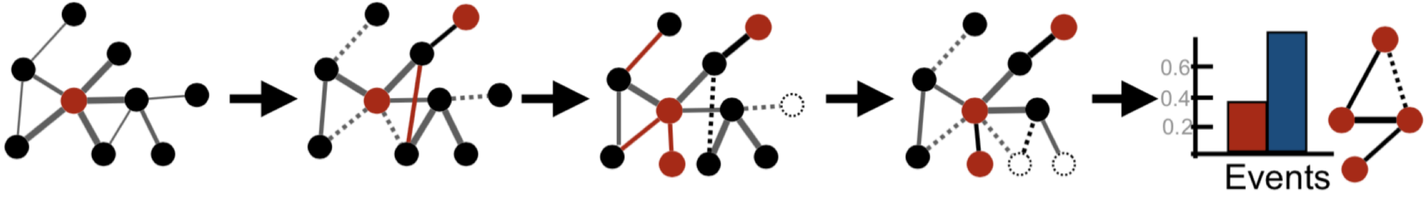
http://www.kabultribune.com/index.php/2016/02/07/kunduz-residents-stage-protest-against-mounting-insecurity/

Security-related protest - precursors




Event Modeling
Event Detection
Event Forecasting
Precursor Discovery
Event Analysis
Information Extraction
Ranking
Population Level
Individual Level
Topical Analysis
Multi-Instance Deep Learning
Multi-Instance Learning
Multi-Task Learning
Graph Neural Networks

Learning Dynamic Context Graphs for Predicting Social Events [S. Deng et al. KDD19]



- Develop a novel graph-based model for predicting events
- Design a mechanism that encodes the dynamic graph structure of words from past input documents to forecast future events.
- Propose a temporal encoding module to alleviate the problem that pre-trained semantic features usually cannot reflect contextual changes over time.

Graph Convolutional Networks

[kipf and welling ICLR17]

Main idea: Pass messages between pairs of nodes

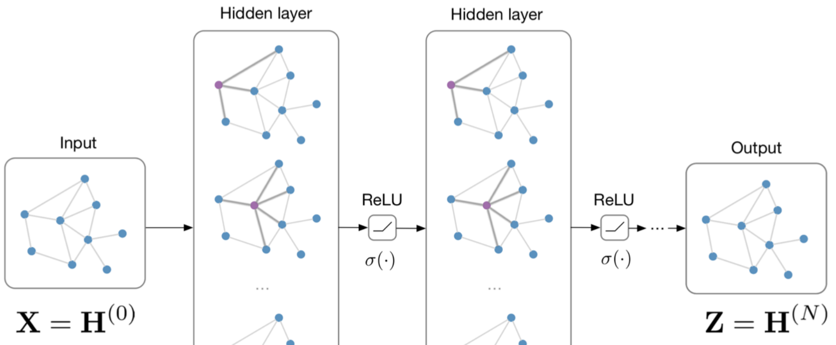
Graph: $G = (\mathcal{V}, \mathcal{E})$

$\mathcal{V}$: Set of nodes $\{v_i\}$, $|\mathcal{V}| = N$

$\mathcal{E}$: Set of edges $\{(v_i, v_j)\}$

Notation: $\mathcal{G} = (\mathbf{A}, \mathbf{X})$

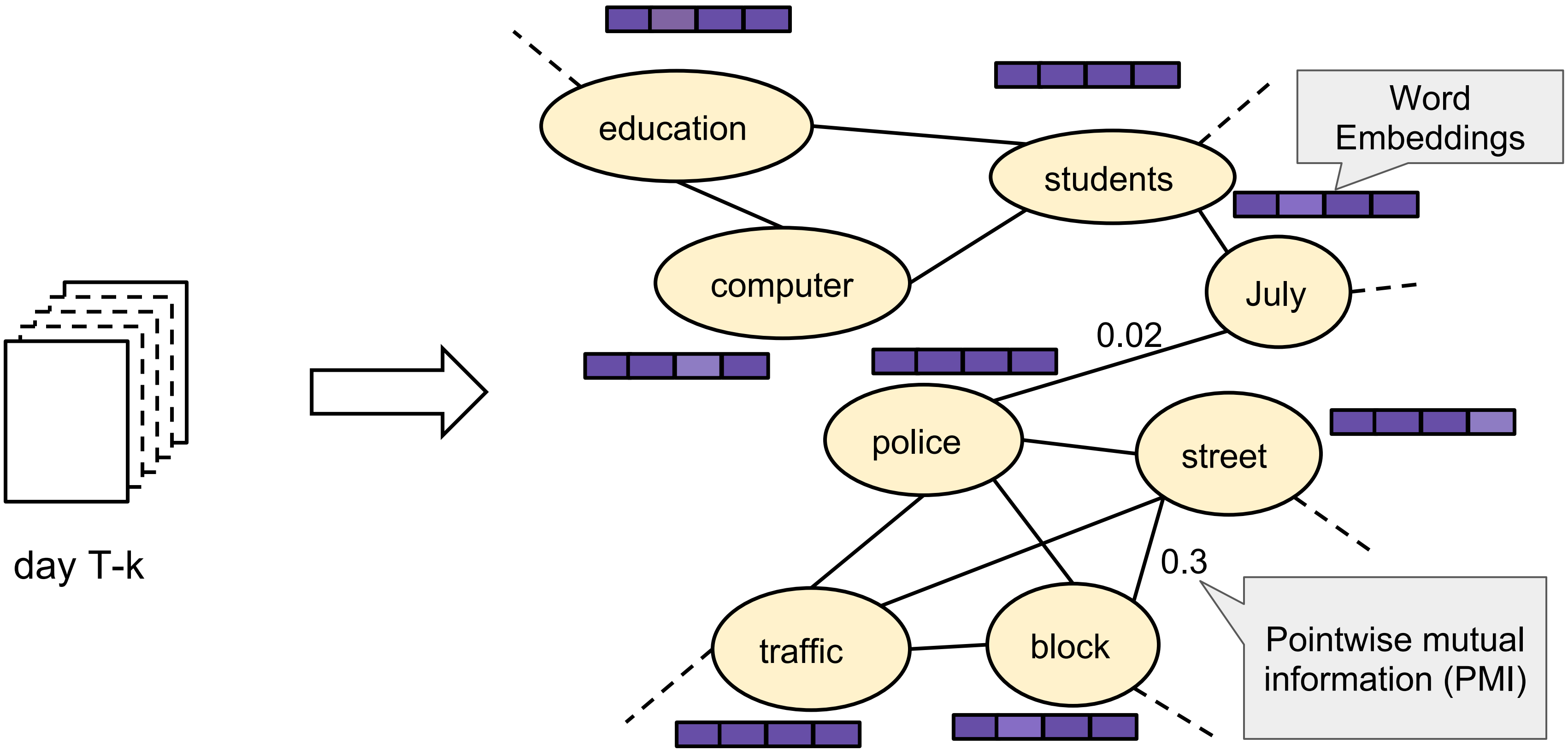
- Adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$
- Feature matrix $\mathbf{X} \in \mathbb{R}^{N \times F}$



$$\mathbf{H}^{(l+1)} = \sigma\left(\hat{\mathbf{A}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right)$$

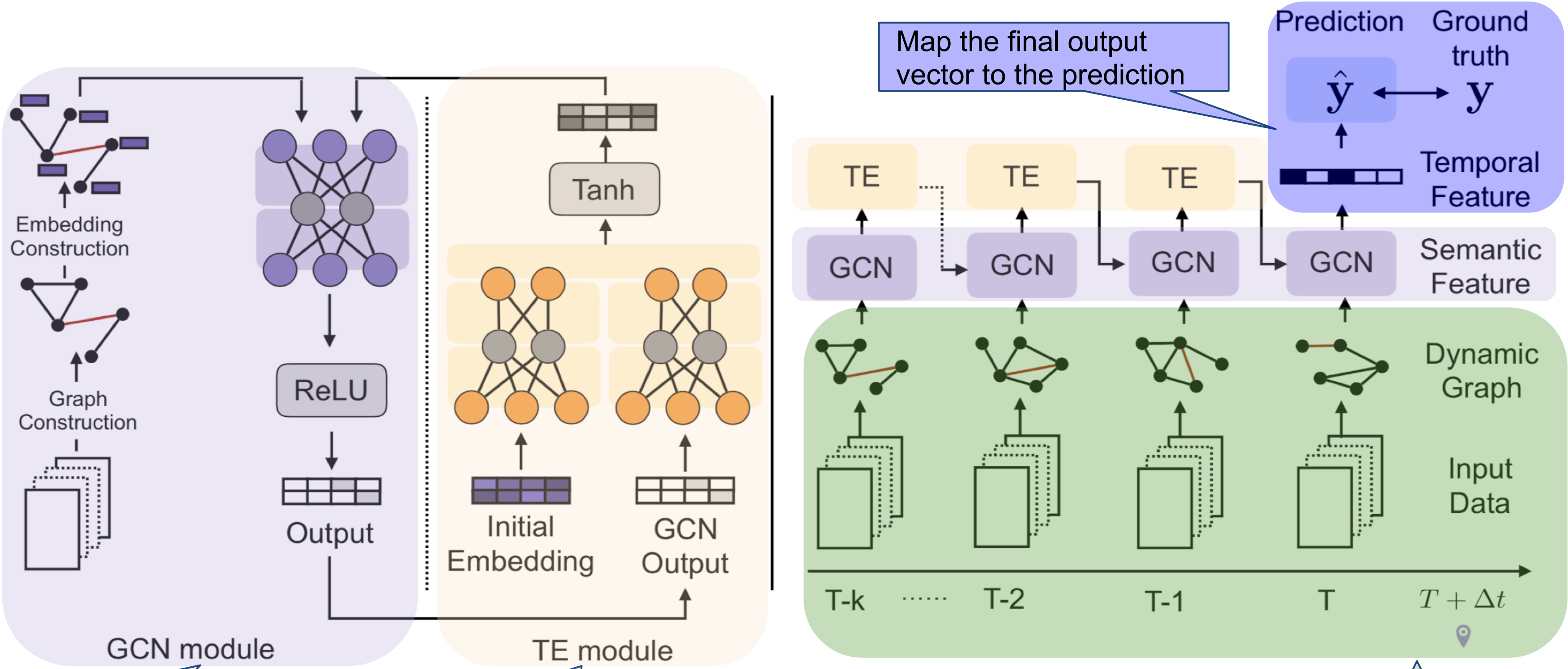
Source: https://tkipf.github.io/graph-convolutional-networks/

Encoding documents into graphs



DynamicGCN: model framework

Map the final output vector to the prediction



Learn the node representation by involving the semantic information from neighbors

Integrate both the semantic information and node embedding

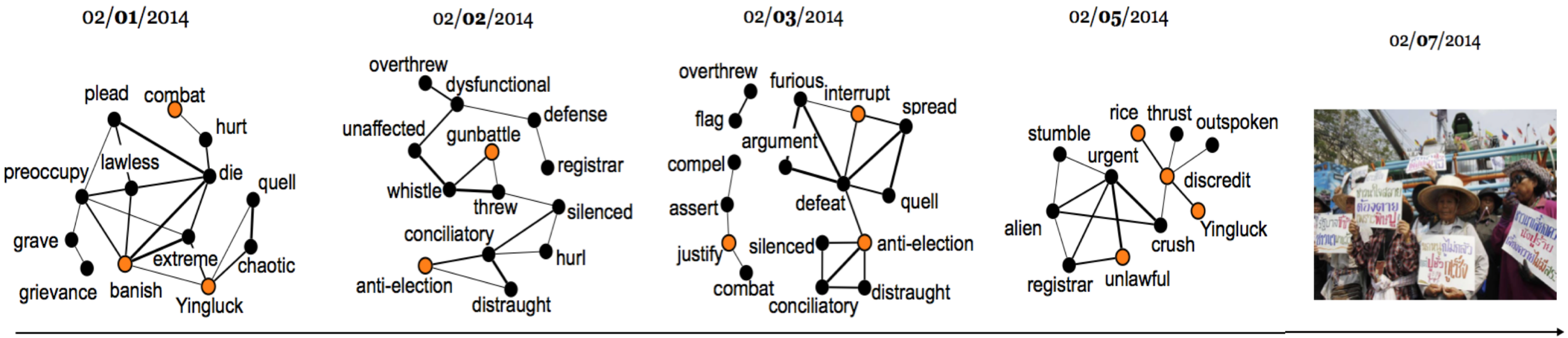
Encode data

DynamicGCN: experimental evaluation

		Thailand		Egypt		India		Russia	
		F1	Rec.	F1	Rec.	F1	Rec.	F1	Rec.
Non temporal	LR-Count	0.77	0.713	0.794	0.747	0.618	0.559	0.739	0.721
	LR-word	0.715	0.634	0.78	0.751	0.543	0.433	0.705	0.689
	LR-NGram	0.7293	0.6535	0.761	0.7039	0.552	0.441	0.714	0.714
	GCN	0.761	0.758	0.849	0.816	0.653	0.627	0.784	**0.826**
Temporal	nMIL	0.73	0.661	0.723	0.797	0.628	**0.719**	0.76	0.769
	GCN+GRU	<u>0.782</u>	0.769	0.85	0.825	<u>0.655</u>	0.621	0.787	<u>0.809</u>
	GCN+LSTM	0.781	<u>0.77</u>	<u>0.851</u>	<u>0.827</u>	0.649	0.614	0.786	0.791
	GCN+RNN	0.757	0.755	<u>0.851</u>	0.82	0.642	0.602	<u>0.787</u>	<u>0.809</u>
	Ours	**0.797**	**0.773**	**0.862**	**0.829**	**0.669**	<u>0.627</u>	**0.804**	0.799

Data:
Integrated Crisis Early Warning System (ICEWS) Dataverse

DynamicGCN: a case study



Violence grips Thai capital on eve of vote called by Yingluck.

Thailand started voting. Voters blocked by anti-election groups squared off with scuffles and hurled objects.

Election Commission asked the national police chief to maintain law and order. Thai Protests Disrupt Vote.

Yingluck's former commerce ministers were suspected of being involved in improper rice deals.

For more details please attend the paper presentation:

Tuesday (Aug. 6) at 1:30-3:00pm, Summit 4, Ground Level, Egan Center

Conclusion and Future Directions – Precursor Identification

- Representation Learning and Deep Learning
  to automatically encode raw input and learn hidden features
- Multi-Instance Learning
  Identify key characteristics in semi-supervised event modeling
- Multi-Task Learning
  to infer relationships across different tasks (locations)

Future directions

- *Data integration for multiple sources*
- *Learning hierarchies of spatial precursors*
- *Semantic encoding and optimization*

References

- Yue Ning, Sathappan Muthiah, Huzefa Rangwala, Naren Ramakrishnan. "Modeling Precursors for Event Forecasting via Nested Multi-Instance Learning." in Proceedings of the 22nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD’16), San Francisco, CA, USA. August 13-17, 2016.
- Wei Wang, Yue Ning, Huzefa Rangwala, Naren Ramakrishnan. A Multiple Instance Learning Framework for Identifying Key Sentences and Detecting Events.In Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM’16). Indianapolis, IN, USA. Oct. 24-28, 2016.
- Yue Ning, Rongrong Tao, Chandan K. Reddy, Huzefa Rangwala, James C. Starz, Naren Ramakrishnan. "STAPLE: Spatio-Temporal Precursor Learning for Event Forecasting" In Proceedings of the 18th SIAM International Conference on Data Mining (SDM’18). San Diego, CA, USA. May 3-5, 2018
- Songgaojun Deng, Huzefa Rangwala, Yue Ning. "Learning Dynamic Context Graphs for Predicting Social Events" in Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'19). Anchorage, Alaska USA. August 4-8, 2019

Coffee Break

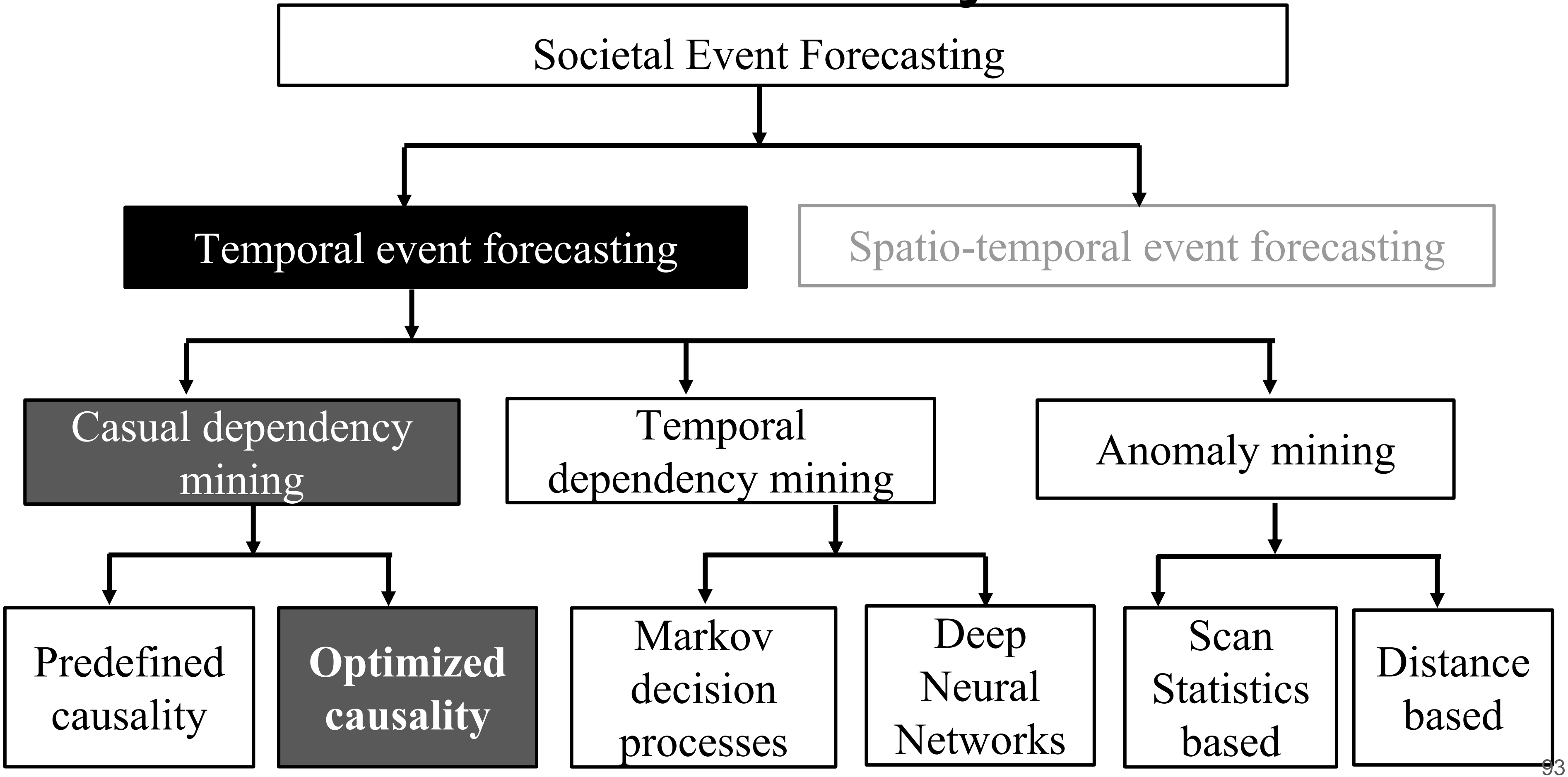
30 Minutes

Part 2: Temporal Event Forecasting

Feng Chen (University of Texas at Dallas)

The University of Texas at Dallas

Taxonomy



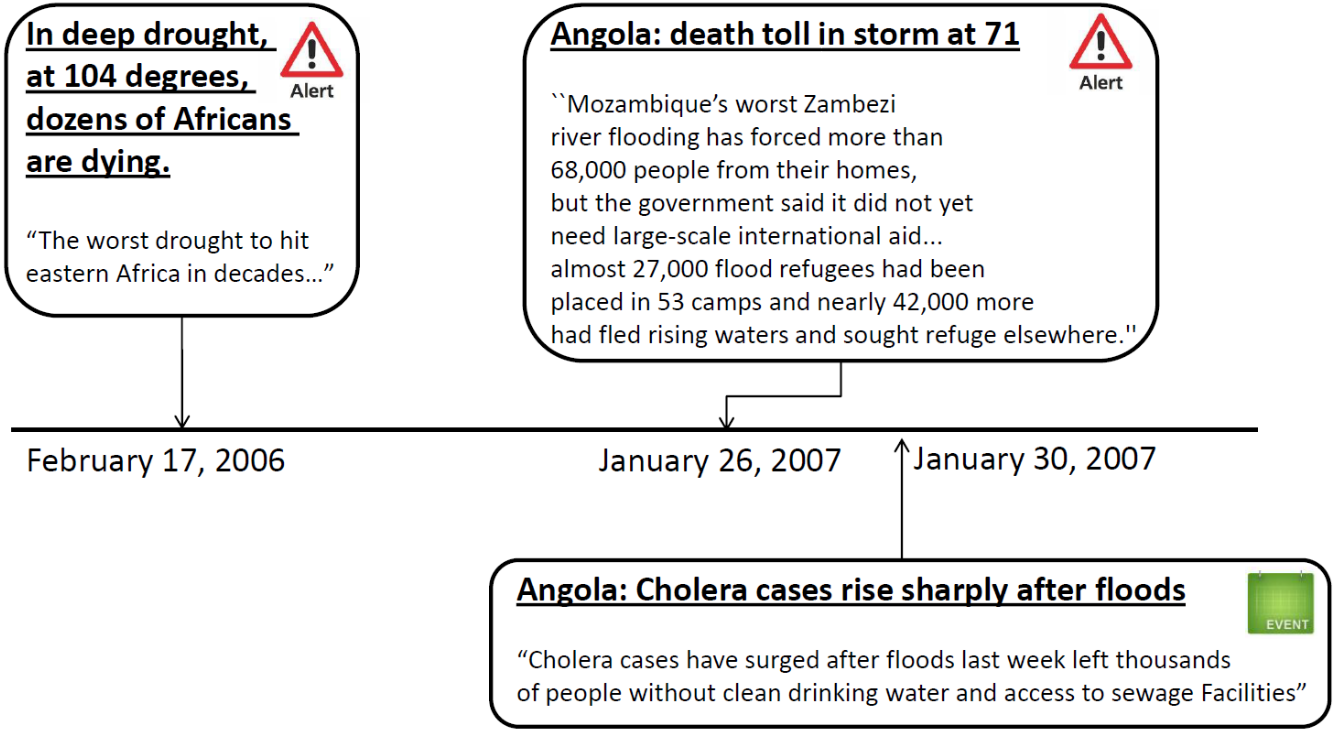
Societal Event Forecasting
Temporal event forecasting
Spatio-temporal event forecasting
Casual dependency mining
Temporal dependency mining
Anomaly mining
Predefined causality
Optimized causality
Markov decision processes
Deep Neural Networks
Scan Statistics based
Distance based

Prediction Learning

Mining the Web to Predict Future Events

(Radinsky and Horvitz, WSDM'13)

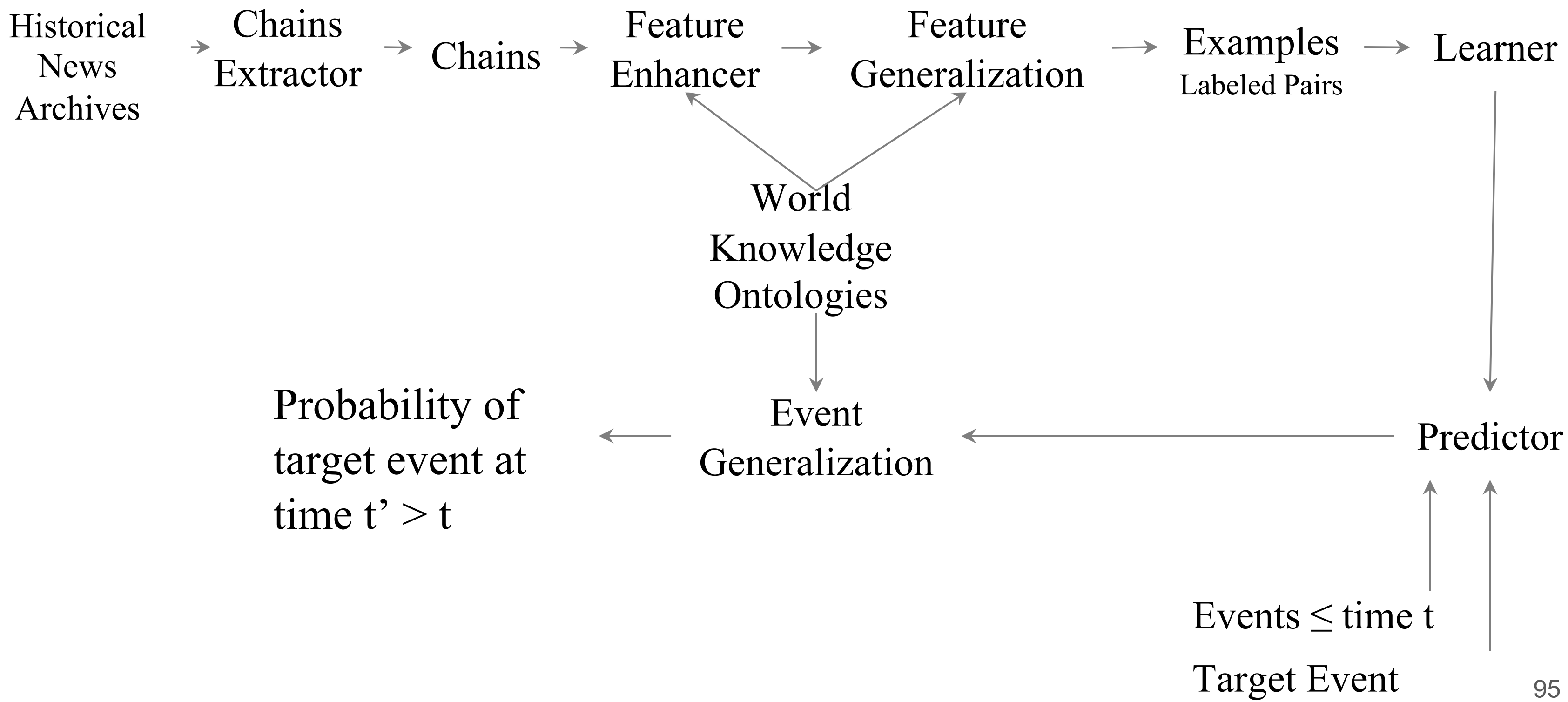
Goal: Predict future events using historical news and web ontologies.



Mining the Web to Predict Future Events

(Radinsky and Horvitz, WSDM'13)

Prediction Learning

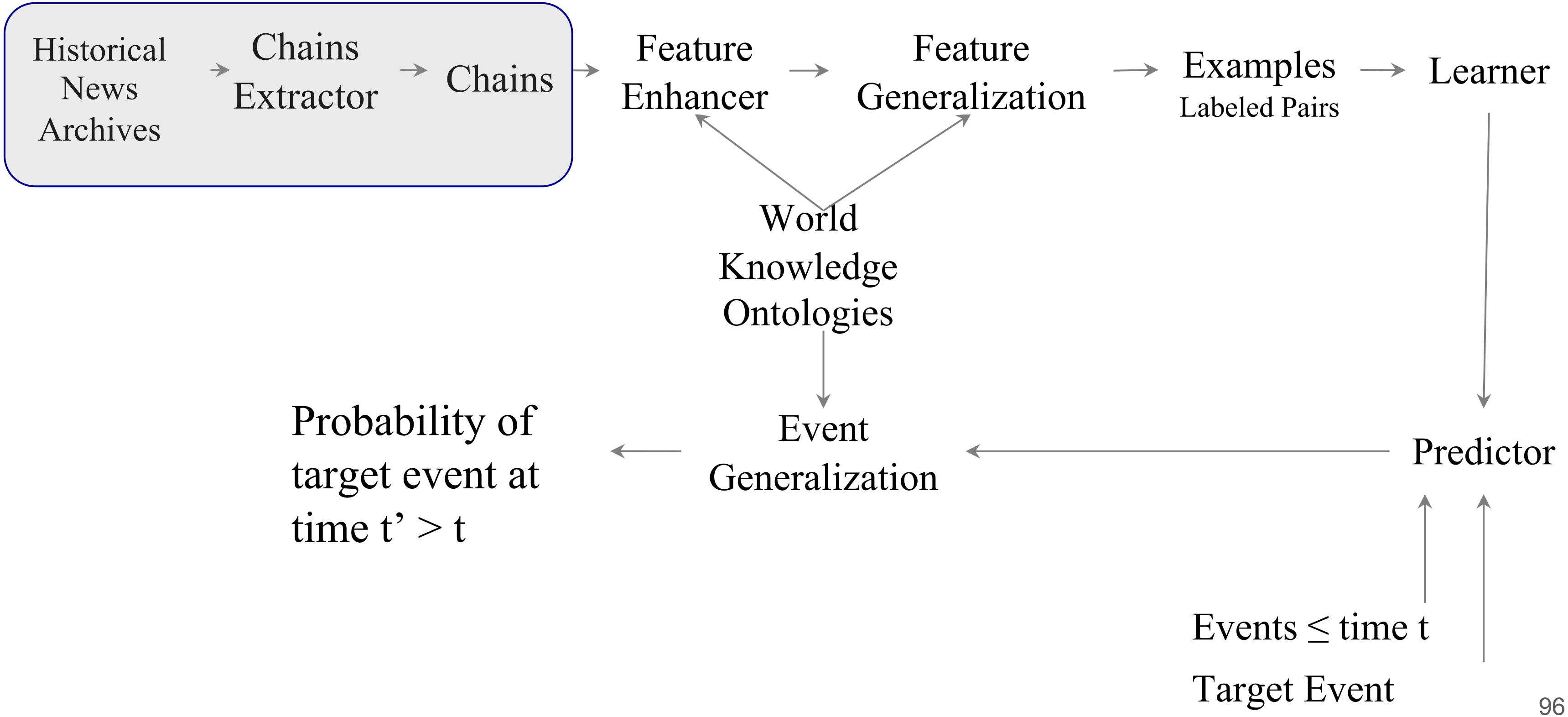


Mining the Web to Predict Future Events

(Radinsky and Horvitz, WSDM'13)

Learning

Prediction



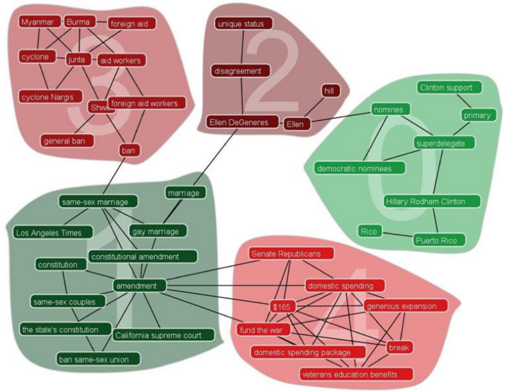
Event Chains (Storylines)

Jan 16, 1992	Jury in Shooting by Officer Hears Conflicting Accounts
Feb 11, 1992	Closing Arguments Conflict on Killing by Teaneck Officer
Feb 12, 1992	Officer Acquitted in Teaneck Killing
Feb 13, 1992	Acquitted Officer Expresses Only Relief, Not Joy
Feb 16, 1992	250 March in Rain to Protest Teaneck Verdict

Event Chains

Cluster documents with similar text

(using bag of words similarity)



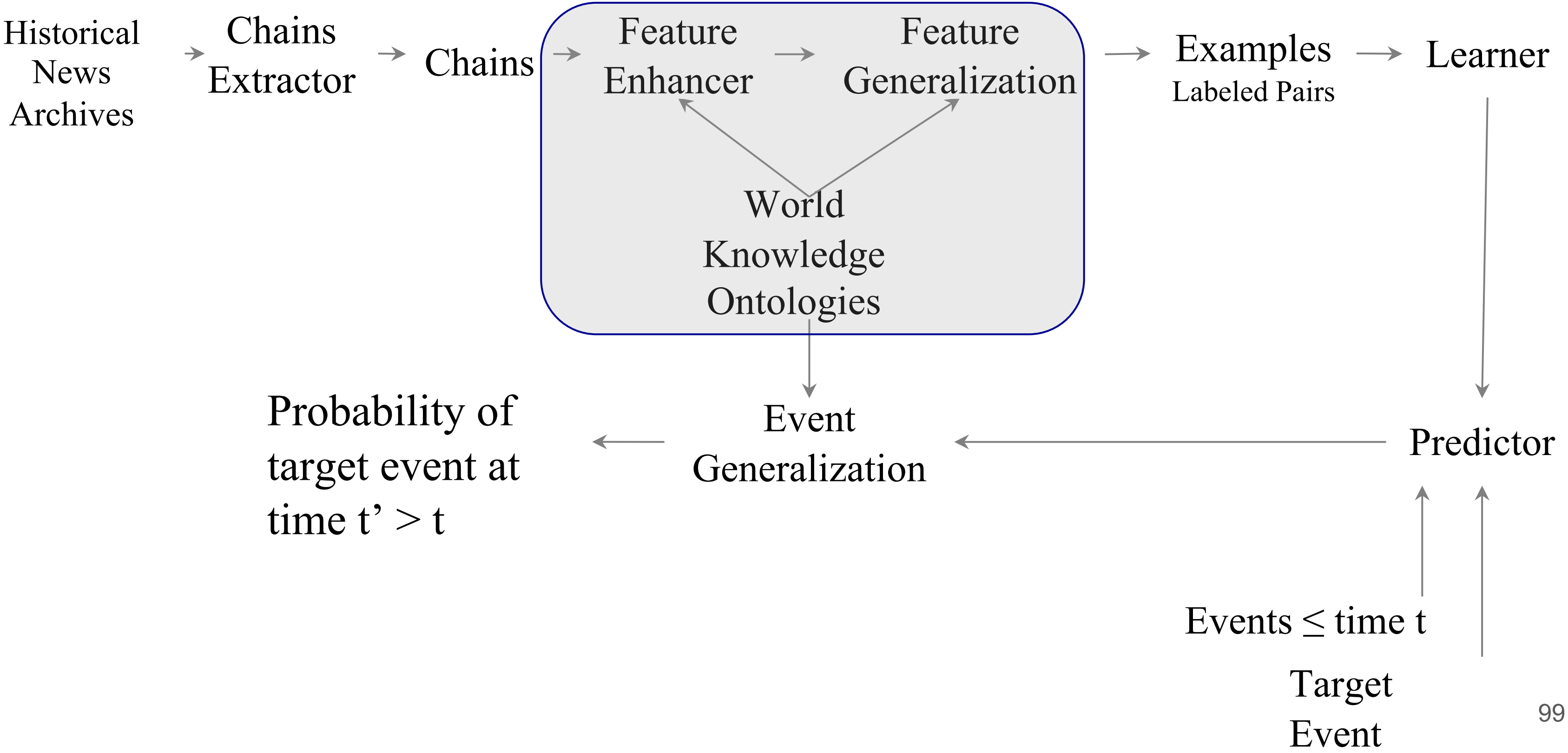
Improve Precision:
Greedily optimize Story Entropy (entropy in its entities)

to grow “slowly”

Prediction Learning

Mining the Web to Predict Future Events

(Radinsky and Horvitz, WSDM'13)



04/14/2011

BBC
News | Sport | Weather | Travel | Future | Autos

NEWS LATIN AMERICA & CARIBBEAN

Home | UK | Africa | Asia | Europe | Latin America | Mid-East | US & Canada | Business | Health | Sci/Enviro

14 April 2011 Last updated at 10:02 GMT

Share

Cuba faces its worst drought for 50 years

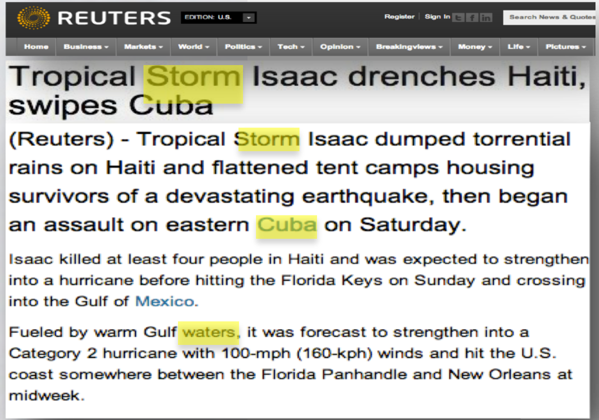
Cuba is facing its worst drought in half a century, with tens of thousands of families almost entirely reliant on water trucks for essential supplies.

The drought started two years ago, and reservoirs are now down to a fifth of their normal levels.

The government is providing road deliveries of water to more than 100,000 people in the worst affected areas of the capital, Havana.

The BBC's Michael Voss asked people in Havana how they were coping

8/25/12



REUTERS EDITION: U.S.
Register | Sign In
Search News & Quotes

Home | Business | Markets | World | Politics | Tech | Opinion | Breakingviews | Money | Life | Pictures

Tropical Storm Isaac drenches Haiti, swipes Cuba

(Reuters) - Tropical Storm Isaac dumped torrential rains on Haiti and flattened tent camps housing survivors of a devastating earthquake, then began an assault on eastern Cuba on Saturday.

Isaac killed at least four people in Haiti and was expected to strengthen into a hurricane before hitting the Florida Keys on Sunday and crossing into the Gulf of Mexico.

Fueled by warm Gulf waters, it was forecast to strengthen into a Category 2 hurricane with 100-mph (160-kph) winds and hit the U.S. coast somewhere between the Florida Panhandle and New Orleans at midweek.

01/19/13

WORLDNEWS on NBCNEWS.com

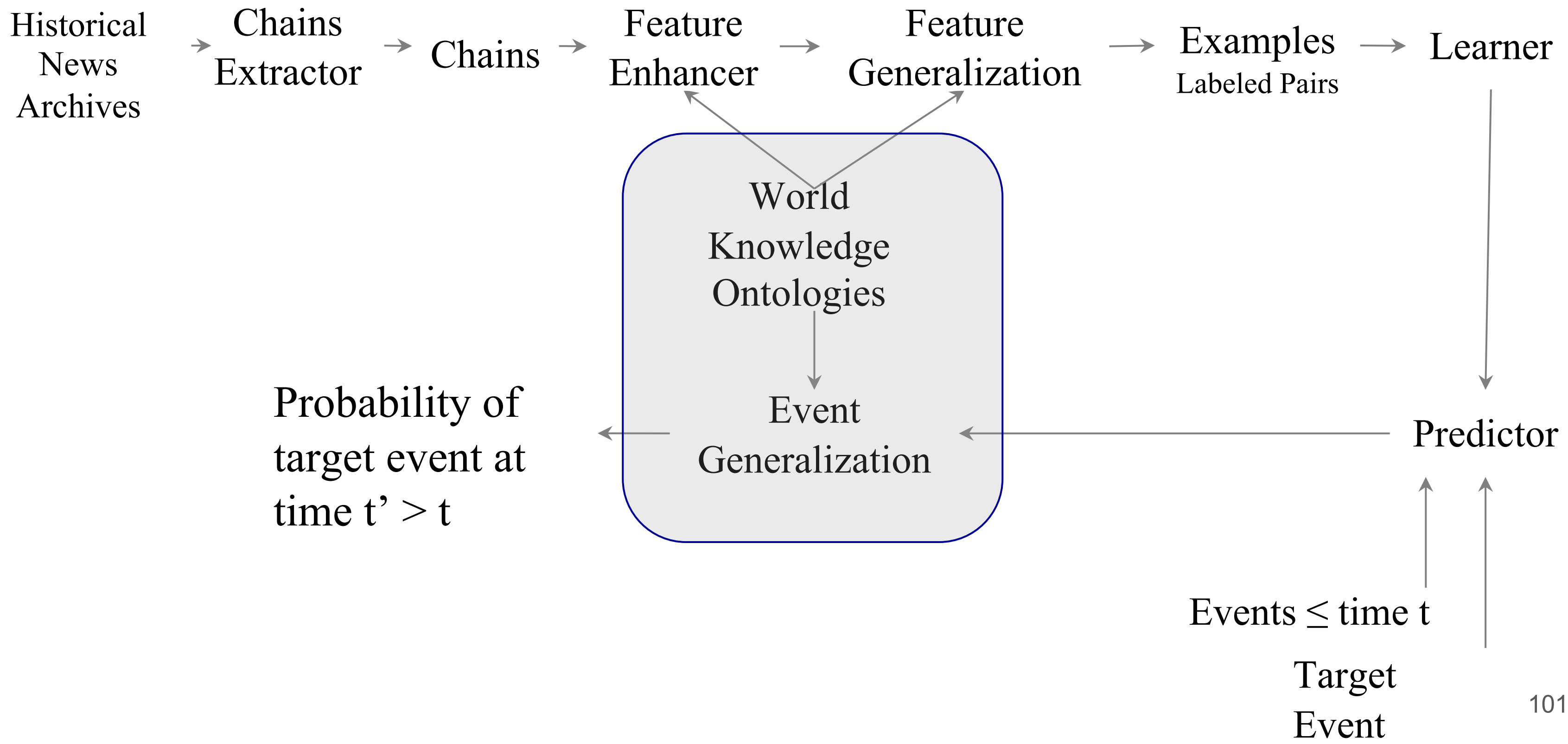
19 Jan 2013 10:46am, EST

After a century without the disease, Cuba fights to contain cholera

Mining the Web to Predict Future Events

(Radinsky and Horvitz, WSDM'13)

Prediction Learning



$$P(\mathit{Cholera\ in\ Havana} \mid \mathit{Cuba, flood})$$

Never appeared in the news archive…

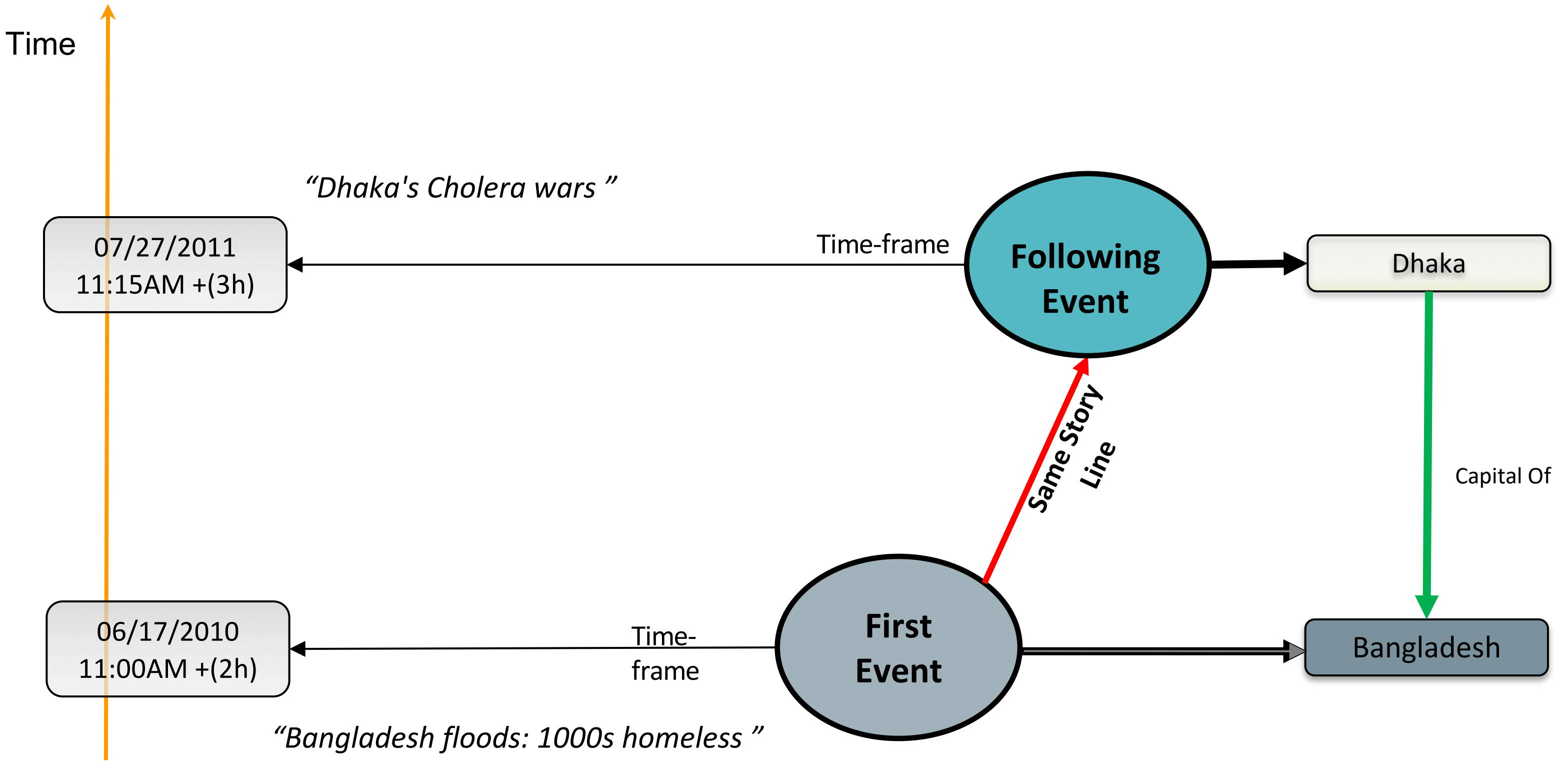
$$
P\left(\begin{array}{l}
\mathit{Cholera\ in\ Havana} \,|\, \\
\mathit{Cuba}, \\
\mathit{AreaTotal}: 109884.0, \\
\mathit{PopulationDensity}: 102.3, \\
\text{GdpNominalPerCapita}: 5100.0, \\
\text{PercentWater}:\ \text{negligible}
\end{array}\right)
$$

$$P\left(\begin{array}{l} Cholera\ in\ Havana \\ Cuba, \\ Social\ States, \\ Island\ Countries \end{array}\right.\left|\begin{array}{l} \\ \\ \\ \\ \end{array}\right)$$

$$P(Cholera\ in\ capital\ of\ [Country] \mid [Country],\ flood)$$

Categories: Cuba | Caribbean countries | Communist states | Eastern Bloc | Former Spanish colonies | Gulf of Mexico | Island countries | Member states of the United Nations | Republics | Single-party states | Socialist states | Spanish-speaking countries | States and territories established in 1902

Abstraction Process



$$P(\text{Cholera in capital of } [\text{Country}] \mid [\text{Country}], \text{flood})$$

Experimental Methodology

- 22 years of NYT (1986–2007)
- Divide to learning and prediction:
  - Learn 1986- 1997
  - Predict 1998-2007
- During prediction, only the first event in the story line (without words containing the prediction target) is given to the predictor
- Predict the last event in the storyline

Algorithm Component Analysis

	General Predictions		Death		Disease		Riots	
	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.	Prec.	Rec.
News alone	19%	100%	80%	59%	44%	34%	88%	38%
News + factual features	19%	100%	81%	62%	52%	31%	87%	42%
News + generalization	21%	100%	81%	67%	53%	28%	88%	42%
Full model	24%	100%	83%	81%	61%	33%	91%	51%

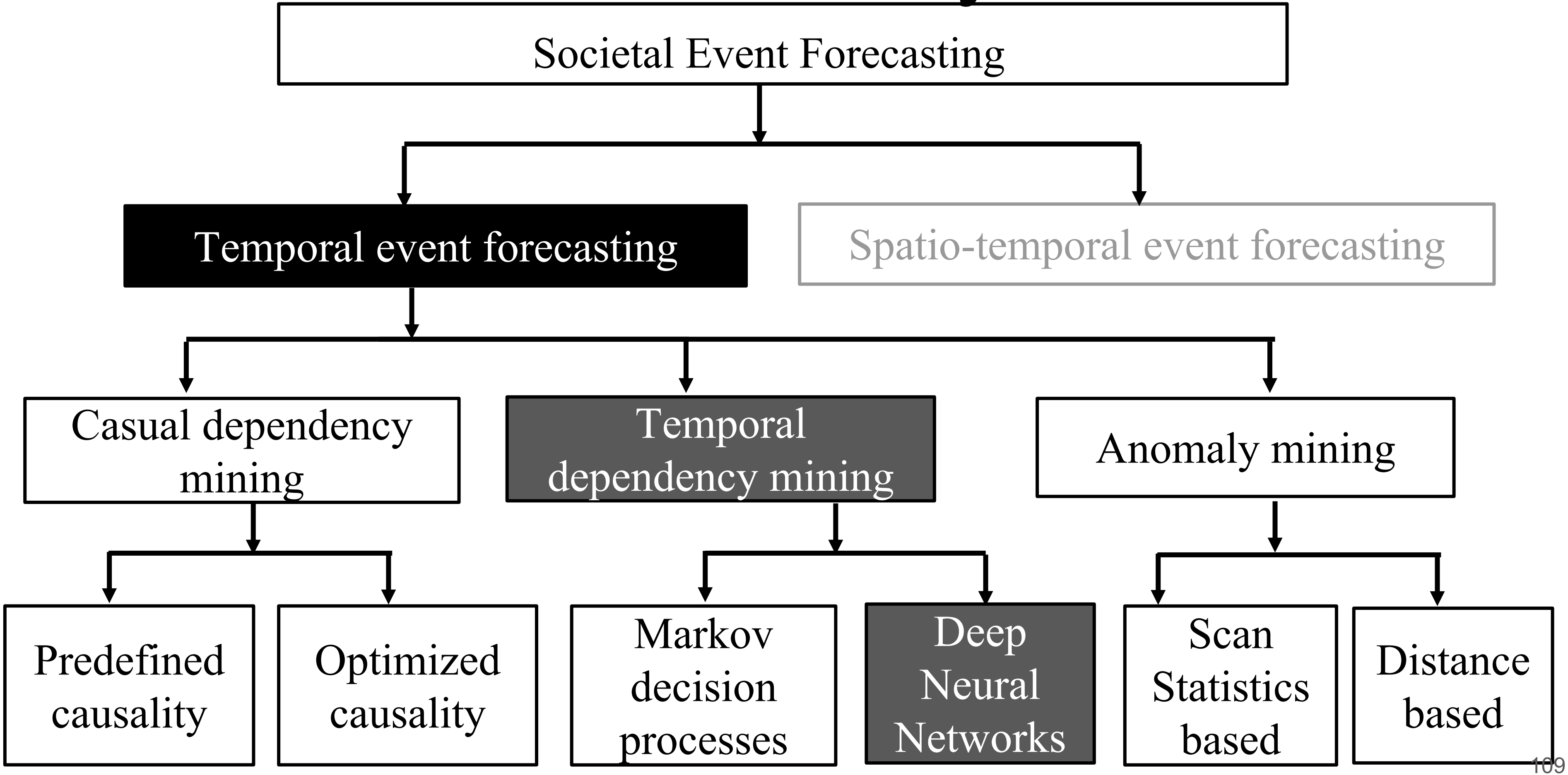
Both factual features and generalization are essential for forecasting.

Alert Time (in days)

General Predictions		Death		Disease Outbreak		Riots	
Med.	Avg.	Med.	Avg.	Med.	Avg.	Med.	Avg.
9	21	8	41	12	273	18	30

Most alerts are given in timely manner providing time for action

Taxonomy


Societal Event Forecasting
Temporal event forecasting
Spatio-temporal event forecasting
Casual dependency mining
Temporal dependency mining
Anomaly mining
Predefined causality
Optimized causality
Markov decision processes
Deep Neural Networks
Scan Statistics based
Distance based

Temporal Dependency based Event Forecasting – Problem Definition

$\mathbb{E}$ is a set of events;
$\mathbb{T}$ is a discrete representation of time

Forecasting function

$$f(e_1, \cdots, e_M) \rightarrow (e'_1, \cdots, e'_D)$$

, s.t.:

$e_1, \cdots, e_M$ occurred at time $t \in \mathbb{T}$

$e'_1, \cdots, e'_D$ occurred at time $t' \in \mathbb{T}, t' > t$

Temporal Dependency based Event Forecasting – Problem Definition

$\mathbb{E}$ is a set of events;
$\mathbb{T}$ is a discrete representation of time

Forecasting function

$$f(e_1, \cdots, e_M) \rightarrow (e'_1, \cdots, e'_D)$$

Instead of modeling the forecasting function $f(e_1, \cdots, e_M)$ based on a causal relational graph, this approach aims to model the function based on a deep neural network.

A Compositional Neural Network Model for Event Forecasting

(Granroth-Wilding and Clark, AAAI'16)

- Training Phase:
  - INPUT: A training collection of news articles
  - OUTPUT: a trained compositional neural network model
  - Step 1: Unsupervised event chain learning
  - Step 2: Train a compositional neural network model to measure the coherence score between a cause event and a candidate next event

Step 1: Unsupervised Event Chain Learning

Text: Robbers made a big score, fleeing after stealing more than $300,000 from Wells Fargo armored-truck guards who were servicing the track's ATMs. , the Police Department said. The two Wells Fargo guards reported they were starting to put money in the clubhouse ATM when a man with a gun approached and ordered them to lie down. . .

Step 1: Unsupervised Event Chain Learning

Text: Robbers made a big score, fleeing after stealing more than $300,000 from Wells Fargo armored-truck guards who were servicing the track's ATMs. , the Police Department said. The two Wells Fargo guards reported they were starting to put money in the clubhouse ATM when a man with a gun approached and ordered them to lie down. . .

Entities: {Wells Fargo armored-truck guards,
The two Wells Fargo guards, they, . . .}

Step 1: Unsupervised Event Chain Learning

Text: Robbers made a big score, fleeing after stealing more than $300,000 from Wells Fargo armored-truck guards who were servicing the track's ATMs, the Police Department said. The two Wells Fargo guard reported they were starting to put money in clubhouse ATM when a man with a gun approached and ordered them to lie down .

Entities: {Wells Fargo armored-truck guards, The two Wells Fargo guards, they, . . .}

Predicates: service, report, put, lie+down.

Arguments: ATM, money, in clubhouse,

Step 1: Unsupervised Event Chain Learning

Text: Robbers made a big score, fleeing after stealing more than $300,000 from Wells Fargo armored-truck guards who were servicing the track's ATMs, the Police Department said. The two Wells Fargo guards reported they were starting to put money in clubhouse ATM when a man with a gun approached and ordered them to lie down .

Entities mentions: {Wells Fargo armored-truck guards,
The two Wells Fargo guards, they, . . .}

Predicates: service, report, put, lie+down.

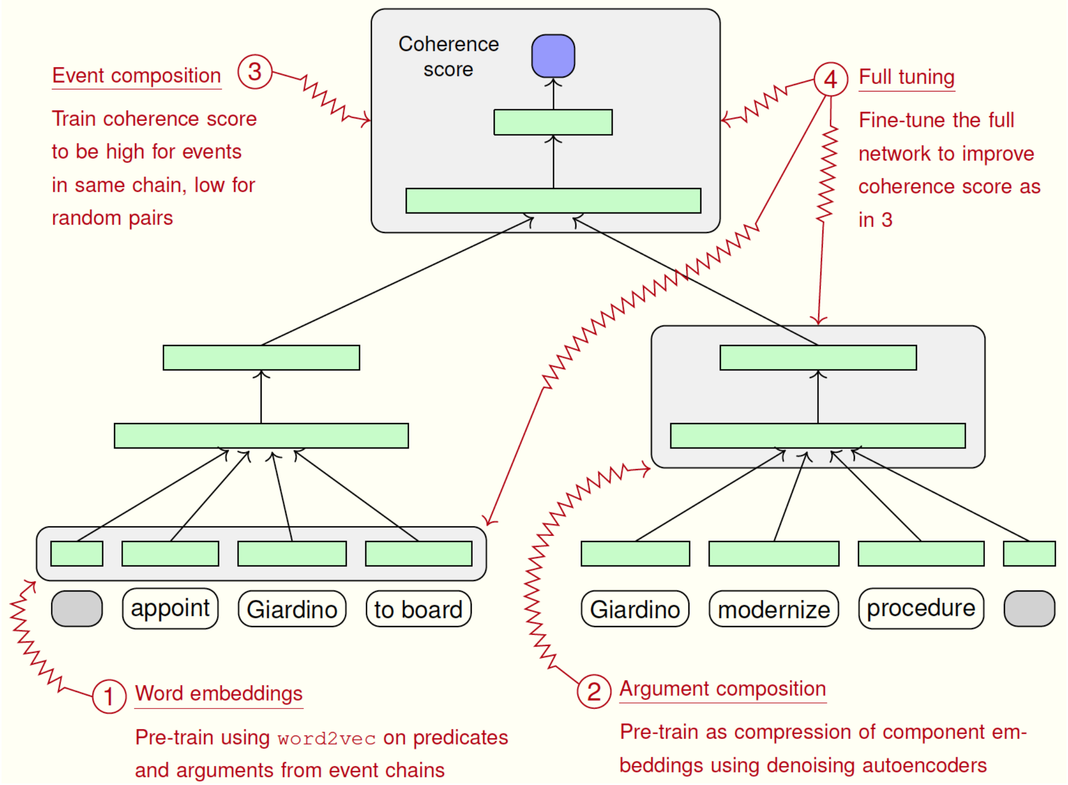
Arguments: ATM, money, in clubhouse,

Event chain: service(x0, ATMs), report(x0),
put(x0, money, in clubhouse), lie+down(x0), . . .

Step 2: Compositional Neural Network Model Training

- **Word Embeddings**
  - Represent predicates and arguments as vectors
- **Argument composition**
  - Compose embeddings into event vector
- **Event Composition**
  - Predict whether two event vectors come from the same chain

Step 2: Compositional Neural Network Model Training

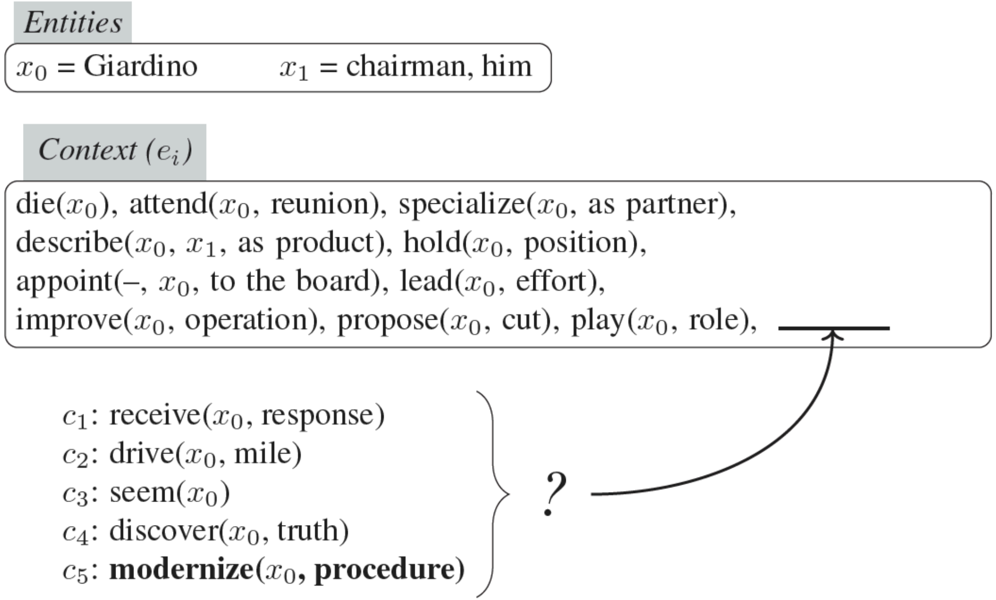


A Compositional Neural Network Model for Event Forecasting

- Testing Phase:
  - INPUT:
    - A testing collection of news articles dated at the current time.
    - A trained compositional neural network model that measures the coherence score between two events.
  - OUPUT:
    - The next candidate event.
  - Step 1: Extraction of the occurred events.
  - Step 2: Ranking of candidate events based on their coherence scores to the occurred events.

A Compositional Neural Network Model – Experiments

- Empirical validations for the multiple choice narrative cloze (MCNC) prediction task



A Compositional Neural Network Model – Experiments

- Empirical validations for the multiple choice narrative cloze (MCNC) prediction task

System	Accuracy (%)
Chance baseline	20.00
C&J08	30.52
BIGRAM	29.67
DIST-VECS	27.94
MIKOLOV-VERB	24.57
MIKOLOV-VERB+ARG	28.97
WORD2VEC-PRED	40.17
WORD2VEC-PRED+ARG	42.23
EVENT-COMP	**49.57**

A Contextual Hierarchical LSTM for Event Forecasting (Hu et al., AAAI'17)

$\mathbb{E}$ is a set of events, in which each event is denoted by its description text (e.g., news headline) which is a sequence of words. For a given $e_i \in \mathbb{E}$,

$$e_i = \left(w_{i,1}, w_{i,2}, \cdots, w_{i,N_i}\right).$$

$\mathbb{T}$ is a discrete representation of time

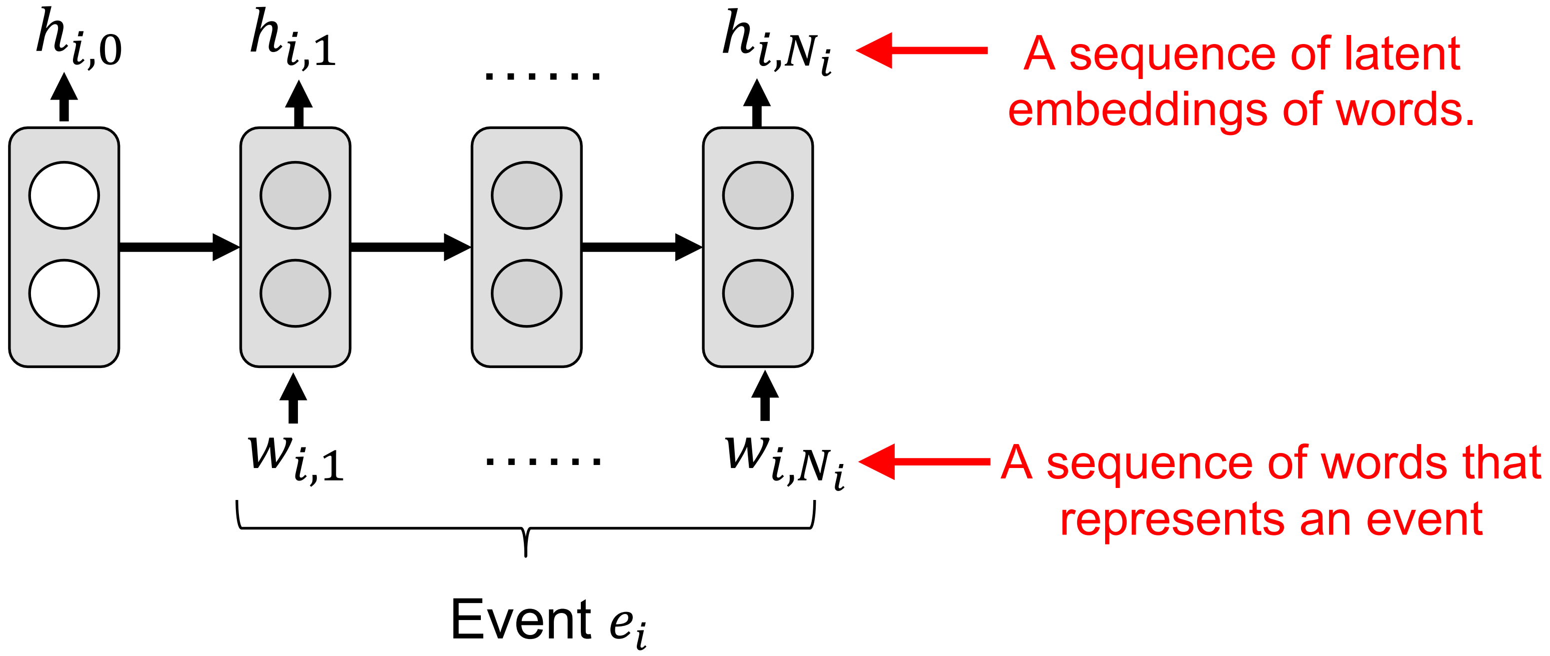
Forecasting function

$$f(e_1, \cdots, e_M) \rightarrow (e'_1, \cdots, e'_D)$$

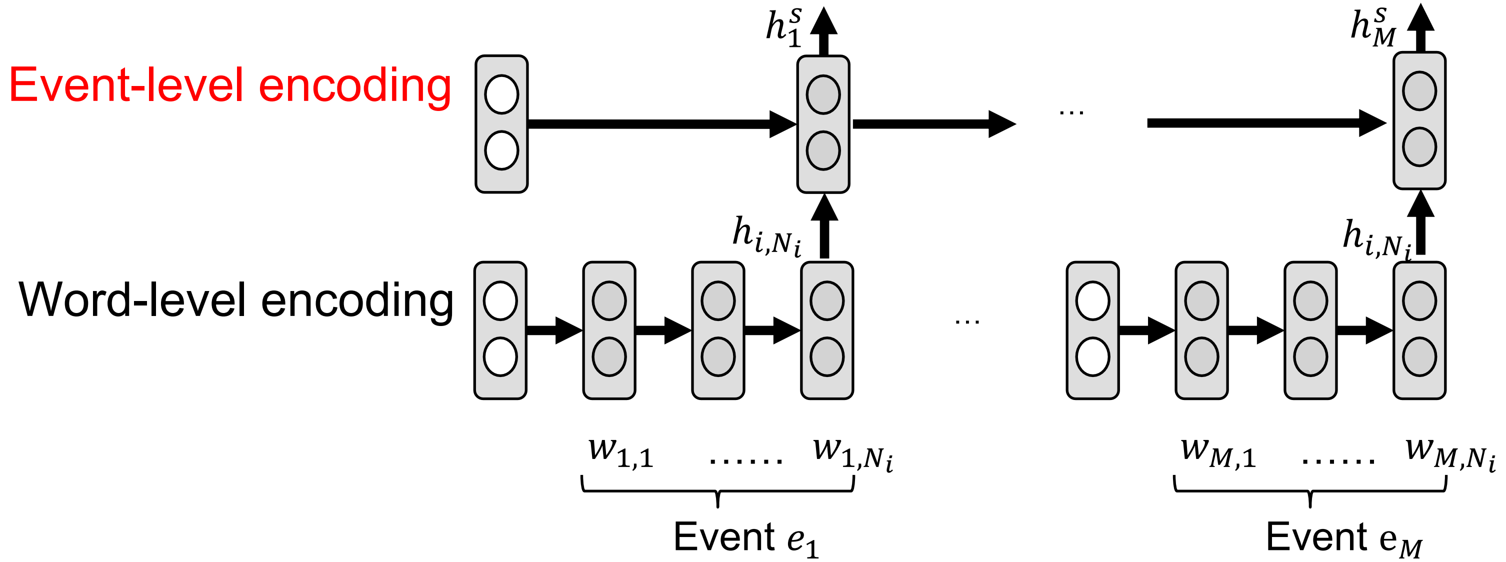
A Contextual Hierarchical LSTM for Event Forecasting

- The proposed contextual hierarchical LSTM (CH-LSTM) model has two main components:
  - Part 1: Word-level LSTM encoding
  - Part 2: Event-level LSTM encoding
  - Part 3: Next event LSTM prediction (decoding)

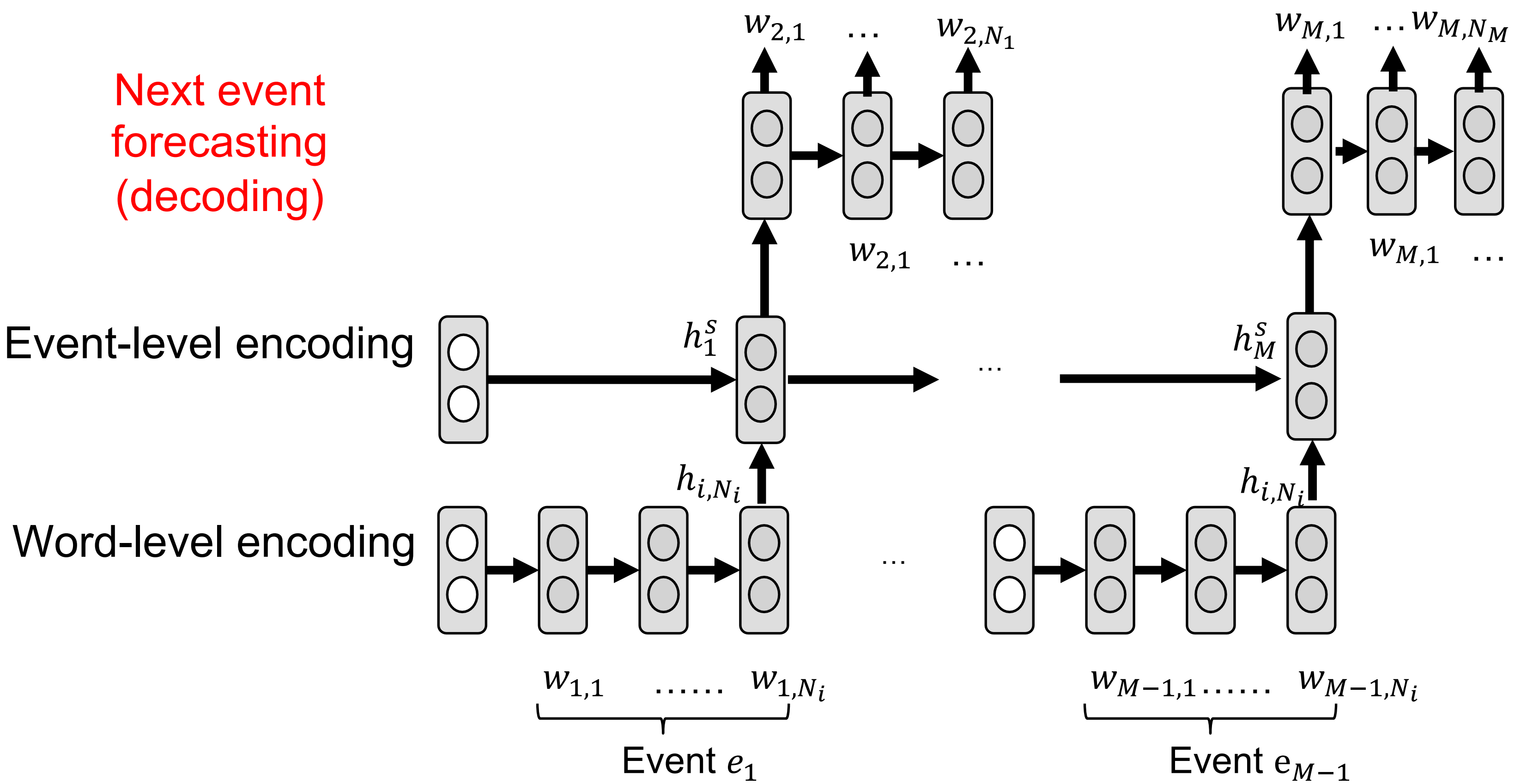
Part 1: Word-level LSTM Encoding



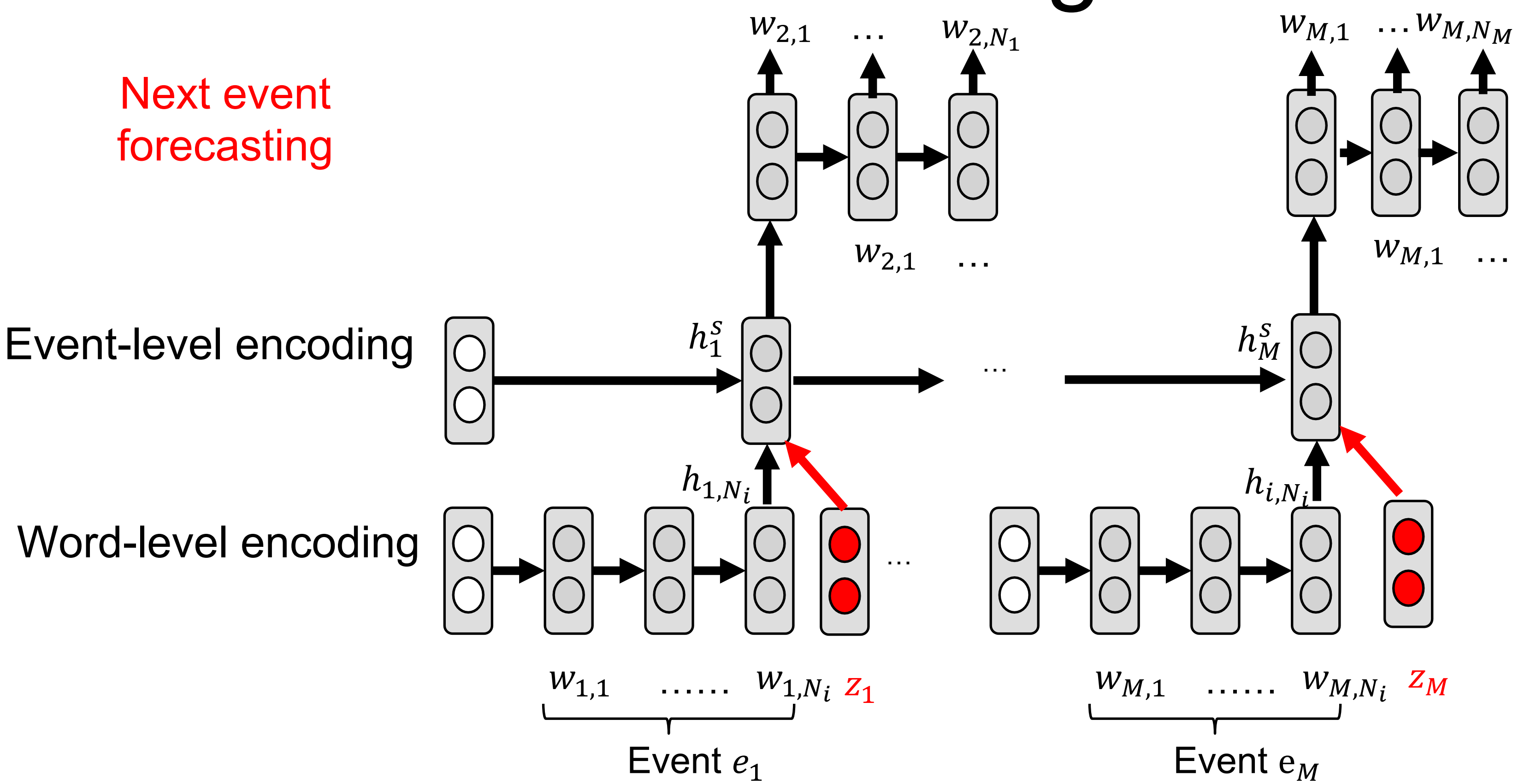
Part 2: Event-level LSTM Encoding



Part 3: Event-level LSTM Event Forecasting



Part 3: Event-level LSTM Event Forecasting



Experiments

- A large-scale Chinese news event dataset containing 15,254 news series from Sina News. Each news series consists of a sequence of news articles (or a chain of relevant events) in temporal order, and the average number of articles for all news series is 50.

十一 过半 华东 华北 方向
(National Day) (more than half) (East China) (North China) (direction)
火车站 电脑 系统 故障
(railway Station) (computer) (system) (failure)
迎来 高峰 铁路 机票
(reach) (peak) (railway) (air ticket)
铁路 高峰 民航 机票 紧张
(railway) (peak) (civil aviation) (air ticket) (scarcity)

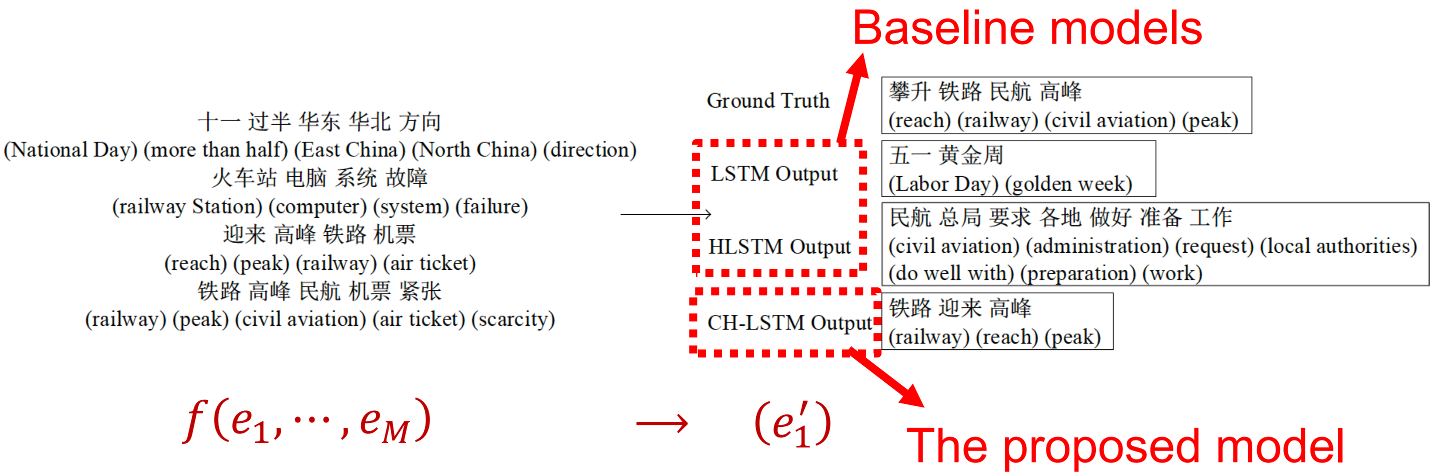
→

Ground Truth	攀升 铁路 民航 高峰 (reach) (railway) (civil aviation) (peak)
LSTM Output	五一 黄金周 (Labor Day) (golden week)
HLSTM Output	民航 总局 要求 各地 做好 准备 工作 (civil aviation) (administration) (request) (local authorities) (do well with) (preparation) (work)
CH-LSTM Output	铁路 迎来 高峰 (railway) (reach) (peak)

$$f(e_1, \cdots, e_M) \quad \longrightarrow \quad (e_1')$$

Experiments

- A large-scale Chinese news event dataset containing 15,254 news series from Sina News. Each news series consists of a sequence of news articles (or a chain of relevant events) in temporal order, and the average number of articles for all news series is 50.



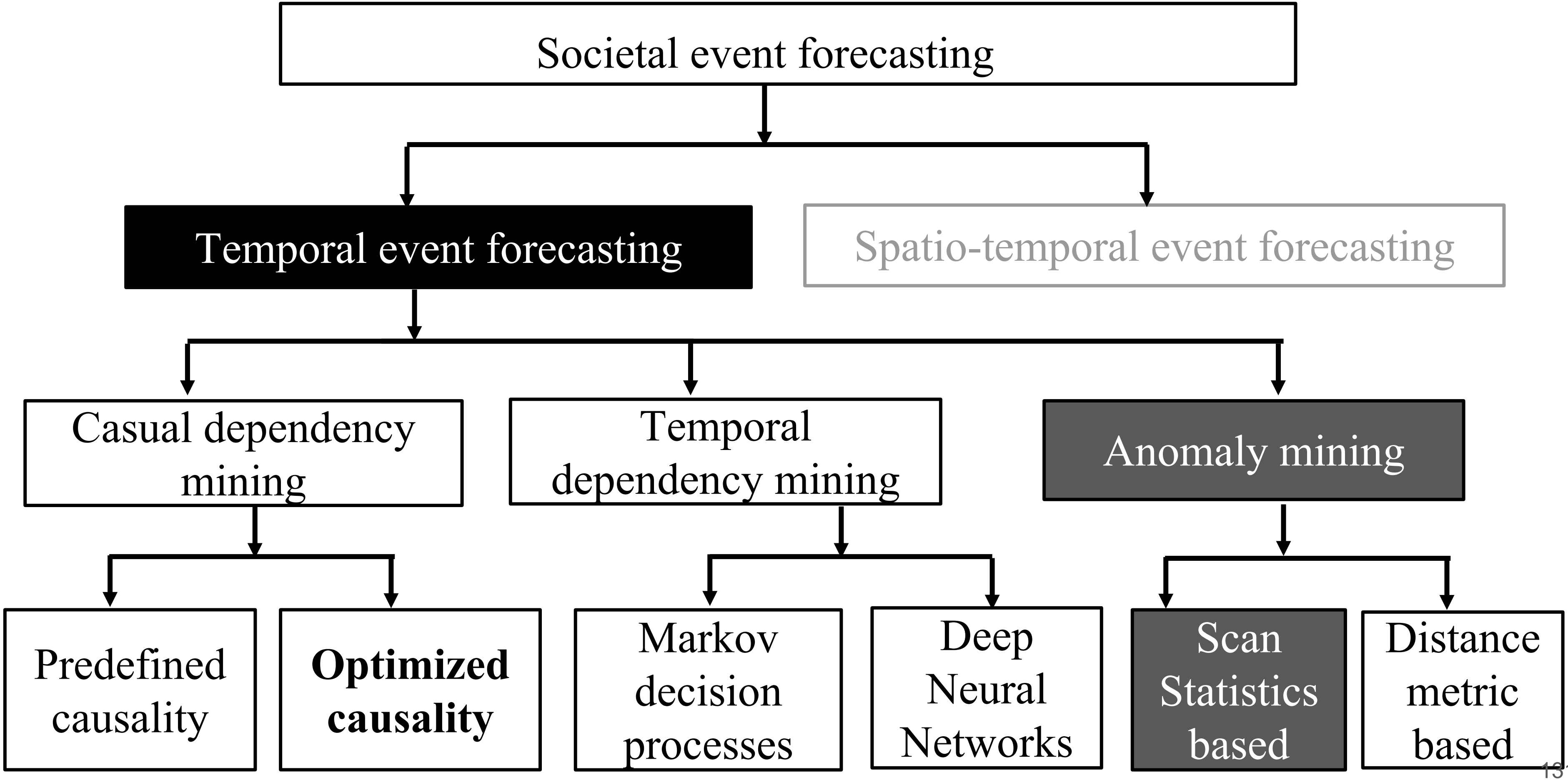
Empirical Results

Per-word classification error

Per-word perplexity of a model

Model	Perp	Error_Rate
Backoff N-Gram	264.07	93.03%
Modified Kneser-Ney	257.24	93.06%
Witten-Bell Discounting N-Gram	255.48	92.60%
LSTM	201.59 ± 0.38	75.22% ± 0.02%
HLSTM	129.44 ± 0.23	71.06% ± 0.02%
CH-LSTM	**127.74 ± 0.21**	**70.02% ± 0.01%**

Taxonomy

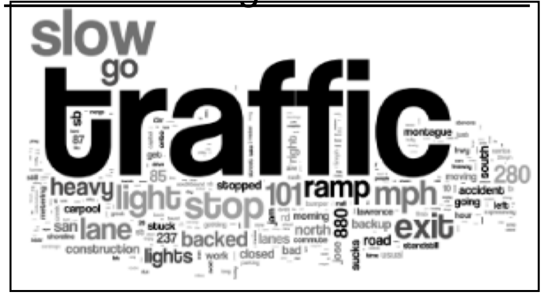

Societal event forecasting
Temporal event forecasting
Spatio-temporal event forecasting
Casual dependency mining
Temporal dependency mining
Anomaly mining
Predefined causality
Optimized causality
Markov decision processes
Deep Neural Networks
Scan Statistics based
Distance metric based

Event Forecasting from Twitter Data

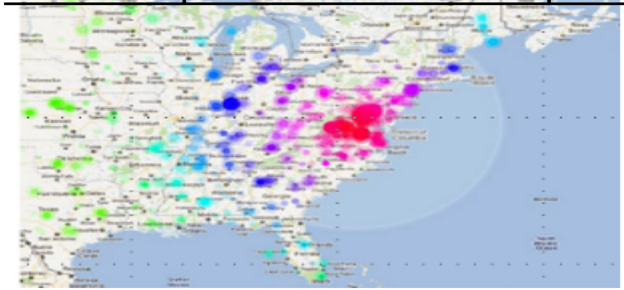
(Chen and Neill, KDD 2014)

<u>Protest in Mexico, 7/14/2012</u>

<u>2012 Washington D.C. Traffic</u>



<u>Tweet Map for 2011 VA Earthquake</u>



Social media is a real-time "sensor" of large-scale population behavior, and can be used for early detection of emerging events…

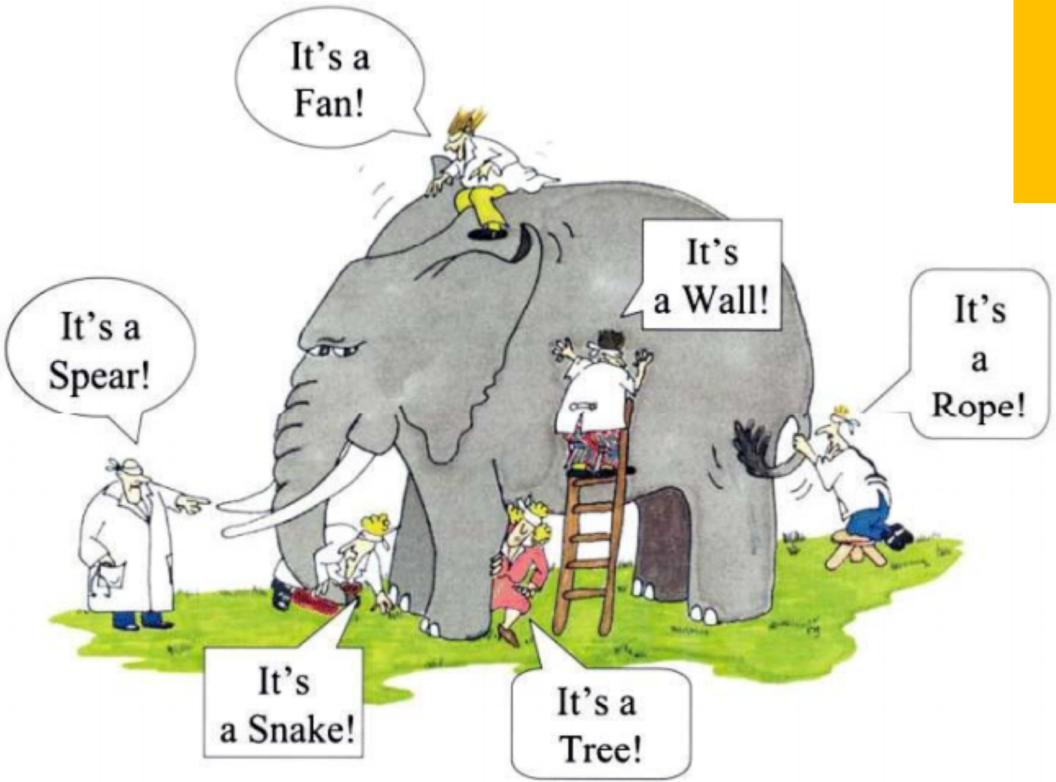
...but it is very complex, noisy, and subject to biases.

We have developed a new event forecasting methodology:

"Non-Parametric Heterogeneous Graph Scan" (NPHGS)

<u>Applied to</u>: civil unrest prediction, rare disease outbreak detection, and early detection of human rights events.

Technical Challenges



Integration of multiple heterogeneous information sources!

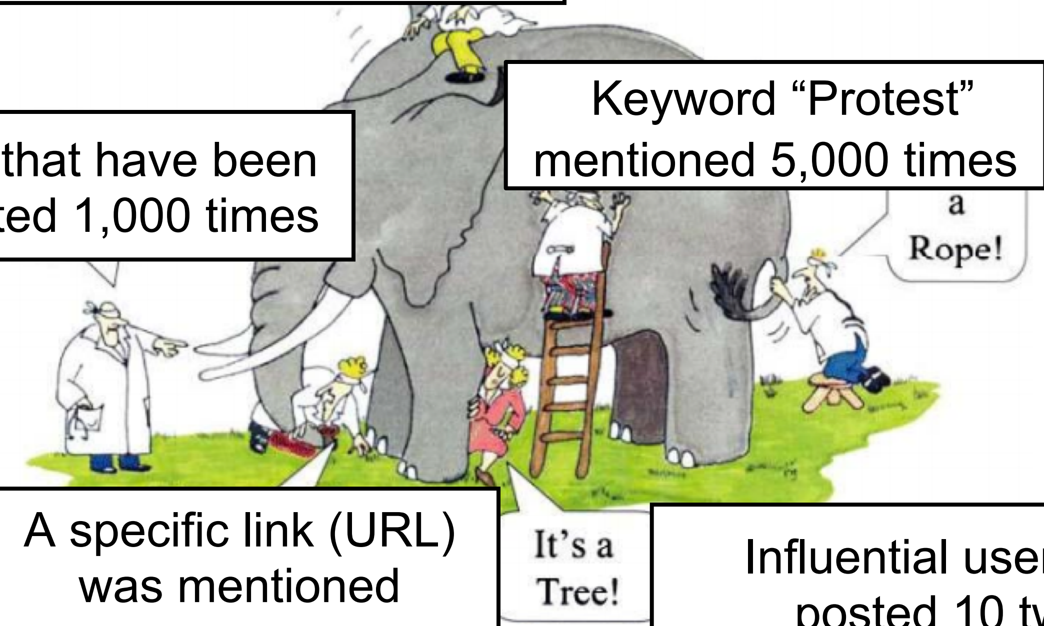
Technical Challenges

One week before Mexico's 2012 presidential election:

Hashtag "#Megamarch" mentioned 1,000 times

Keyword "Protest" mentioned 5,000 times

Tweets that have been re-tweeted 1,000 times



Mexico City has 5,000 active users and 100,000 tweets

A specific link (URL) was mentioned 866 times

Influential user "Zeka" posted 10 tweets

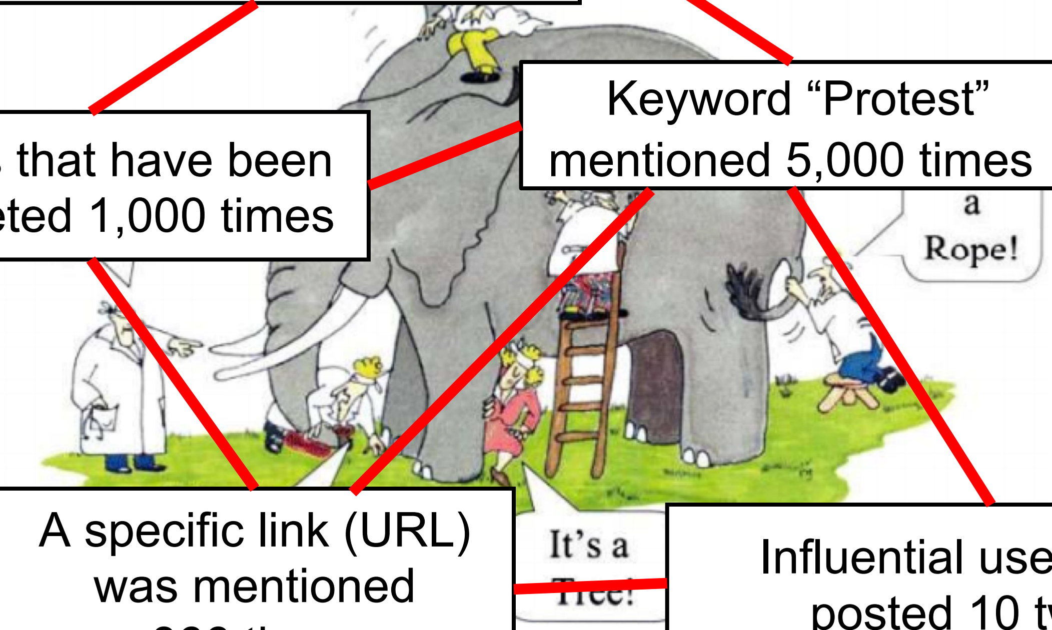
Technical Challenges

One week before Mexico's 2012 presidential election:

Hashtag "#Megamarch" mentioned 1,000 times

Keyword "Protest" mentioned 5,000 times

Tweets that have been re-tweeted 1,000 times

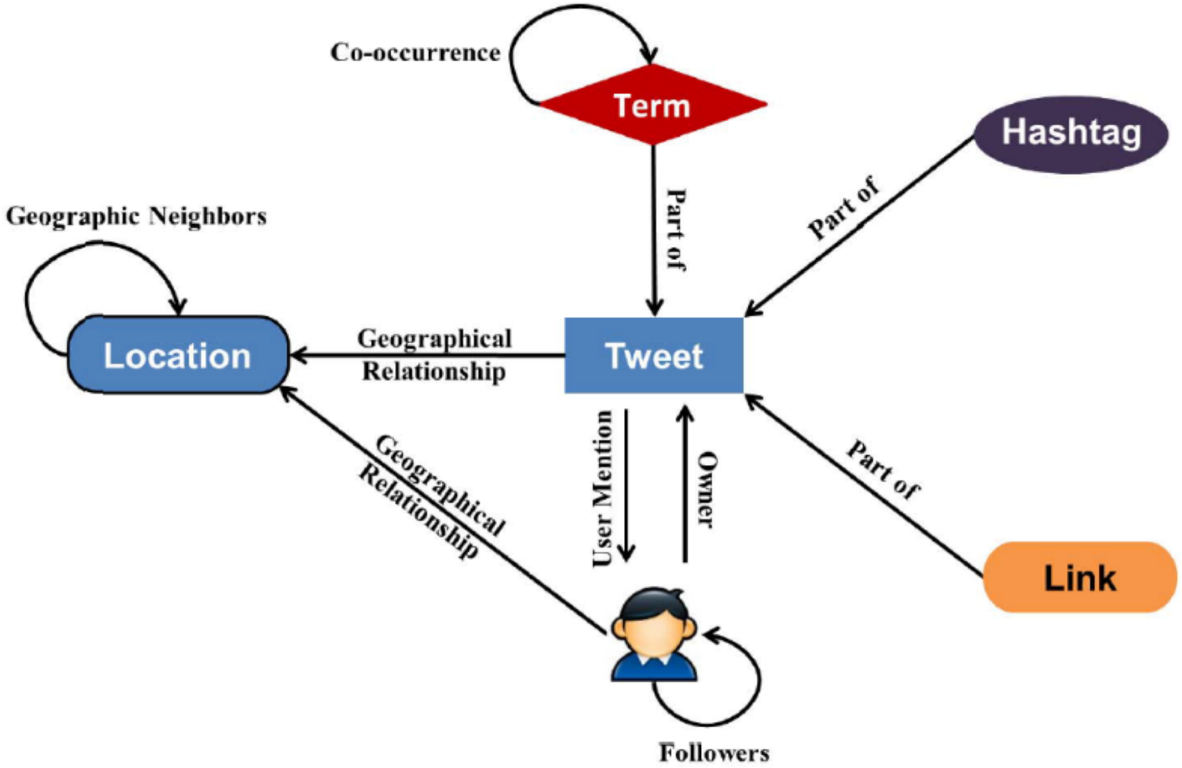


Mexico City has 5,000 active users and 100,000 tweets

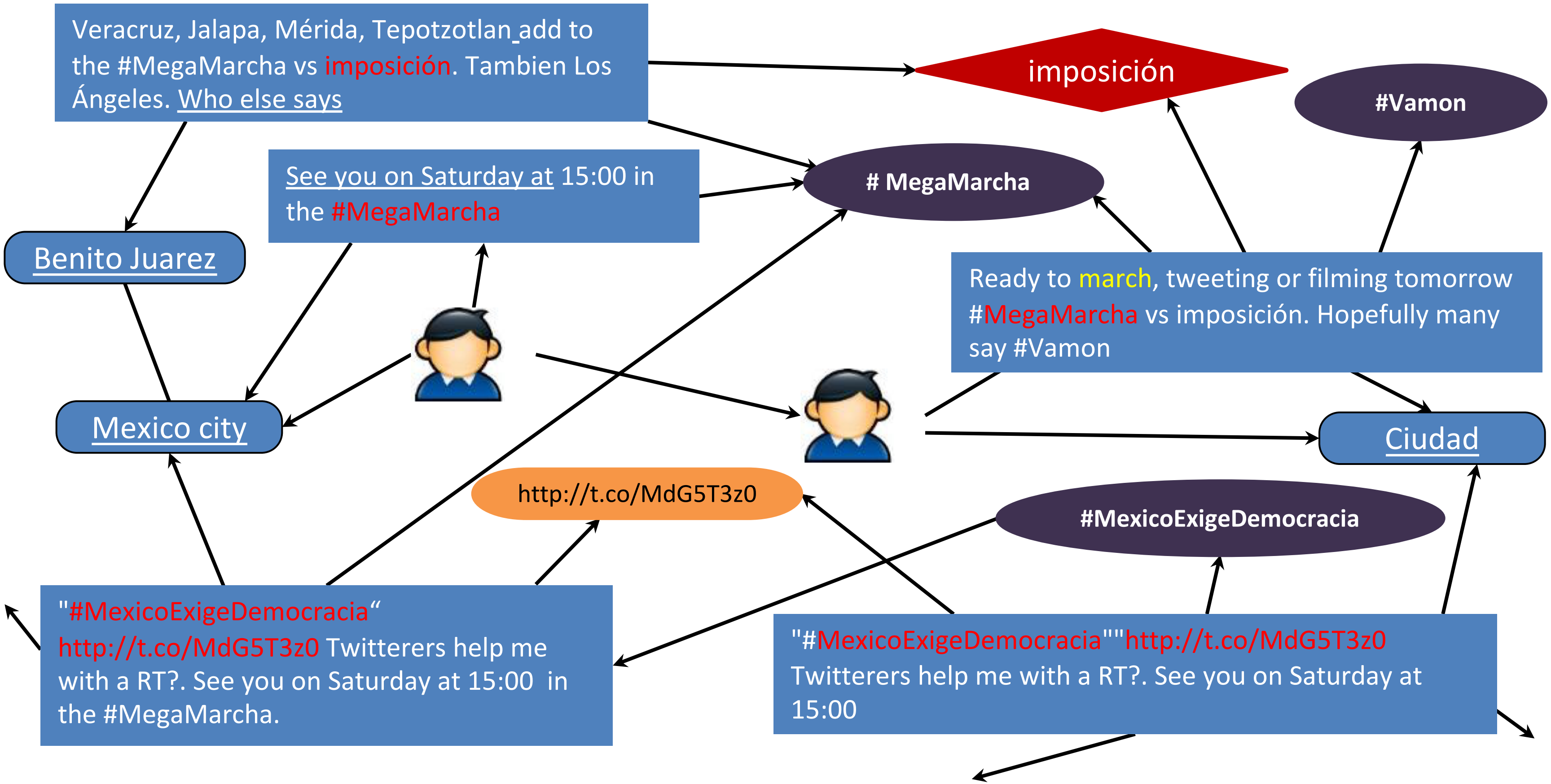
A specific link (URL) was mentioned 866 times

Influential user "Zeka" posted 10 tweets

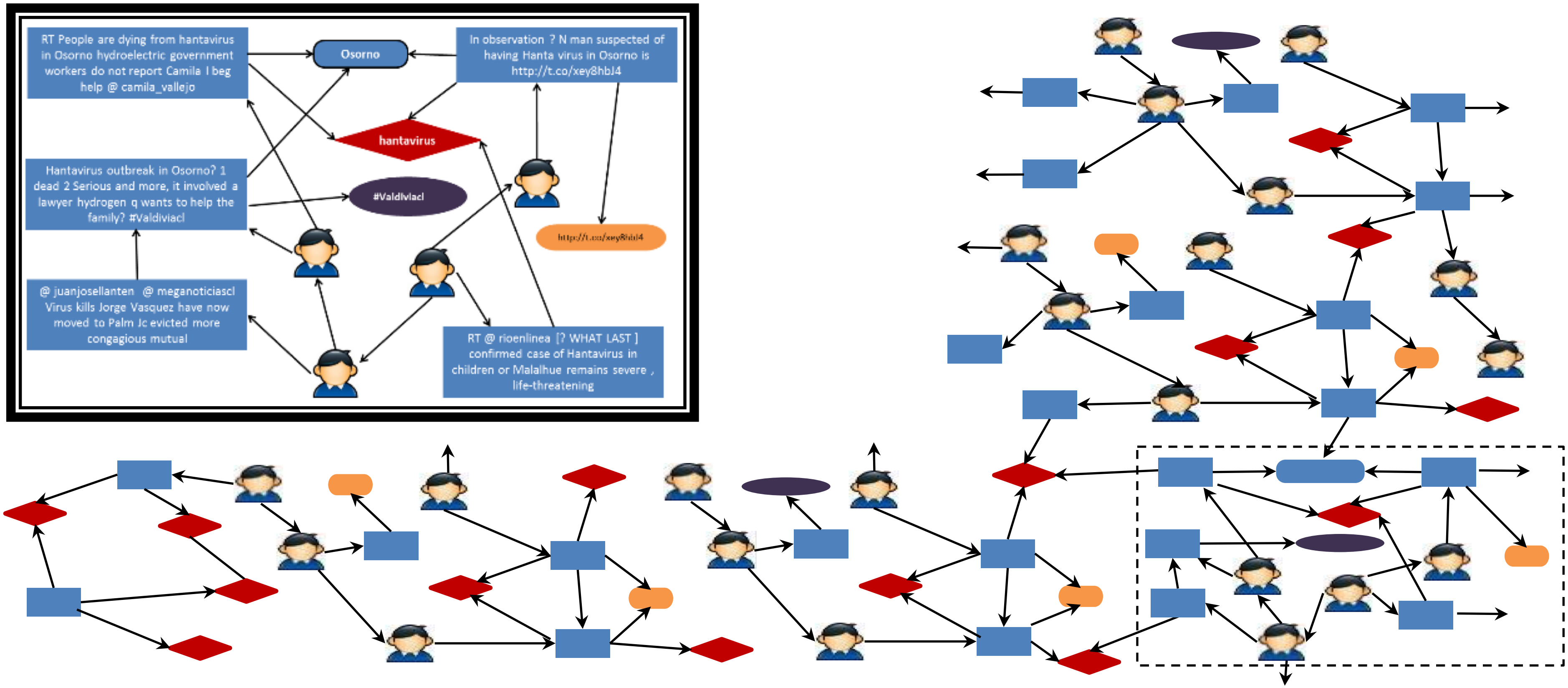
Twitter Heterogeneous Network



Twitter Heterogeneous Network



Twitter Heterogeneous Network



Nonparametric Heterogeneous Graph Scan

(Chen and Neill, KDD 2014)

1) We model the heterogeneous social network as a **sensor network**. Each node senses its local neighborhood, computes multiple features, and reports the overall degree of anomalousness.

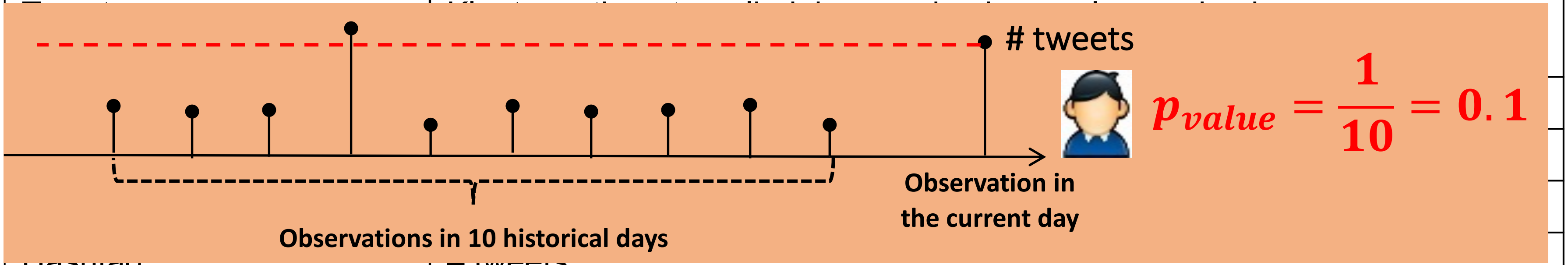
2) We compute an **empirical p-value** for each node:
   - Uniform on [0,1] under the null hypothesis of no events.
   - We search for subgraphs of the network with a higher than expected number of low (significant) empirical p-values.

3) We can scale up to very large heterogeneous networks:
   - Heuristic approach: **iterative subgraph expansion** (“greedy growth” to subset of neighbors on each iteration).
   - We can efficiently find the best subset of neighbors, ensuring that the subset remains connected, at each step.

Sensor network modeling

Each node reports an empirical p-value measuring the current level of anomalousness for each time interval (hour or day).

Node (Entity) Type	Features
User	# tweets, # retweets, # followers, #followees, #mentioned_by, #replied_by, diffusion graph depth, diffusion graph size



Features for each node $\xrightarrow{\text{empirical calibration}}$ Individual p-value for each feature $\xrightarrow{\text{min}}$ Minimum empirical p-value for each node $\xrightarrow{\text{empirical calibration}}$ Overall p-value for each node

Nonparametric scan statistics

Number of nodes in S with p-values $\leq\alpha$.

Subgraph

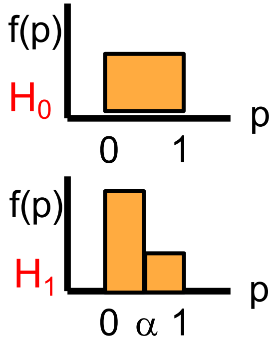
$$F(\boxed{S}) = \max_{\boxed{\alpha}\leq\alpha_{max}} F_\alpha(S) = \max_{\alpha\leq\alpha_{max}} \phi(\alpha, \boxed{N_\alpha(S)}, \boxed{N(S)})$$

Significance level

Number of nodes in S

Berk-Jones (BJ) statistic:

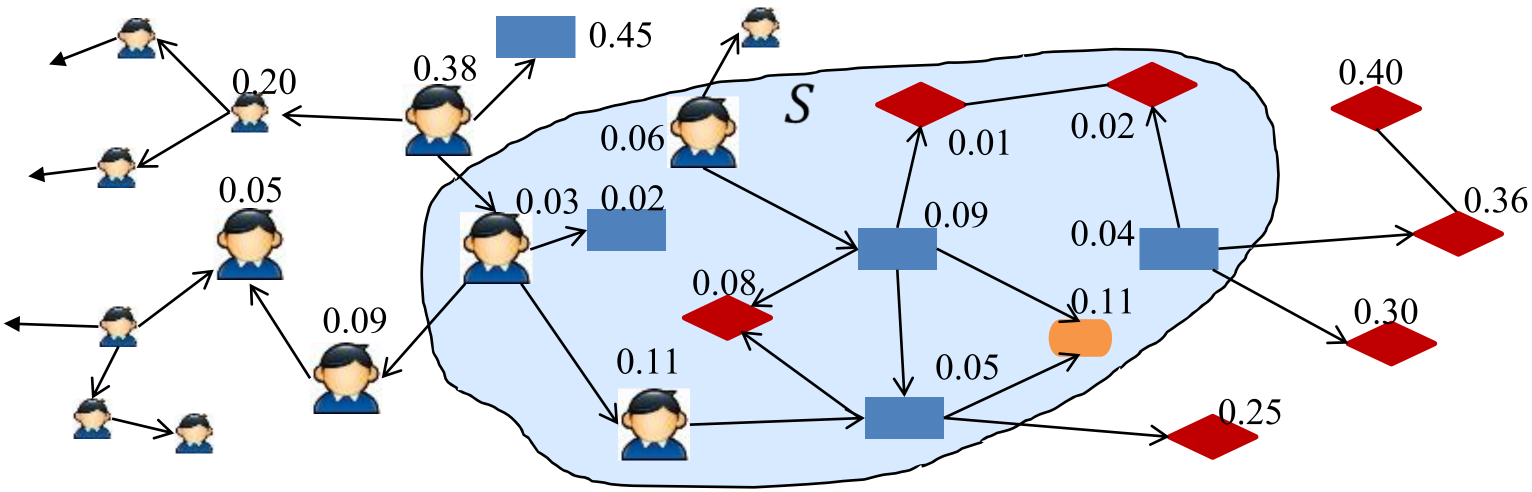
$$\phi_{BJ}(\alpha, N_\alpha(S), N(S)) = N(S)K\left(\frac{N_\alpha}{N}, \alpha\right)$$



Kullback-Liebler divergence:

$$K(x,y) = x\log\left(\frac{x}{y}\right) + (1-x)\log\left(\frac{1-x}{1-y}\right)$$

Nonparametric graph scanning



$$\boxed{S^\star = \underset{S \subseteq \mathbb{V}: S \text{ is connected}}{\operatorname{argmax}} \boxed{F(S)}}$$

We propose an approximate algorithm with time cost O(|V| log |V|).

NPHGS evaluation- civil unrest

Country	# of tweets	News source*
Argentina	29,000,000	Clarín; La Nación; Infobae
Chile	14,000,000	La Tercera; Las Últimas Notícias; El Mercurio
Colombia	22,000,000	El Espectador; El Tiempo; El Colombiano
Ecuador	6,900,000	El Universo; El Comercio; Hoy

Gold standard dataset: 918 civil unrest events between July and December 2012.

Example of a gold standard event label:

PROVINCE = "El Loa" COUNTRY = "Chile"
DATE = "2012-05-18" LINK = "http://www.pressenza.com/2012/05/..."
DESCRIPTION = "A large-scale march was staged by inhabitants of the northern city of Calama, considered the mining capital of Chile, who demanded the allocation of more resources to copper mining cities"

We compared the detection performance of our NPHGS approach to homogeneous graph scan methods and to a variety of state-of-the-art methods previously proposed for Twitter event forecasting.

NPHGS results- civil unrest

Method	FPR (FP/Day)	TPR (Forecasting)	TPR (Forecasting & Detection)	Lead Time (Days)	Lag Time (Days)	Run Time (Hours)
ST Burst Detection	0.65	0.07	0.42	1.10	4.57	30.1
Graph Partition	0.29	0.03	0.15	0.59	6.13	18.9
Earthquake	0.04	0.06	0.17	0.49	5.95	18.9
RW Event	0.10	0.22	0.25	0.93	5.83	16.3
Geo Topic Modeling	0.09	0.06	0.08	0.01	6.94	9.7
NPHGS (FPR=.05)	0.05	0.15	0.23	0.65	5.65	38.4
NPHGS (FPR=.10)	0.10	0.31	0.38	1.94	4.49	38.4
NPHGS (FPR= .15)	0.15	0.37	0.42	2.28	4.17	38.4
NPHGS (FPR=.20)	0.20	0.39	0.46	2.36	3.98	38.4

Table 3: Comparison between NPHGS and Existing Methods on the civil unrest datasets

NPHGS outperforms existing representative techniques for both event detection and forecasting, increasing **detection power**, **forecasting accuracy**, and **forecasting lead time** while reducing **time to detection**.

Similar improvements in performance were observed on a second task:

Early detection of rare disease outbreaks, using gold standard data about 17 hantavirus outbreaks from the Chilean Ministry of Health.

Part 2: References

(a) Causal dependency mining

i. Predened causality

- Muthiah, S., Butler, P., Khandpur, R. P., Saraf, P., Self, N., Rozovskaya, A., ... & Marathe, A. (2016, August). Embers at 4 years: Experiences operating an open source indicators forecasting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 205-214). ACM.
- Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. Icwsm, 10(1), 178-185.
- Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. Journal of computational science, 2(1), 1-8.

ii. Optimized causality

- Arias, M., Arratia, A., & Xuriguera, R. (2013). Forecasting with twitter data. ACM Transactions on Intelligent Systems and Technology (TIST), 5(1), 8.
- Kruengkrai, C., Torisawa, K., Hashimoto, C., Kloetzer, J., Oh, J. H., & Tanaka, M. (2017). Improving Event Causality Recognition with Multiple Background Knowledge Sources Using Multi-Column Convolutional Neural Networks. In AAAI (pp. 3466-3473).
- Radinsky, K., Davidovich, S., & Markovitch, S. (2012). Learning to predict from textual data. Journal of Artificial Intelligence Research, 45, 641-684.
- Radinsky, K., & Horvitz, E. (2013, February). Mining the web to predict future events. In Proceedings of the sixth ACM international conference on Web search and data mining (pp. 255-264). ACM.

Part 2: References

(b) Temporal dependency mining

i. Markov decision processes

- Qiao, F., Li, P., Zhang, X., Ding, Z., Cheng, J., & Wang, H. (2017). Predicting social unrest events with hidden Markov models using GDELT. Discrete Dynamics in Nature and Society, 2017.
- Schrodt, P. A. (2006). Forecasting conflict in the Balkans using hidden Markov models. In Programming for Peace (pp. 161-184). Springer, Dordrecht.

ii. Deep neural networks

- Granroth-Wilding, M., & Clark, S. (2016, February). What happens next? event prediction using a compositional neural network model. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (pp. 2727-2733). AAAI Press.
- Hu, L., Li, J., Nie, L., Li, X. L., & Shao, C. (2017). What Happens Next? Future Subevent Prediction Using Contextual Hierarchical LSTM. In AAAI (pp. 3450-3456).
- Pichotta, K., & Mooney, R. J. (2016, February). Learning Statistical Scripts with LSTM Recurrent Neural Networks. In AAAI (pp. 2800-2806).
- Wang, Z., & Zhang, Y. (2017, August). DDoS event forecasting using Twitter data. In Proceedings of the 26th International Joint Conference on Artificial Intelligence (pp. 4151-4157). AAAI Press.

Coffee Break

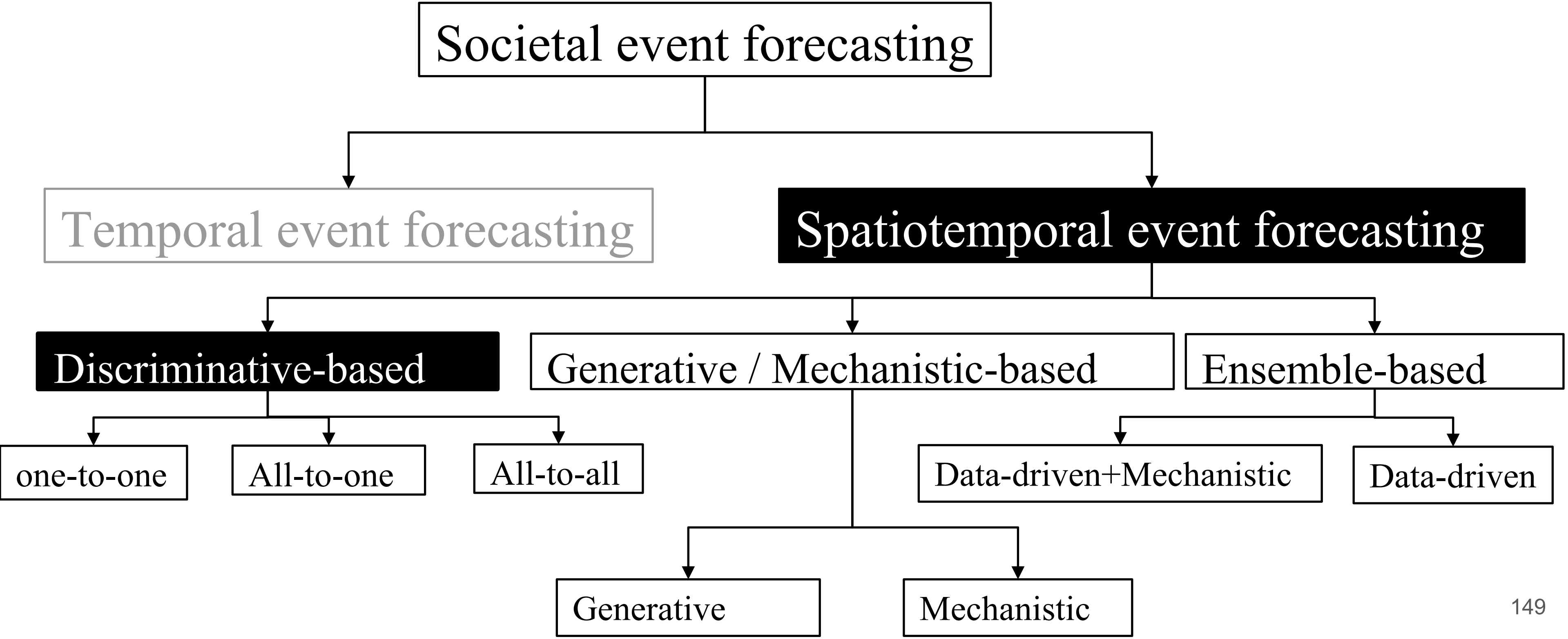
15 Minutes

Part 3: Spatio-Temporal Event Forecasting

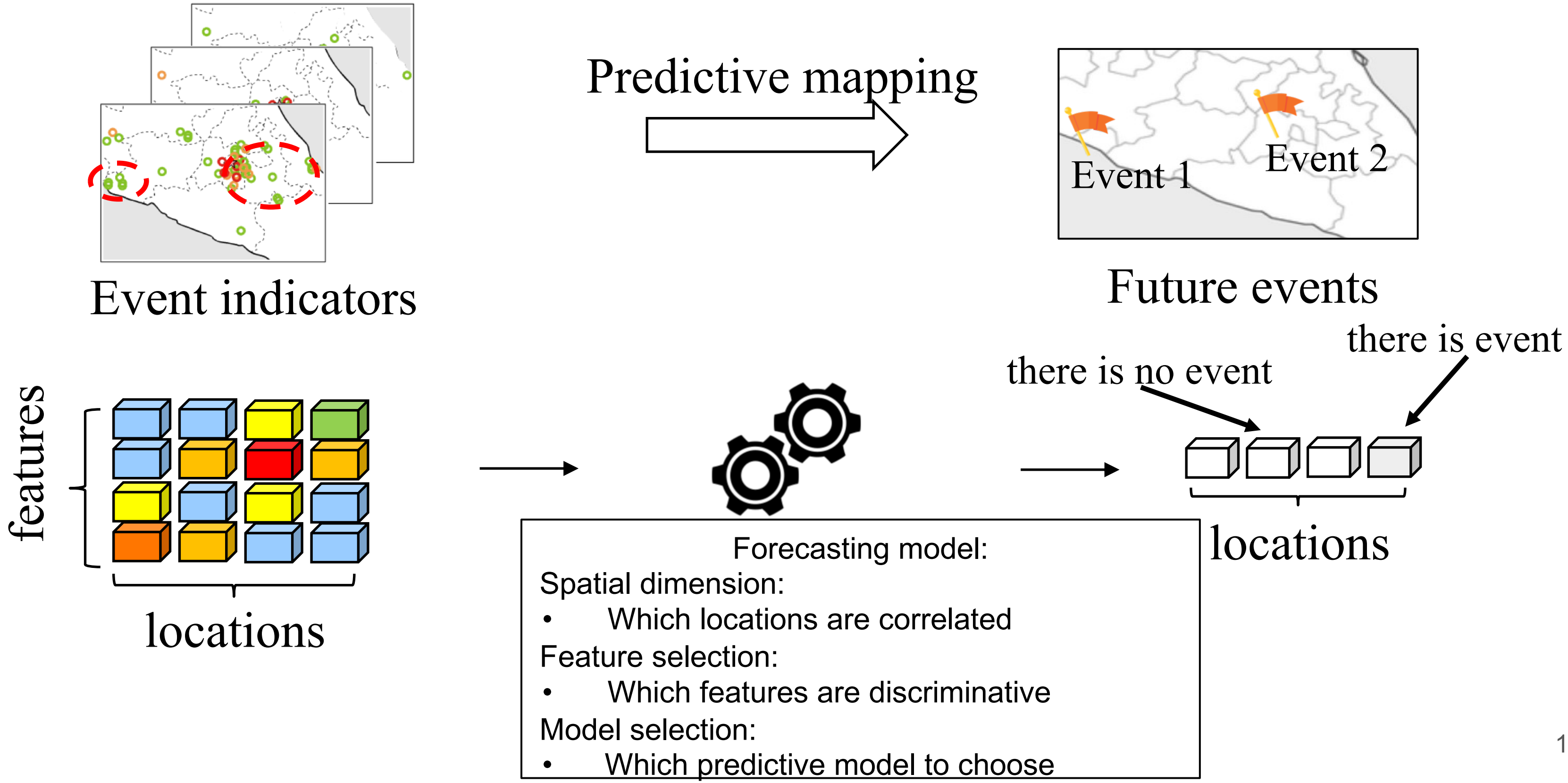
Liang Zhao (George Mason University)

GEORGE MASON UNIVERSITY

Taxonomy

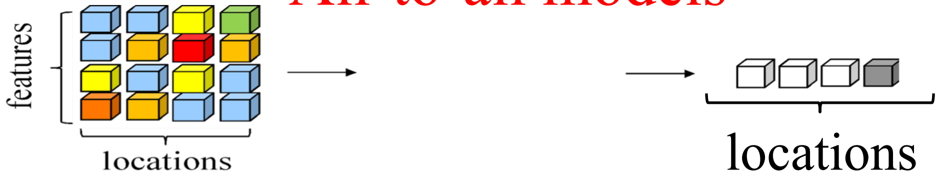

Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative-based
Generative / Mechanistic-based
Ensemble-based
one-to-one
All-to-one
All-to-all
Data-driven+Mechanistic
Data-driven
Generative
Mechanistic

Discriminative Learning-based



Categorization

All-to-all models



S: number of locations, K: number of features

Advantages:

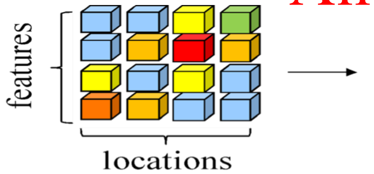
- Consider spatial dependency of inputs
- Consider spatial autocorrelation of outputs

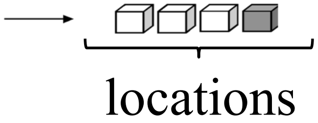
Disadvantages:

- Time&memory consuming, complexity $\geq S^2 \cdot K$
- Complex model
- Large data required
- Tricky to define how locations are auto-correlated.

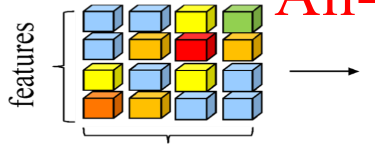
Categorization

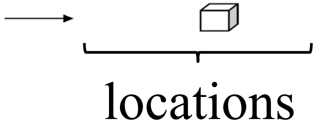
All-to-all models





All-to-one models





S: number of locations, K: number of features

Advantages:
- Consider spatial dependency of inputs
- Consider spatial autocorrelation of outputs

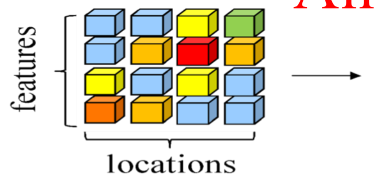
Disadvantages:
- Time&memory consuming, complexity $\geq S^2 \cdot K$
- Complex model
- Large data required
- Tricky to define how locations are auto-correlated.

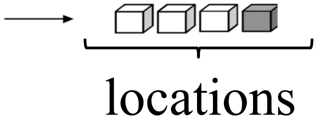
Advantages:
- Consider spatial dependency of inputs
- More efficient: Complexity $\geq S \cdot K$

Disadvantages:
- Complex model
- Large data required
- Ignore potential correlation among the events

Categorization

All-to-all models





S: number of locations, K: number of features

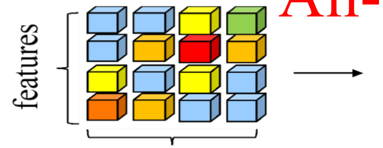
Advantages:

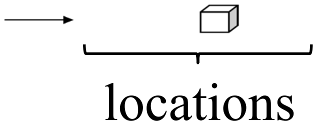
- Consider spatial dependency of inputs
- Consider spatial autocorrelation of outputs

Disadvantages:

- Time&memory consuming, complexity $\geq S^2 \cdot K$
- Complex model
- Large data required
- Tricky to define how locations are auto-correlated.

All-to-one models





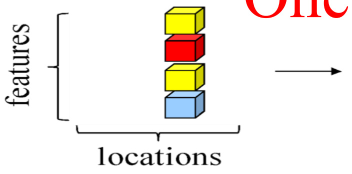
Advantages:

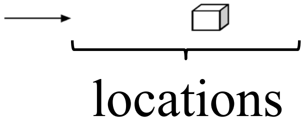
- Consider spatial dependency of inputs
- More efficient: Complexity $\geq S \cdot K$

Disadvantages:

- Complex model
- Large data required
- Ignore potential correlation among the events

One-to-one models





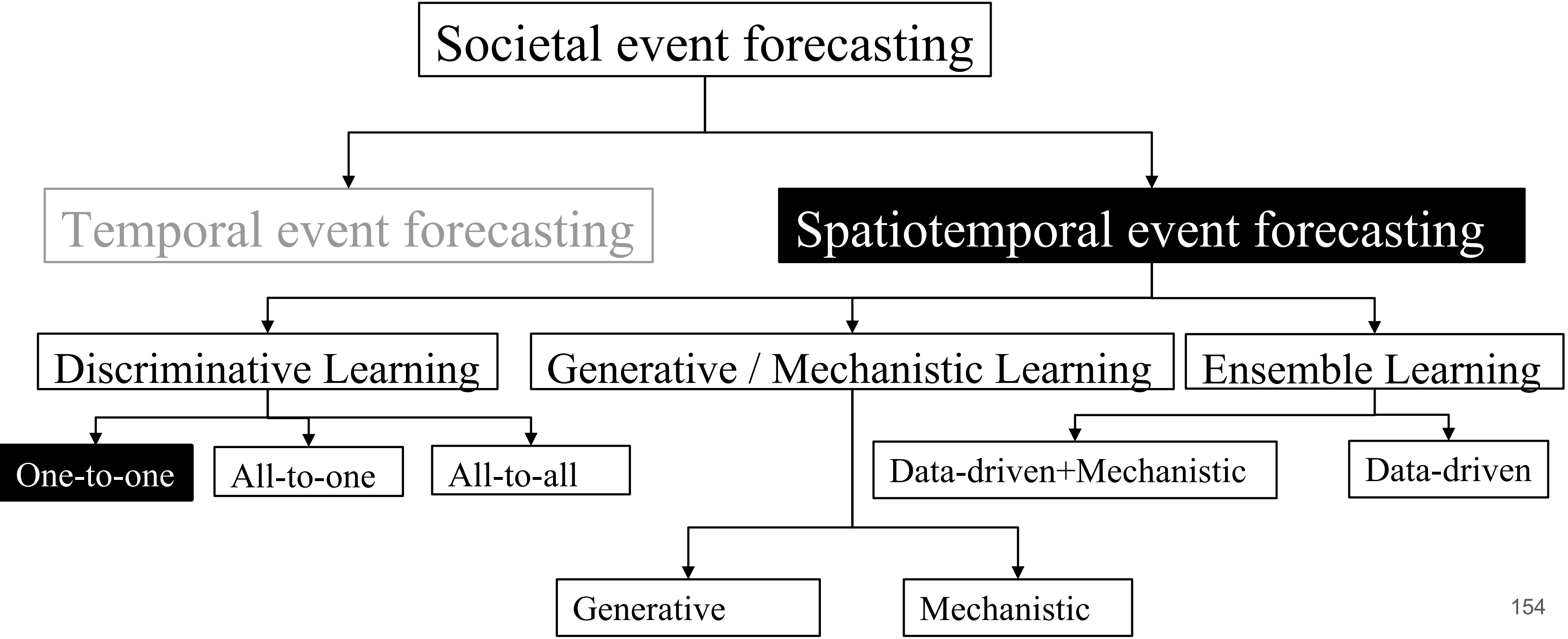
Advantages:

- Simple model, easy to train
- Small data is needed
- More efficient: Complexity $\geq K$

Disadvantages:

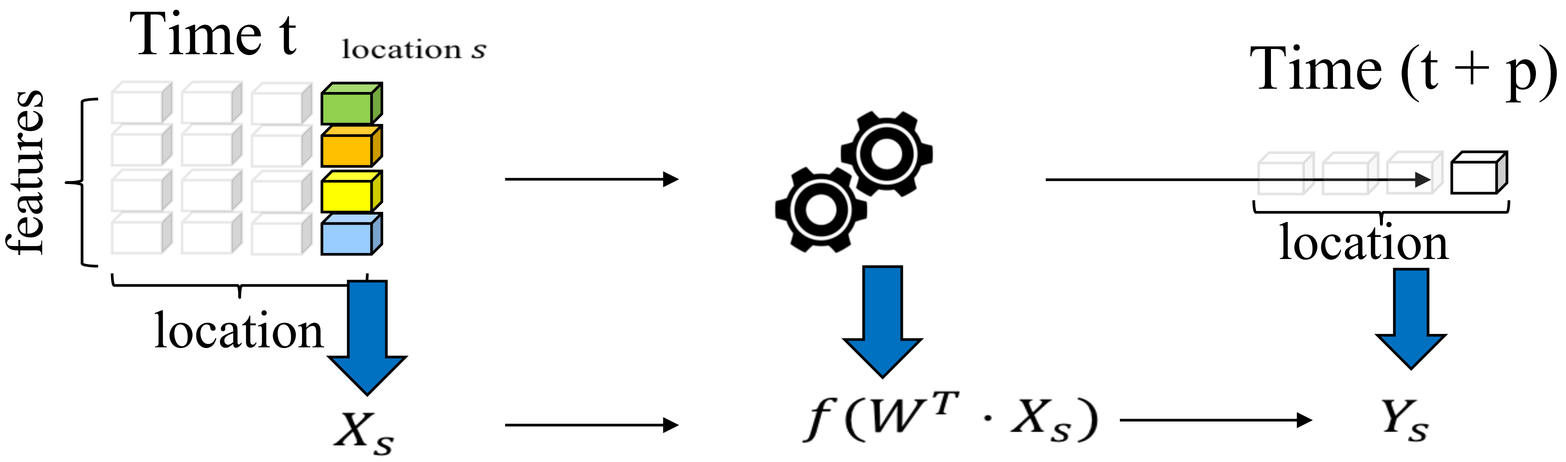
- Cannot consider spatial dependency of inputs
- Ignore potential correlation among the events

Taxonomy


Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative Learning
Generative / Mechanistic Learning
Ensemble Learning
One-to-one
All-to-one
All-to-all
Data-driven+Mechanistic
Data-driven
Generative
Mechanistic

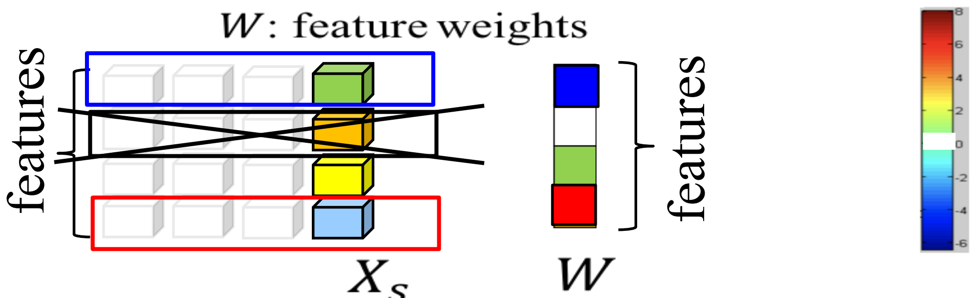
One-to-One models

Use individual location to forecast for each corresponding individual location

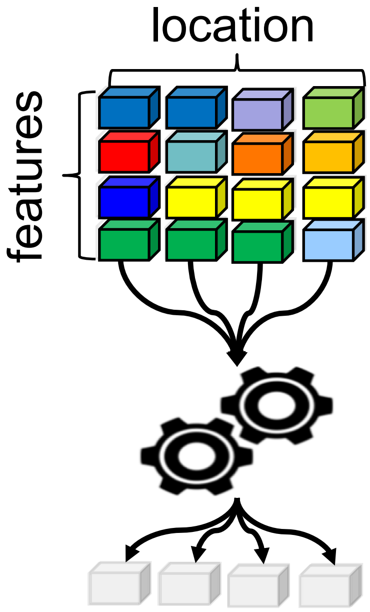


f: predictive model

- Logistic regression [Gerber, DSS'15]
- Linear regression [Gerber, TCSS'18]
- So on so forth…



Category 1: All locations share a single model



Pro: sufficient data to train model

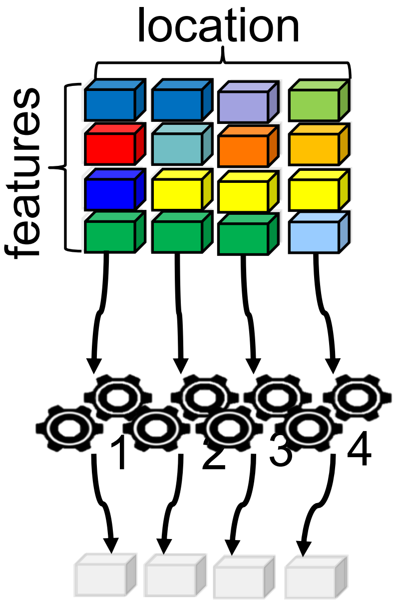
Con: ignore the individual city's exclusive characteristics (size, population, etc.)

City: Mexico City
Population: 8M
Size: 573 mi²

1K protest tweets have different meanings to these two locations

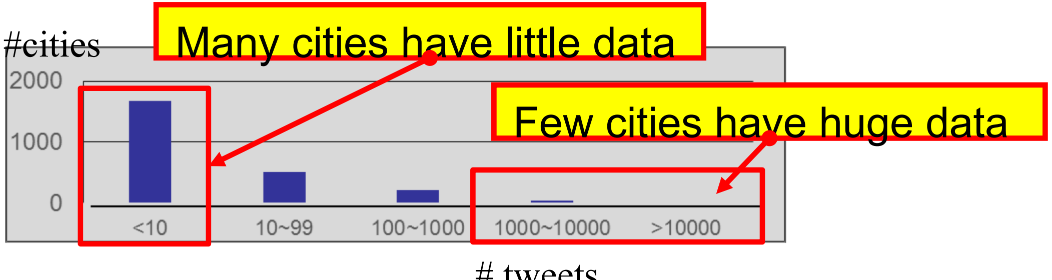
City: Taxco
Population: 39K
Size: 134 mi²

Category 2: Each model for each location



Pro: consider the individual location's exclusive characteristics

Con: 1. Insufficient data for small cities.



2. Ignore the relatedness among different locations

Relatedness among locations

- Similar expressions
- Same languages
- Shared keywords
- Relevant events
- Similar topics

Multi-task learning for Spatiotemporal Event Forecasting [Zhao et al., KDD'15]

Each model for each location + All locations share a single model

Pro: consider the exclusive characteristics

~~Con: 1. Ignore the relatedness among different locations~~

~~2. Insufficient data for small cities.~~

Pro: Sufficient training data

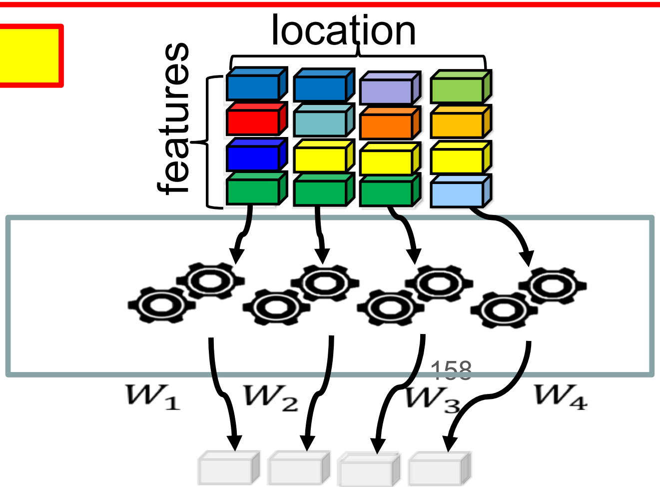
~~Con: Ignore the individual city's exclusive characteristics~~

Combine

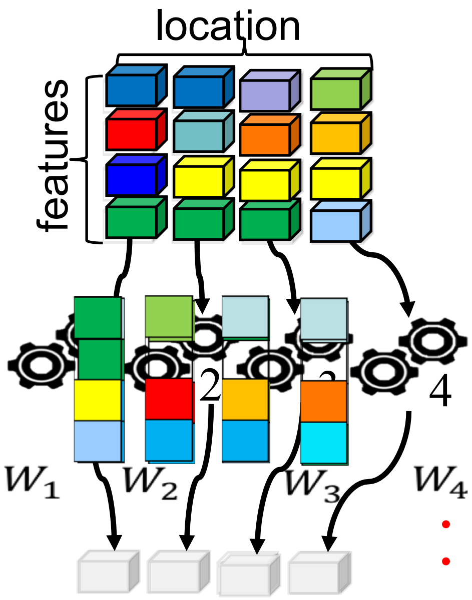
Jointly preserve:

- Spatial dependency
- Spatial heterogeneity

Regularize all the models
Enforce knowledge sharing



Multi-task learning for Spatiotemporal Event Forecasting [Zhao et al., KDD'15]



$$\min_{W} \sum_{i=1}^{S} \mathcal{L}(W_i^T X_i, Y_i) + \lambda \cdot \mathcal{R}(W)$$

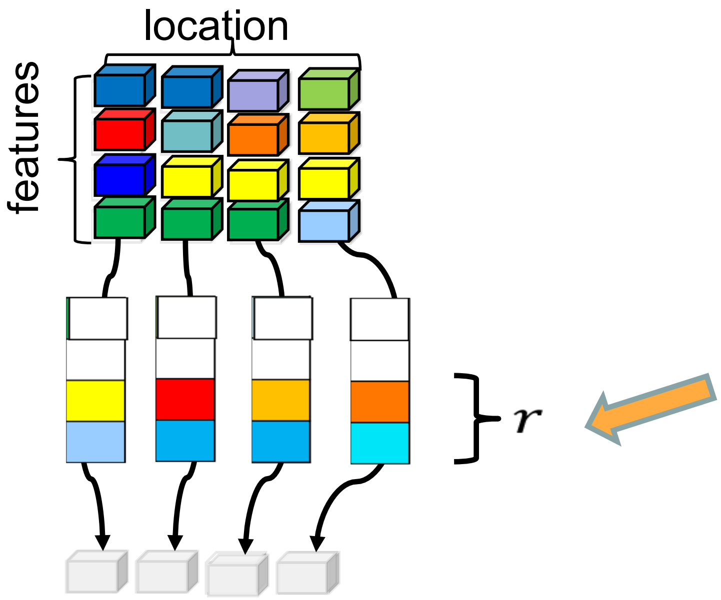
$$\mathcal{R}(W) = \|W\|_{2,1} = \sum_i \|W_i\|_2$$

Minimizing $l_{2,1}$ norm will make the matrix row sparse

- A feature important for a location will also tend to be also important
- Their weights value can be different.

- Keywords "basketball" and "music" is unimportant for "influenza outbreaks" for various locations;
- Keywords "cold" and "cough" is important to forecast "influenza outbreaks" for various locations;
- However, their weights are different in different locations (e.g., due to different population size in each location.)

More constraints



$$\min_{W} \sum_{i=1}^{L} \mathcal{L}(W_i^T X_{i,}) \quad + \lambda \cdot \mathcal{R}(W)$$

$$\text{s.t. } \sum_j^m I(\|W^j\| > 0) \leq r$$

Sometimes, the users have preference on how many features to select

Model optimization algorithm: Solved by projected gradient descent.

Experiments: Event Forecasting Performance

precision, recall, F-measure

Training set: Twitter data from July 1, 2012 to December 31, 2012
Testing set: Twitter data from January 1, 2013 to May 31, 2013

Label set: Authoritative news reports on civil unrest events

method	Mexico	Paraguay	Brazil	All Countries
DQEF	0.56, 0.40, 0.47	0.90, 0.15, 0.26	0.37, 0.34, 0.35	0.54, 0.38, 0.45
LASSO-K	0.68, 0.32, 0.44	1.00, 0.17, 0.29	0.62, 0.44, 0.51	0.72, 0.28, 0.40
DQEF+LASSO	0.57, 0.49, 0.53	1.00, 0.11, 0.20	0.42, 0.49, 0.45	0.55, 0.44, 0.49
LASSO	0.70, 0.36, 0.48	1.00, 0.17, 0.29	0.63, 0.43, 0.51	0.73, 0.30, 0.43
rMTFL-D	**0.96**, 0.12, 0.21	1.00, 0.02, 0.04	1.00, 0.07, 0.13	**0.77**, 0.15, 0.25
rMTFL-K	0.78, 0.45, 0.57	0.93, 0.43, 0.59	**0.79**, 0.55, 0.65	0.71, 0.51, 0.59
rMTFL	0.70 0.70 0.70	**0.96**, 0.32, 0.48	0.71, 0.52, 0.60	0.68, 0.57, 0.62
CMTFL-I	0.59, 0.87, 0.70	0.95, 0.39, 0.55	0.72, 0.60, **0.66**	0.62, 0.68, 0.65
CMTFL-II	0.71, **0.79**, **0.75**	0.78, **0.81**, **0.79**	0.76, **0.57**, 0.65	0.69, **0.71**, **0.70**

- Multitask models outperform the traditional LASSO models
- The proposed CMTFL II is generally the BEST

Selected Features

Few and not relevant keywords, due to the sparsity of the training data for small state

... the United States

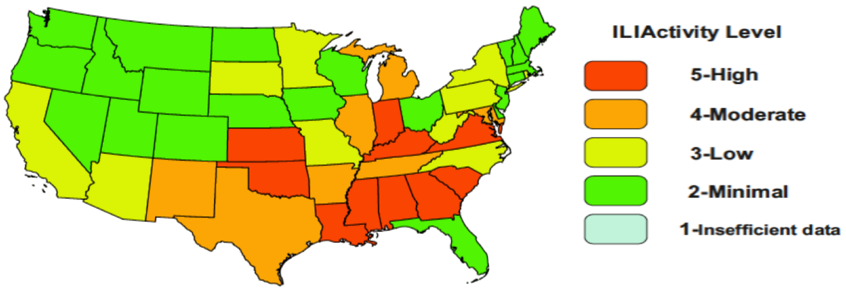
Methods	Features	Wyoming	Nebraska	Washington	New York	California	Alaska	Florida	New Mexico
LASSO	Static	four excuse works job diet cancelled boss ankle complicate NIH	birds drop thinks dealing warm body pissed practice masks class	jadi tired 101 birds 2nd cancer classes hands miss recover	drop chicken vomiting late bottle quickly miserable ate brought hrs	fast sleep decided ill started quite normal less years gak	immune	kalo four past 12s pigs pissed heard tea infected wasn	officially tea juga drop strains die nausea swear fight gettin
	Dynamic	TRUE	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	TRUE
rMTFL	Static	catching jab vaccination excuse daughter quickly outbreak poor died four	warm goin drop practice thinks class pissed excuse dealing body	ankle poor pray gym disease jadi finally quarantine thera severe	drop chicken begginning hospitalize month infections kind throat bro barely	fast appetite tired quite lemon energy vomit sleep normal killing	immune fever strep bug bird week flu virus vaccination tomorrow	12s pigs ebola past wasn helps tea practice heard kalo	coming slime thanks vomiting tea less positive catch starting weak
	Dynamic	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
CMTFL-II	Static	[illegible] doctor around	[illegible] week tomorrow	[illegible] school week	[illegible] home tonight	house school fever days sucks tonight bug stay bed tomorrow	sick cold bed school around home swine away throat bug	year soon tonight bug symptoms coming since tomorrow around work	days stay coming tomorrow away strep bug house soon sick
	Dynamic	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
CMTFL-III	Static	flu sick cold days bed feeling stomach week work soon	stomach cold sick feeling week days bed soon work flu	cold sick bed week days flu sucks stomach soon feeling	bed stomach cold days soon family sucks week feeling sick	bed days feeling cold week sick soon work sucks family	chills illness trip official wanted bring decided cancelle[illegible] avoid taking	[illegible] least pretty	[illegible] work soon
	Dynamic	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE

Does not ensure to include the dynamic features

Does not ensure to include the dynamic features

Multi-task Event Scale Forecasting
[Gao, and Zhao, AAAI'18]

Event Scale Forecasting (Gao et al., AAAI'18)



Influenza outbreaks in Week 13 of 2016

Generalize the output to ordinal!

Ordinal regression

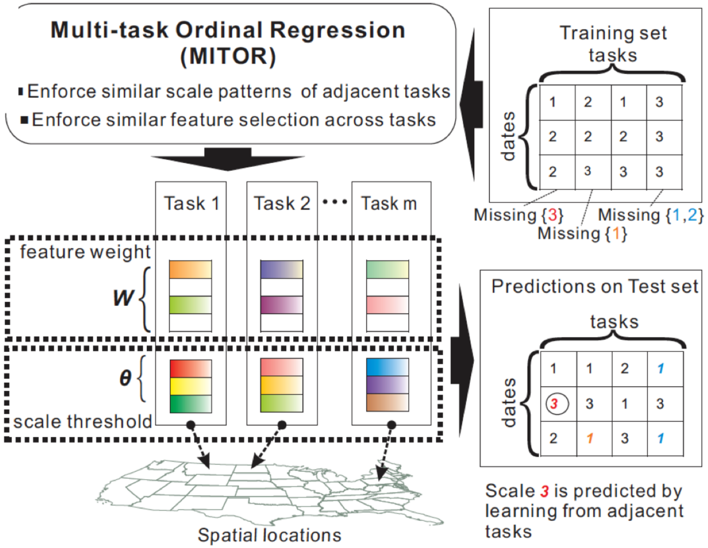
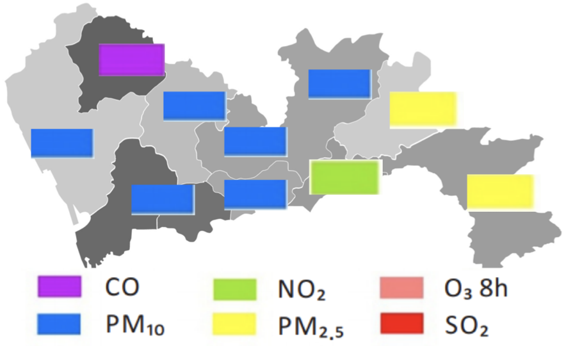


Figure 2: Flowchart of the proposed MITOR model

Multi-task Event Subtype Forecasting

[Gao, et al. AAAI'19]

Event Subtype Forecasting (Gao et al., AAAI'19)

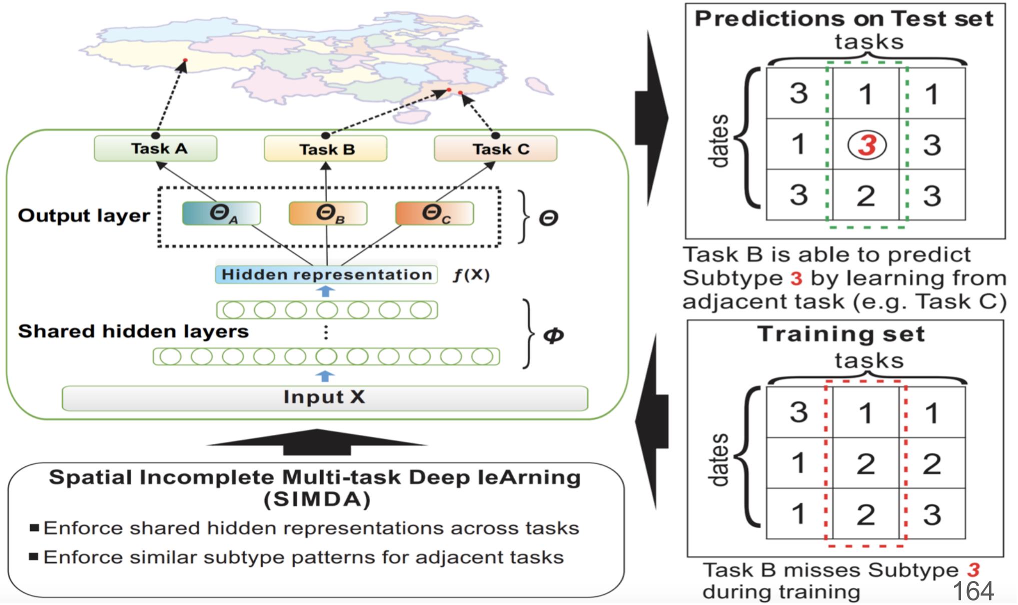


Primary Pollutant in one day in Shenzhen, China, 2013.

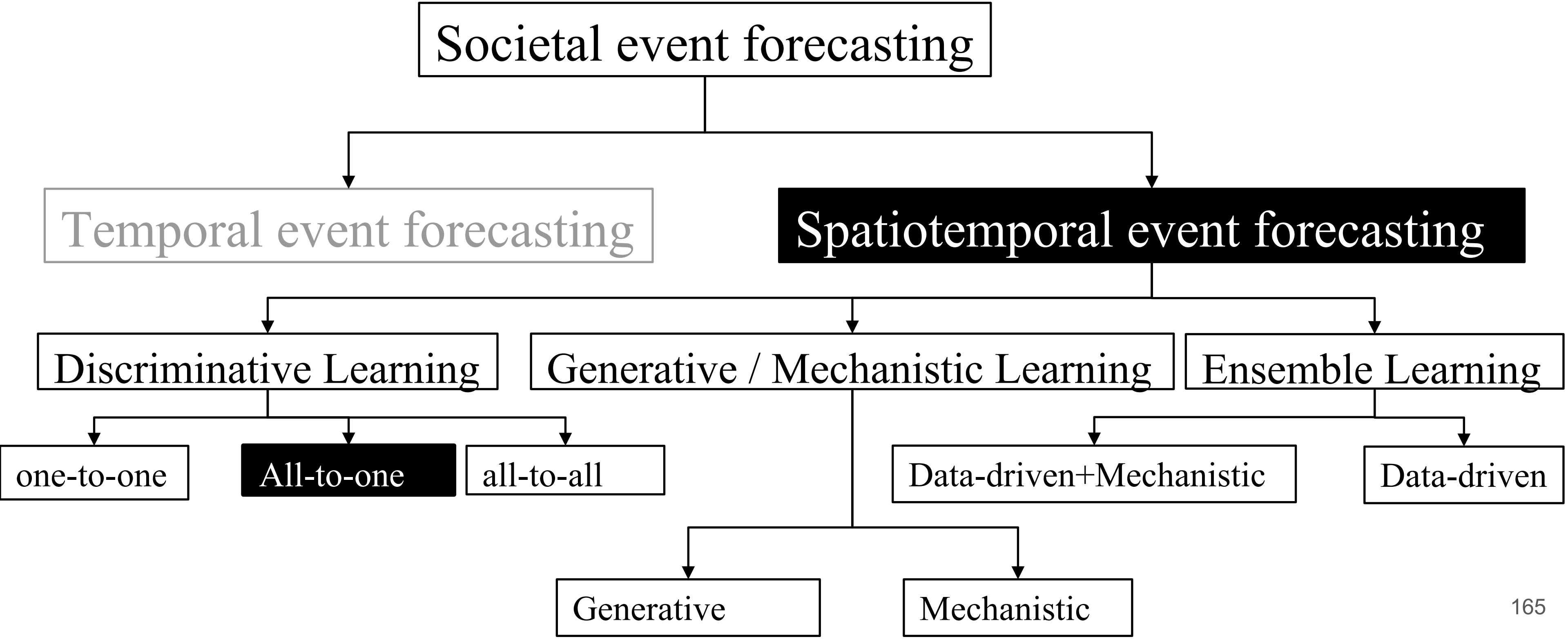
Multi-class classification

Generalize the output to multi-class!

Multi-class Classification



Taxonomy


Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative Learning
Generative / Mechanistic Learning
Ensemble Learning
one-to-one
All-to-one
all-to-all
Data-driven+Mechanistic
Data-driven
Generative
Mechanistic

All-to-one models

Use Multiple locations to forecast for each individual location

All the locations
(spatial dependency among indicators)

Each individual location
(spatial independency among events)

When the inputs have strong:

Spatial hierarchy

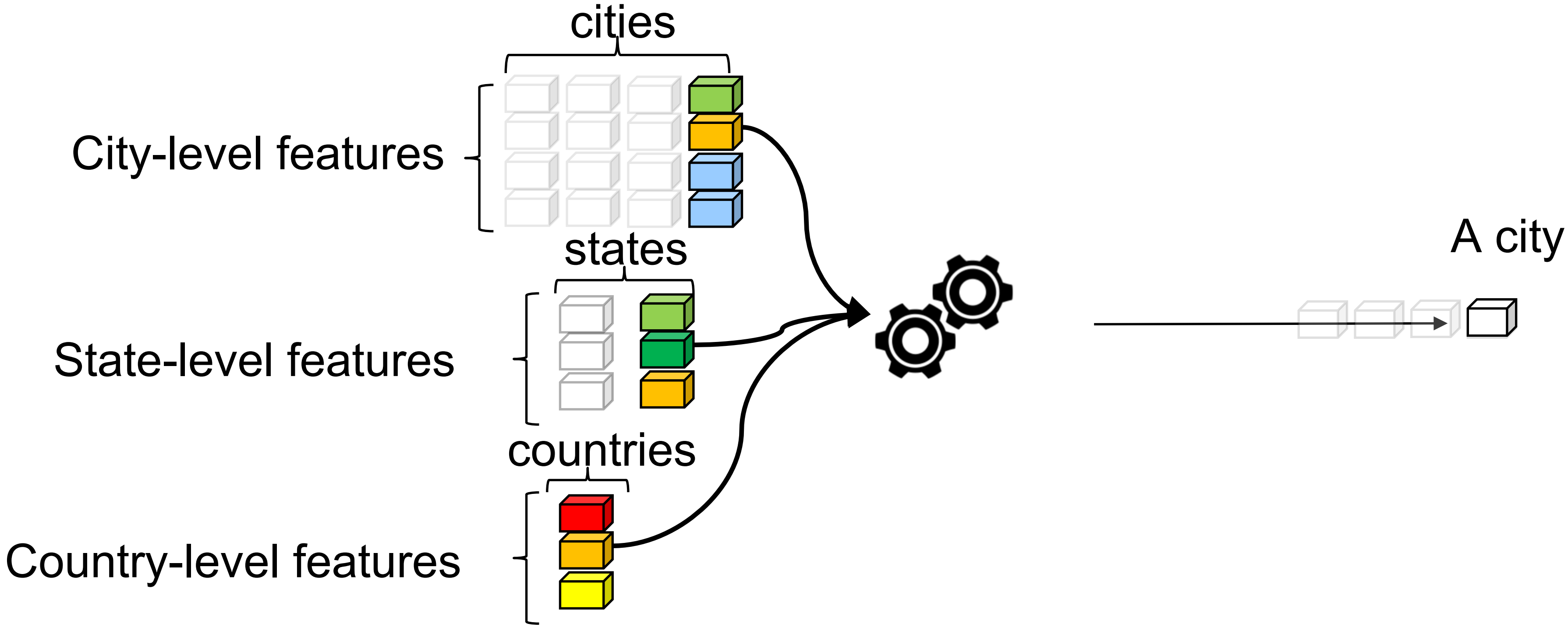
Missing values

Spatial dependency

Spatial multi-resolution

Hierarchical Incomplete Multisource Feature Learning [Zhao et al., KDD'16]

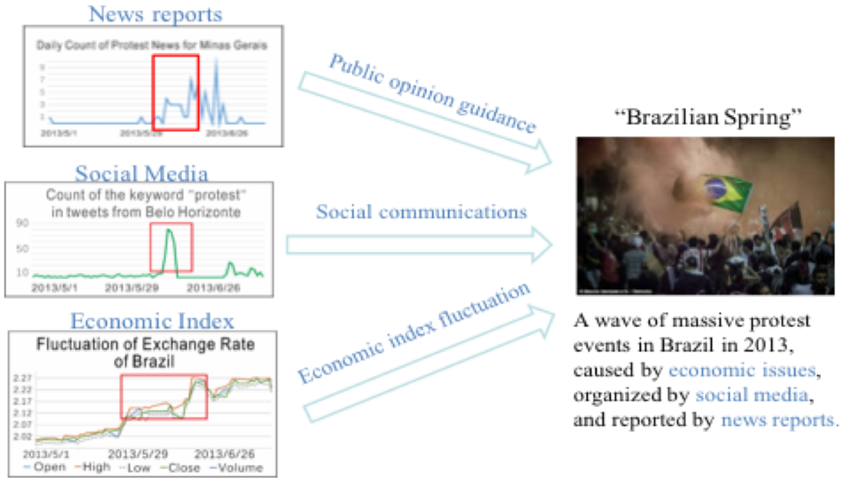
Different feature in different spatial levels



Applications: Multi-source Event Forecasting

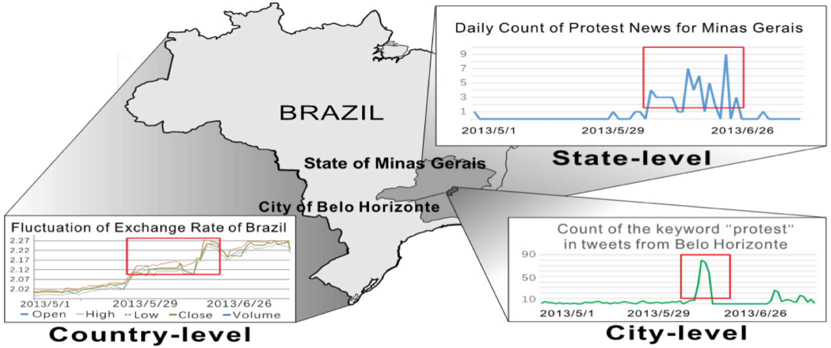
Why multiple data sources?

- Spatiotemporal events are often influenced by different aspects of the society.
- Different data sources complement each other.
- One single source cannot cover all aspects of an event.

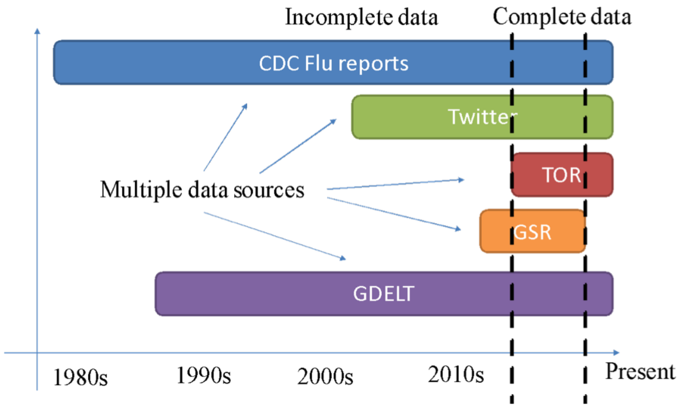


Spatial hierarchy among inputs

Challenge 1: Hierarchical topology
- E.g., country-level, state-level, city-level
- Higher-level features can influence lower-level ones



Challenge 2: Interactive missing values
- Different data sources, different spans
- Need to consider the interactions among different sources.



Hierarchical Incomplete Multisource Feature Learning

Given the multi-source data for a location l at time t, predict whether the event will happen at time τ

$$f : \{\underbrace{X_{t,l_1}}_{\text{city}}, \underbrace{\cdots}_{\text{state,...}}, \underbrace{X_{t,l_N}}_{\text{country}}\} \rightarrow Y_{\tau,l}$$

- Each location has features at multiple levels *l=(l1, l2, ...lN)* E.g., (San Francisco, CA, USA)

Variables are dependent on the variables in their parent level

$$(level-1)\ \ Y_{\tau,l} = \alpha_0 + \sum\nolimits_{i=1}^{|\mathcal{F}_1|} \alpha_i^T \cdot [X_{t,l_1}]_i + \varepsilon$$ city-level

$$(level-2)\ \ \alpha_i = \beta_{i,0} + \sum\nolimits_{j=1}^{|\mathcal{F}_2|} \beta_{i,j}^T \cdot [X_{t,l_2}]_j + \varepsilon_i$$ state-level

$$(level-3)\ \ \beta_{i,j} = W_{i,j,0} + \sum\nolimits_{k=1}^{|\mathcal{F}_3|} W_{i,j,k}^T \cdot [X_{t,l_3}]_k + \varepsilon_{i,j}$$ country-level

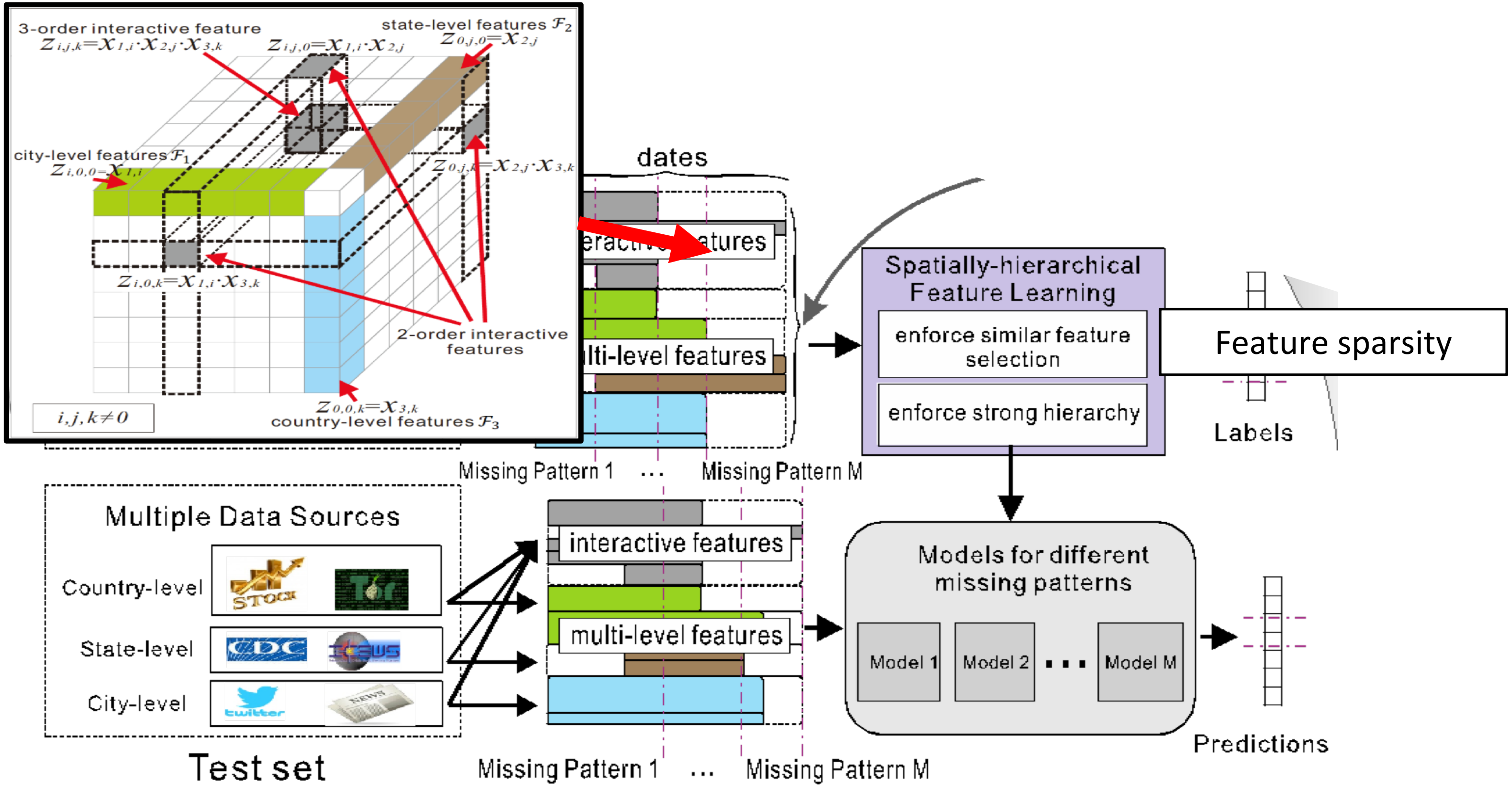
Encode hierarchical feature correlation by nth-order strong hierarchy

$$\Rightarrow\ \ Y_{\tau,l} = \sum_{i=0}^{|\mathcal{F}_1|}\sum_{j=0}^{|\mathcal{F}_2|}\sum_{k=0}^{|\mathcal{F}_3|} W_{i,j,k} \cdot [X_{t,l_1}]_i \cdot [X_{t,l_2}]_j \cdot [X_{t,l_3}]_k + \varepsilon$$

Tensor form:

$$\Rightarrow\ \ Y_{\tau,l} = W \odot Z_{t,l} + \varepsilon$$

Model Framework



(a) hierarchical multi-source interactive feature learning

Dataset

CU: Civil Unrest

News reports

Dataset	Domain	Label sources [1]	#Events
Argentina	CU	Clarín; La Nación; Infobae	1306
Brazil	CU	O Globo; O Estado de São Paulo; Jornal do Brasil	3226
Chile	CU	La Tercera; Las Últimas Notícias; El Mercurio	706
Colombia	CU	El Espectador; El Tiempo; El Colombiano	1196
El Salvador	CU	El Diáro de Hoy; La Prensa Gráfica; El Mundo	657
Mexico	CU	La Jornada; Reforma; Milenio	5465
Paraguay	CU	ABC Color; Ultima Hora; La Nacíon	1932
Uruguay	CU	El País; El Observador	624
Venezuela	CU	El Universal; El Nacional; Ultimas Notícias	3105
U.S.	FLU	CDC Flu Activity Map	1027

FLU: Influenza

Disease surveillance reports

Hierarchical features and missing values

Multi-level sources

	Civil Unrest (yyyy-mm-dd)			Influenza (yyyy-week)		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Geo-level	City	State	Country	State	Region	Country
data sources: training period	Twitter: 2013-04-01∼ 2013-12-31	ICEWS: 2013-04-01∼2013-07-10 2013-10-21∼2013-12-31	CURRENCY: 2013-04-01∼2013-10-21 TOR: 2013-04-01∼2013-10-21	Twitter: 2011-1∼2013-52	ILI-Net: 2009-35∼2013-52	FluSurv-NET: 2009-1∼2011-12 2011-36∼2012-13 2012-36∼2013-52

Block-wise missing values

Multi-source features

Multi-source features

domain	data sources	features
CU	CURRENCY	Open,High,Low,Close
	TOR	Tor daily usage statistics
	ICEWS	CAMEO Codes
	Twitter	982 keywords
FLU	FluSurv-NET	Influenza Hospitalization Ratio by age groups: 0-4 yr, 5-17 yr, 18-49 yr, 50-64 yr, and 65+ yr
	ILI-Net	un/weighted ILI ratios,positive perctentage, #cases of flu types: A(H1N1), A(N1), A(H3), A, B, H3N2v
	Twitter	522 keywords

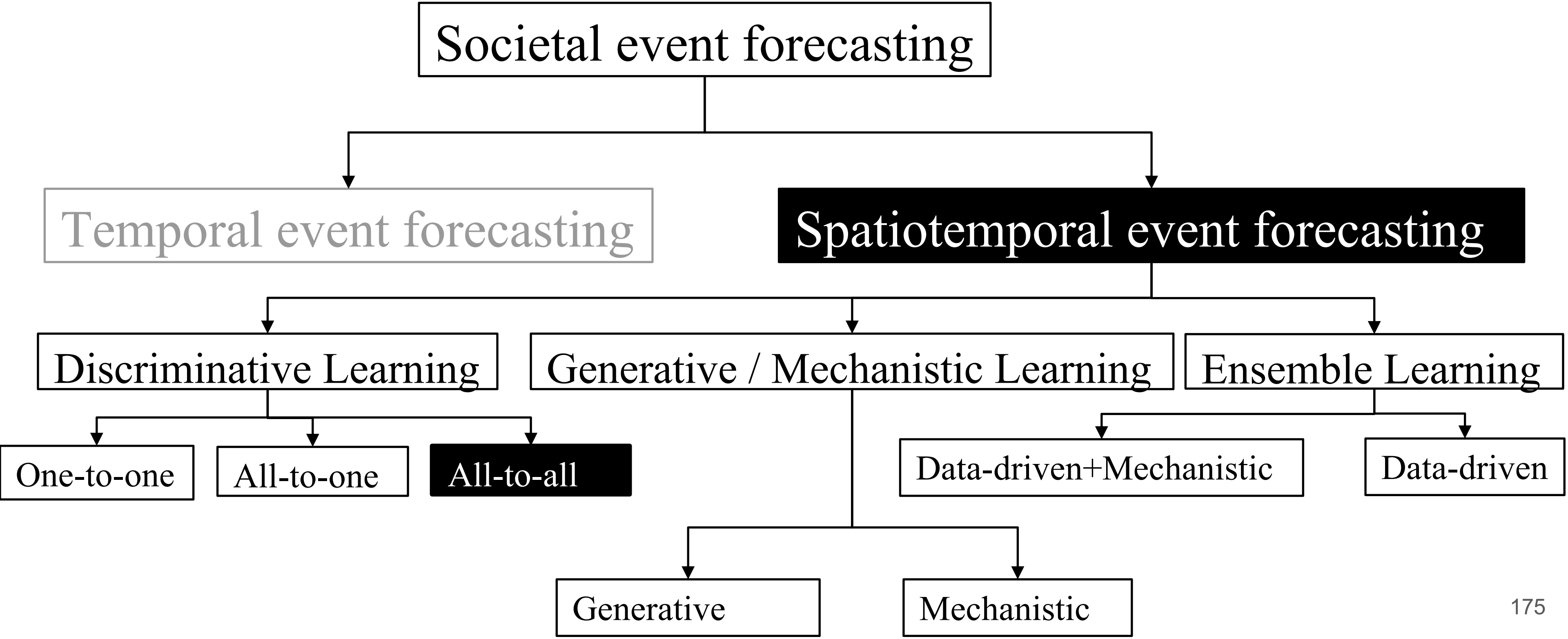
AUC for different missing ratios

(AUC: area under ROC curve)

Method	Argentina	Brazil	Chile	Colombia	El Salvador	Mexico	Paraguay	Uruguay	Venezuela
Missing data ratio (3%)									
LASSO	0.5267	0.7476	0.5624	0.8032	0.3148	0.7823	0.5572	0.4693	0.8073
LASSO-INT	0.5268	0.7191	0.5935	0.7861	0.5269	0.777	0.4887	0.5069	0.7543
iMSF	0.4795	0.4611	0.5033	0.7213	0.5	0.5569	0.4486	0.4904	0.5
MTL	0.3885	0.5017	0.5011	0.4334	0.3452	0.4674	0.4313	0.3507	0.5501
Baseline	0.5065	0.7317	**0.6148**	0.8084	**0.777**	**0.8037**	0.7339	0.7264	**0.7846**
HIML	**0.5873**	**0.8353**	0.5705	**0.8169**	0.7191	0.7973	**0.7478**	**0.8537**	0.7488
Missing data ratio (30%)									
LASSO	0.5035	0.7362	0.588	0.8412	0.3785	0.7896	0.478	0.6749	0.681
LASSO-INT	0.4976	0.6361	0.5912	0.8151	0.3852	0.7622	0.426	0.7177	0.6428
iMSF	0.4797	0.4611	0.4959	0.6845	0.5	0.5569	0.4811	0.4898	0.5
MTL	0.4207	0.5156	0.5023	0.5978	0.3413	0.4666	0.4318	0.347	0.4397
Baseline	0.5012	0.7724	**0.6245**	0.8032	**0.7626**	0.7598	0.738	0.8205	**0.7621**
HIML	**0.5854**	**0.8497**	0.6072	**0.8449**	0.726	**0.7907**	**0.7471**	**0.8576**	0.7378
Missing data ratio (50%)									
LASSO	0.5128	0.7461	0.5301	0.8167	0.3139	0.7552	0.5285	0.5992	0.6678
LASSO-INT	0.504	0.6145	0.5537	0.7339	0.4283	0.7309	0.4745	0.5396	0.6155
iMSF	0.4796	0.4611	0.4962	0.7467	0.4899	0.5488	0.4804	0.487	0.5
MTL	0.5104	0.4818	0.4715	0.65	0.3375	0.4744	0.436	0.3578	0.3839
Baseline	0.5101	0.7717	**0.639**	0.8142	**0.7665**	**0.8079**	0.7324	0.8112	**0.7759**
HIML	**0.5795**	**0.8463**	0.548	**0.8432**	0.7126	0.7892	**0.7477**	**0.856**	0.7176
Missing data ratio (70%)									
LASSO	0.5162	0.6674	0.5947	0.8344	0.2597	0.7485	0.4075	0.2652	0.6699
LASSO-INT	0.4691	0.5557	0.5469	0.7167	0.2116	0.7	0.3808	0.2256	0.6503
iMSF	0.4796	0.4611	0.5503	0.7855	0.5	0.557	0.4795	0.5221	0.5
MTL	0.4128	0.5023	0.5069	0.6195	0.3323	0.4702	0.4283	0.3569	0.6464
Baseline	0.5188	0.7741	**0.6059**	0.8121	**0.7557**	**0.8097**	0.7136	0.72	0.6993
HIML	**0.5484**	**0.7812**	0.3887	**0.8416**	0.7181	0.8001	**0.7146**	**0.8453**	**0.716**

- The proposed HIML performs the best
- Methods considers hierarchical features performs better
- Performance decreases when missing ratio increases
- Methods that can handle incomplete data decreases slower in performance

Taxonomy



Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative Learning
Generative / Mechanistic Learning
Ensemble Learning
One-to-one
All-to-one
All-to-all
Data-driven+Mechanistic
Data-driven
Generative
Mechanistic

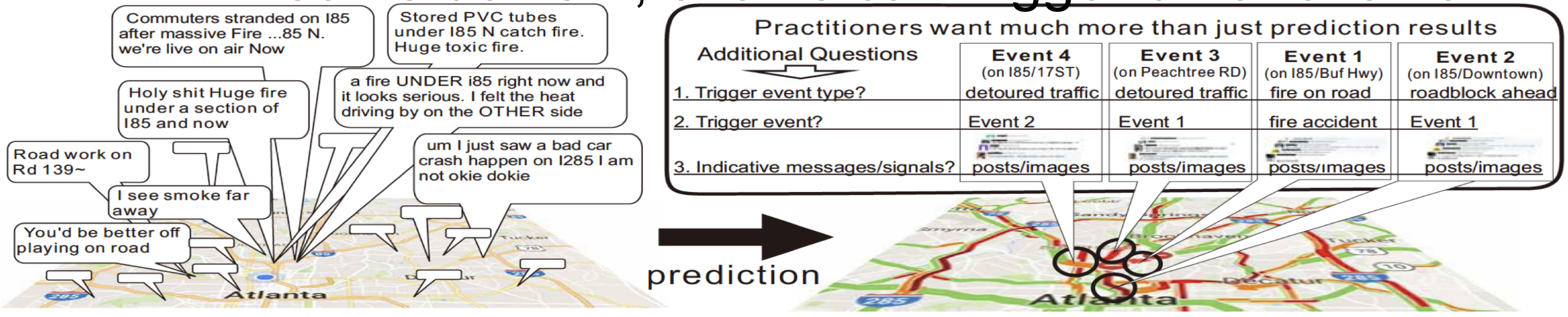
All-to-all models

Use all the locations to forecast for all the locations simultaneously

All the locations (spatial dependency among events)

All the locations (spatial dependency among indicators)

In some domain, events can trigger other events



(a) current crowdsourced observations

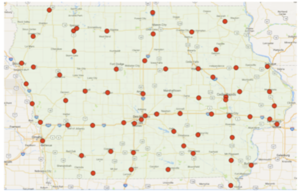
(b) future congestion events

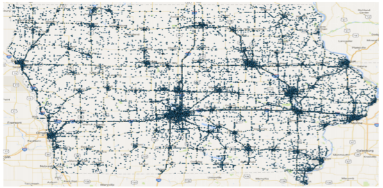
Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data

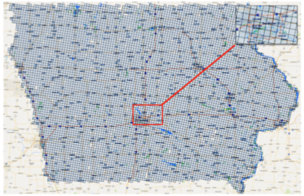
[Zhuoning et al, KDD'18]

Challenges:

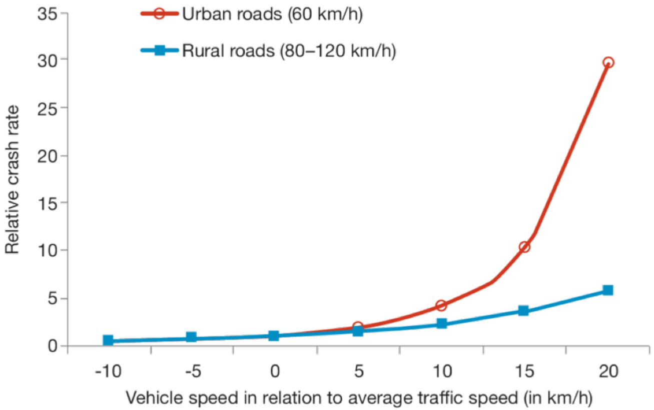
- Existing methods fail to sufficiently utilize all different sources.



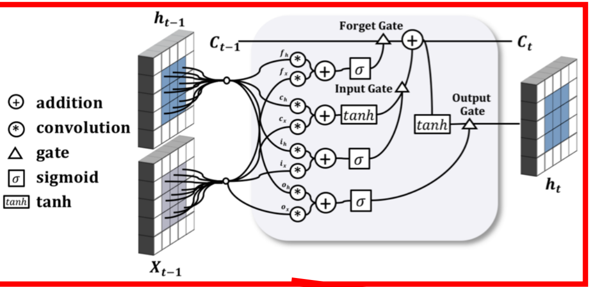


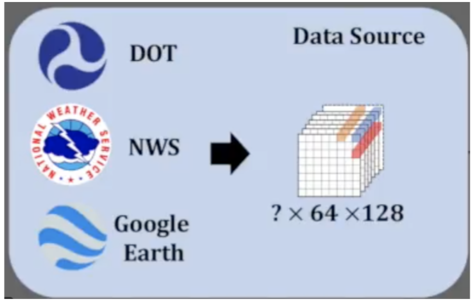


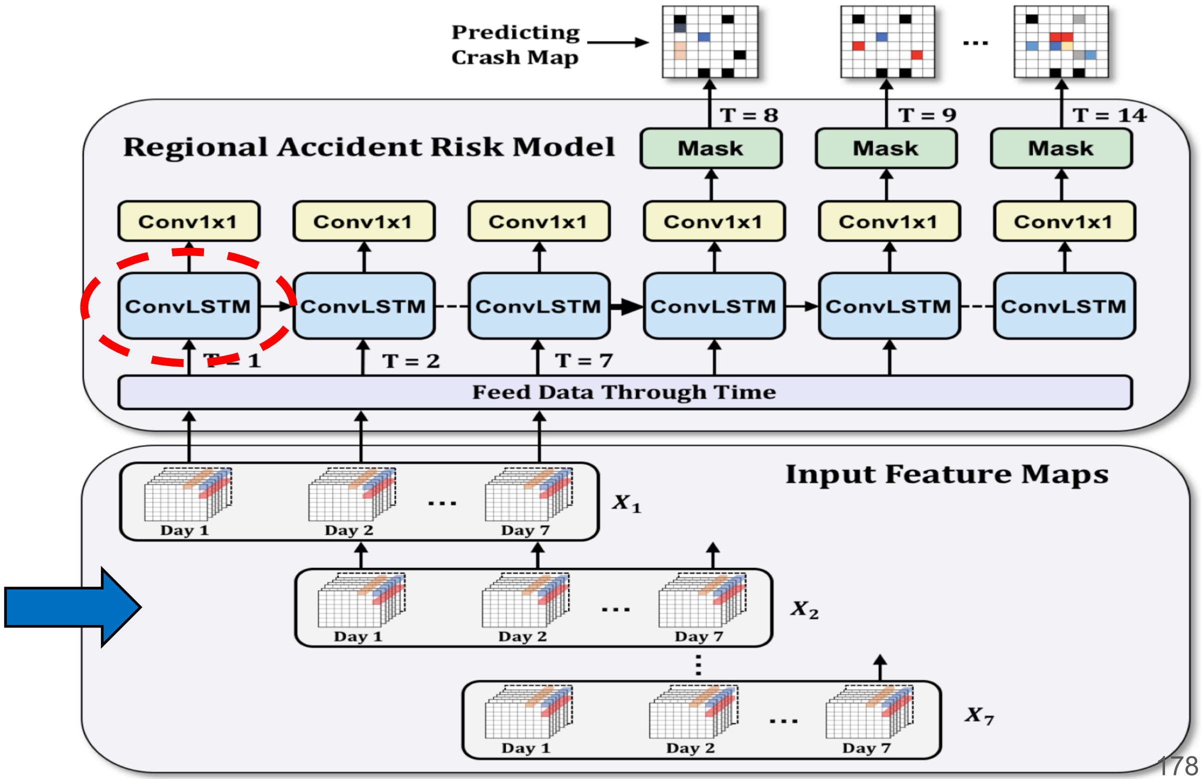
- Spatial heterogeneity
  - e.g., rural vs urban
- Class imbalance
  - a.k.a., accidents are rare



The structure of the regional ConvLSTM model







Model Performance

Table 1: Model Performance

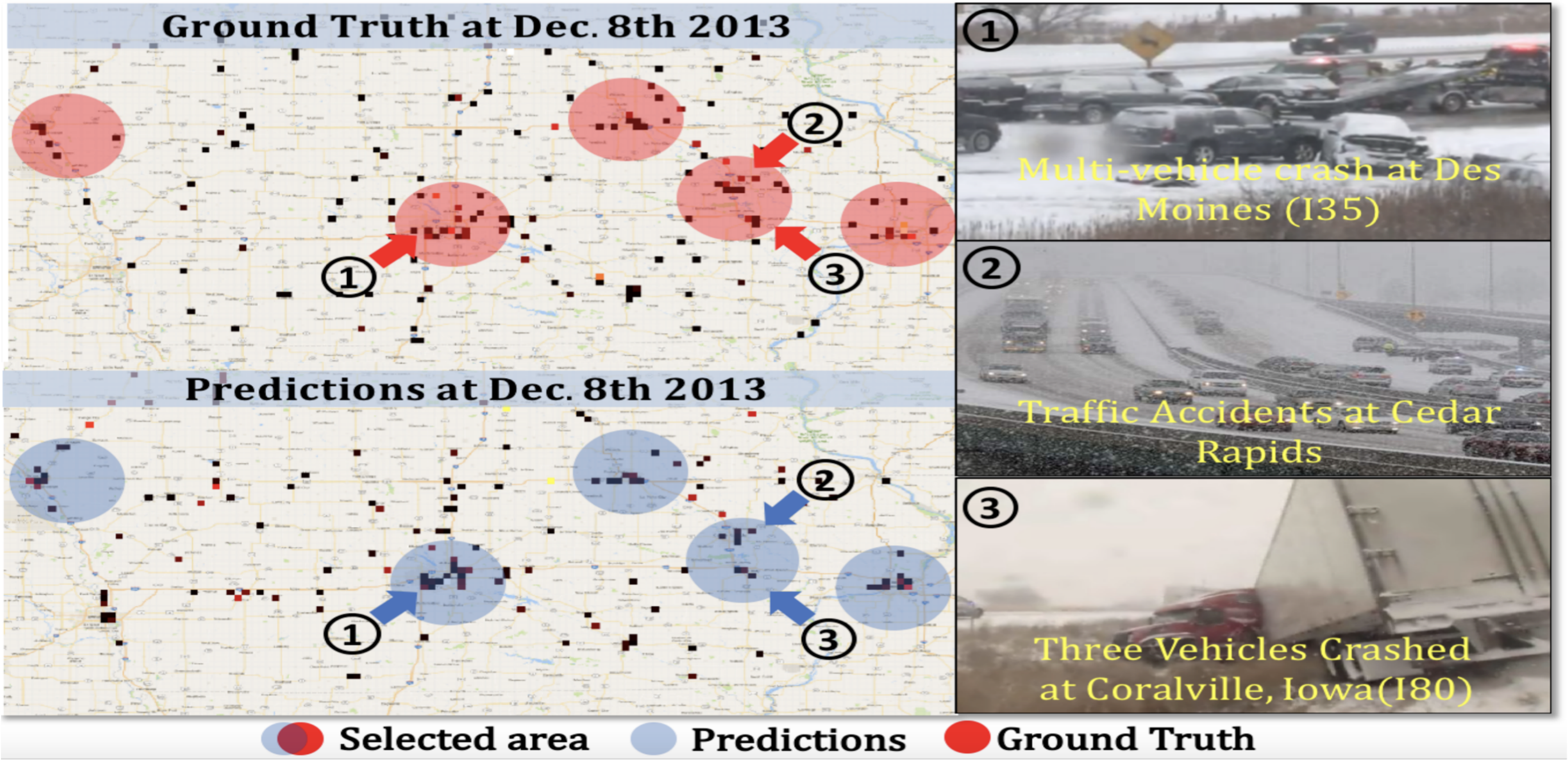
Model	Type-1 Urban			Type-2 Rural			Type-3 Mixed		
	MSE	RMSE	CE	MSE	RMSE	CE	MSE	RMSE	CE
LR(C=0.1)	0.146	0.382	0.051	0.040	0.199	0.002	0.086	0.294	0.014
DTR(depth=30)	0.172	0.415	0.243	0.056	0.237	0.123	0.111	0.334	0.230
DNN(2048x2048)	0.140	0.374	0.033	0.036	0.190	0.023	0.082	0.286	0.011
FC-LSTM(2048x2048)	0.187	0.434	0.419	0.042	0.205	0.419	0.089	0.298	0.001
ConvLSTM (128x128x128x128)	0.117	0.343	0.074	0.037	0.192	0.025	0.077	0.278	0.071
Historical Average (7 years)	0.050	0.224	0.340	0.015	0.121	0.219	0.033	0.181	0.295
Hetero-ConvLSTM (128x128x128x128)	**0.021**	**0.144**	**0.014**	**0.006**	**0.078**	**0.001**	**0.013**	**0.116**	**0.010**

Table 2: Impact of Feature Groups

Model	Type-1 Urban			Type-2 Rural			Type-3 Mixed			All Regions		
	MSE	RMSE	CE	MSE	RMSE	CE	MSE	RMSE	CE	MSE	RMSE	CE
N	0.120	0.346	0.089	0.063	0.251	0.212	0.082	0.286	0.068	0.049	0.222	0.047
N+RW+RA	0.126	0.356	0.073	0.038	0.195	0.046	0.076	0.276	0.087	0.056	0.237	0.074
N+RW+RA+V+RC	0.123	0.351	0.127	0.039	0.199	0.006	0.100	0.316	0.256	0.049	0.221	0.037
N+RW+RA+V+RC+G	0.148	0.384	0.247	0.038	0.194	0.039	0.080	0.283	0.050	0.048	0.219	0.043
N+RW+RA+V+RC+G+CL	0.118	0.344	0.075	0.046	0.216	0.100	0.082	0.286	0.018	0.048	0.220	0.030
N+RW+RA+V+RC+G+CL+E	0.117	0.343	0.074	0.037	0.192	0.025	0.077	0.278	0.071	0.049	0.222	0.026

Using heterogeneous data sources is advantageous!

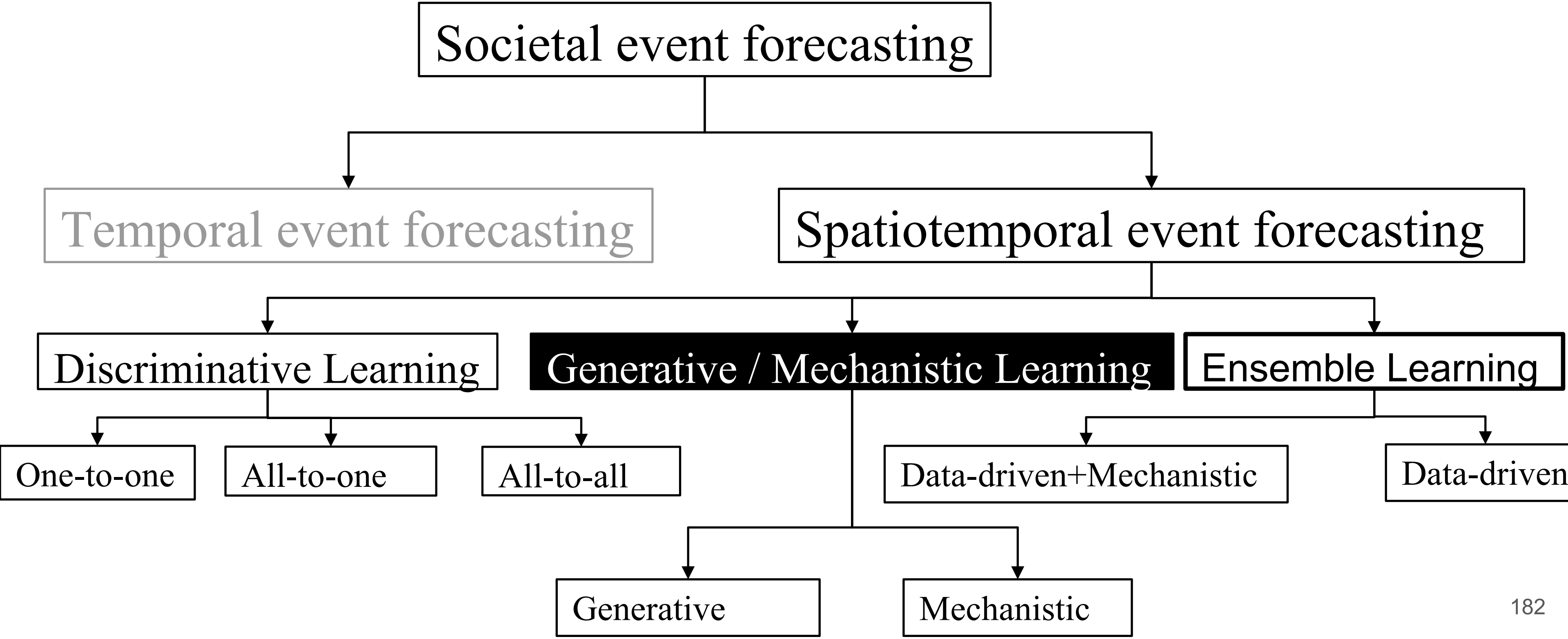
Case study of traffic accidents



References

- Matthew S. Gerber. Predicting crime using Twitter and kernel density estimation. Decision Support Systems. Volume 61, May 2014, Pages 115-125.
- Wu, Congyu, and Matthew S. Gerber. "Forecasting Civil Unrest Using Social Media and Protest Participation Theory." *IEEE Transactions on Computational Social Systems* 5, no. 1 (2018): 82-94.
- Zhuoning Yuan, Xun Zhou, Tianbao Yang. Hetero-ConvLSTM: A Deep Learning Approach to Traffic Accident Prediction on Heterogeneous Spatio-Temporal Data. In 24th ACM SIGKDD International Conference on Knowledge Discovery from Data (KDD), 2018 (Accepted).
- Liang Zhao, Qian Sun, Jieping Ye, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. "Multi-Task Learning for Spatio-Temporal Event Forecasting." in Proceedings of the 21st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2015),Sydney, Australia, pp. 1503-1512, Aug 2015.
- Yuyang Gao, Liang Zhao, Lingfei Wu, Yanfang Ye, Hui Xiong, Chaowei Yang. Incomplete Label Multi-task Deep Learning for Spatio-temporal Event Subtype Forecasting.Thirty-third AAAI Conference on Artificial Intelligence (AAAI 2019), Hawaii, USA, Feb 2019, to appear.
- Yuyang Gao and Liang Zhao. Incomplete Label Multi-Task Ordinal Regression for Spatial Event Scale Forecasting. Thirty-Second AAAI Conference on Artificial Intelligence (AAAI 2018), New Orleans, US, Feb 2018, pp. 2999-3006, New Orleans, US, Feb 2018.
- Liang Zhao, Jieping Ye, Feng Chen, Chang-Tien Lu, Naren Ramakrishnan. “Hierarchical Incomplete Multisource Feature Learning for Spatiotemporal Event Forecasting.” in Proceedings of the 22st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2016), San Francisco, California, pp. 2085-2094, Aug 2016.
- Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. "Multi-resolution Spatial Event Forecasting in Social Media." in Proceedings of the IEEE International Conference on Data Mining (ICDM 2016), pp. 689-698, Barcelona, Spain, Dec 2016.

Taxonomy



Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative Learning
Generative / Mechanistic Learning
Ensemble Learning
One-to-one
All-to-one
All-to-all
Generative
Mechanistic
Data-driven+Mechanistic
Data-driven

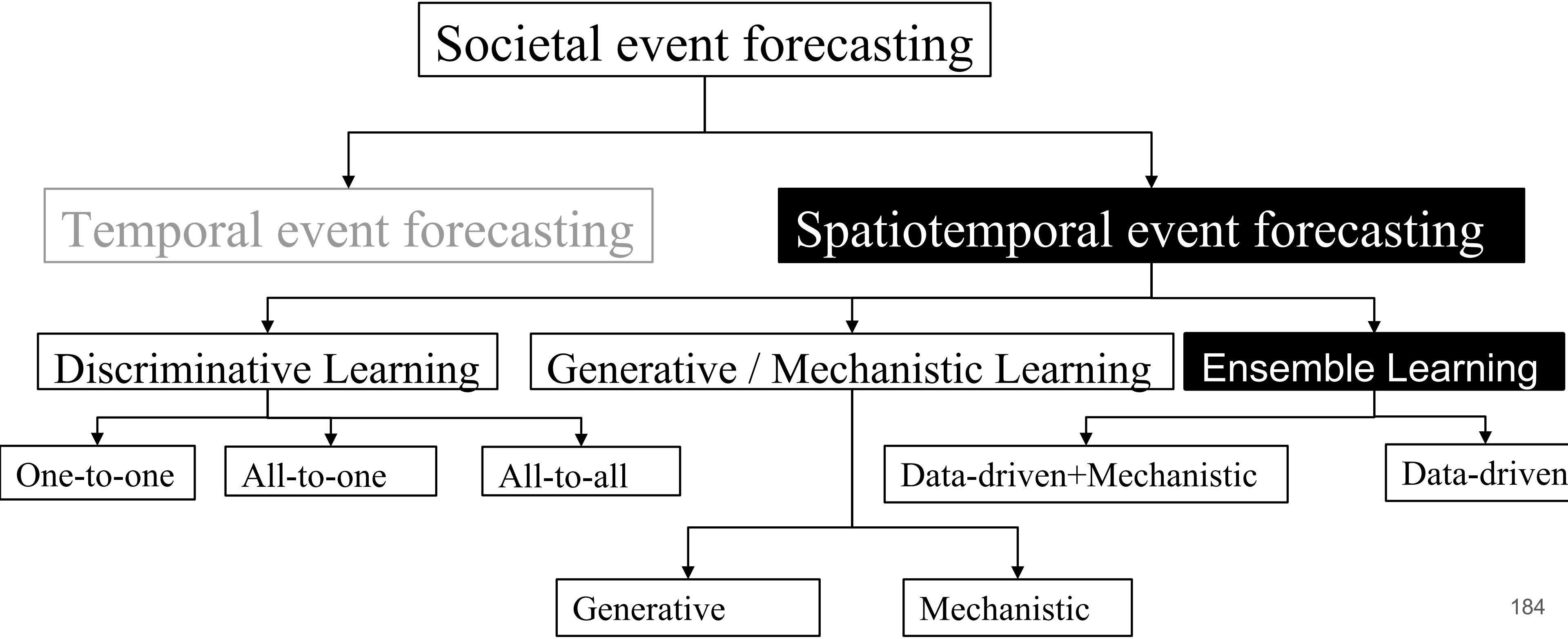
References

•Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. "Spatiotemporal Event Forecasting in Social Media." in Proceedings of the SIAM International Conference on Data Mining (SDM 2015), Vancouver, BC, pp. 963-971, Apr-May 2015.

•Liang Zhao, Feng Chen, Chang-Tien Lu, and Naren Ramakrishnan. "Online Spatial Event Forecasting in Microblogs.", ACM Transactions on Spatial Algorithms and Systems (TSAS), Volume 2 Issue 4, Acticle No. 15, pp. 1-39, November 2016.

•Fang Jin, Rupinder Khandpur, Nathan Self, Edward Dougherty, Sheng Guo, Feng Chen, B. Aditya Prakash, Naren Ramakrishnan. Modeling Mass Protest Adoption in Social Network Communities using Geometric Brownian Motion, in Proceedings of the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'14), Aug 2014.

Taxonomy


Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative Learning
Generative / Mechanistic Learning
Ensemble Learning
One-to-one
All-to-one
All-to-all
Data-driven+Mechanistic
Data-driven
Generative
Mechanistic

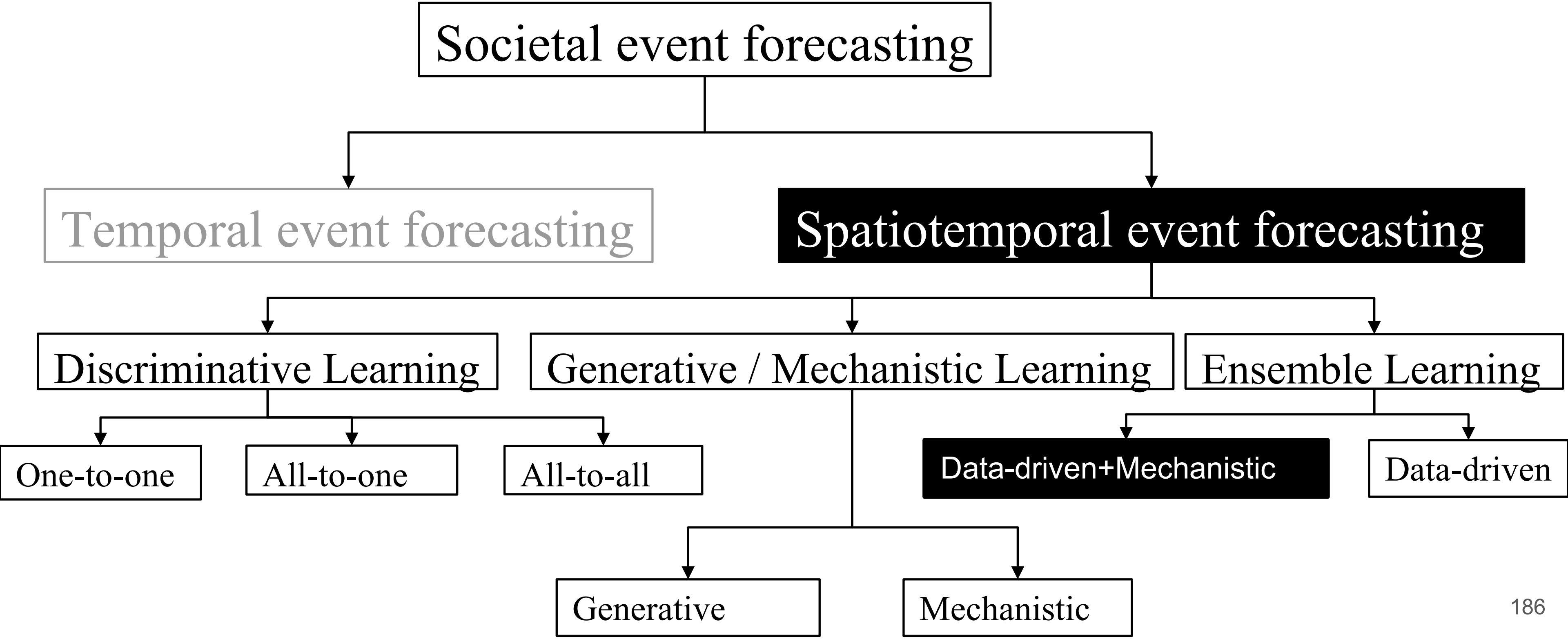
Ensemble Learning for Spatiotemporal Event Forecasting

- Due to the complexity of the societal phenomena.
  - Each data source may only cover one part
  - Each model may only explain a portion of the truth
  - Some truth are unobservable.

Ensemble learning:

- Leverage the complementary strength of different models
- Sufficiently utilize different data sources in modeling different phenomena

Taxonomy


Societal event forecasting
Temporal event forecasting
Spatiotemporal event forecasting
Discriminative Learning
Generative / Mechanistic Learning
Ensemble Learning
One-to-one
All-to-one
All-to-all
Data-driven+Mechanistic
Data-driven
Generative
Mechanistic

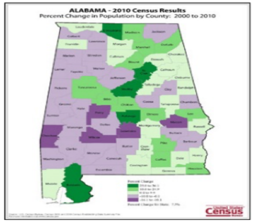
SimNest: Social Media Nested Epidemic Simulation via Online Semi-supervised Deep Learning
[Zhao, et al., ICDM'15, Geoinformatica, 2019]

- Goal: Utilize social media data and disease mechanism to model the underlying influenza epidemics progression.
- Model characteristics:
    - Ensembles of Data-driven and Mechanistic Models
    - Online Learning

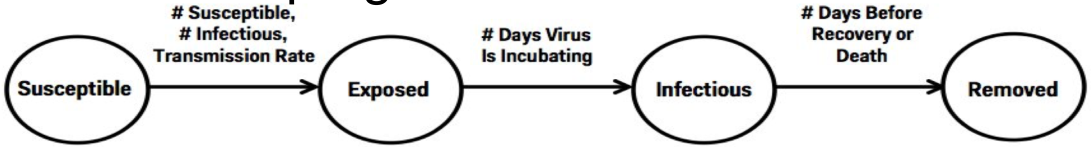
Epidemics Modeling (Category 1): Computational Epidemiology

1. Model the following mechanisms

a. Demographics and social contact network



b. Disease progression: SEIR model



c. Interventions

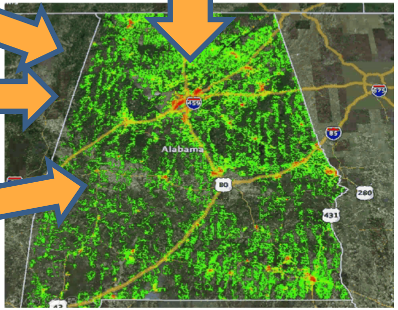
School Closure

Vaccination

Isolation

2. Tune parameters against surveillance data

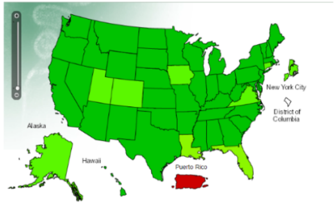
3. Run simulation model



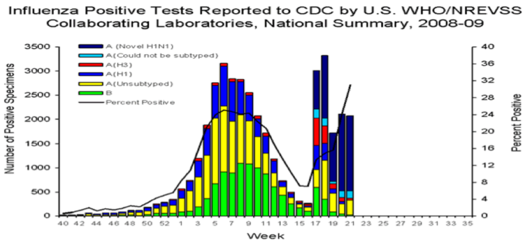
Epidemics Modeling (Category 1): Computational Epidemiology

- Challenges
  - Challenge 1: Coarse-grained surveillance data

State-wise:

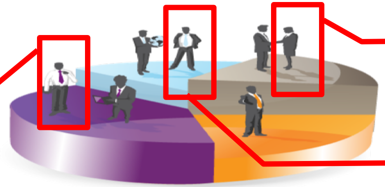


Week-wise:



  - Challenge 2: Dynamics of contact networks

This year much more people get flu shot



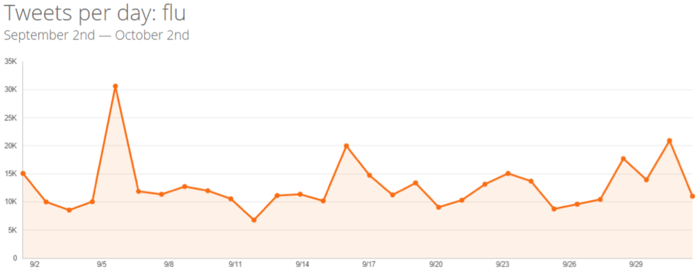
Peter moved out to another city because he lost job.

Jim is suddenly on vacation.

  - Challenge 3: Poor timeliness
    - Surveillance data is at least one week behind.

Epidemics Modeling (Category 2): Data-driven Techniques on Social Media

- Fast monitoring real-time epidemics



- Spatially & Temporally fine-grained
- No delay

1. Identify the response to flu

2. Identify the individual's disease progression

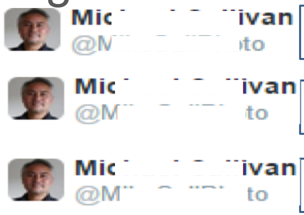
- Individual-wise health condition mining

in flu season, What Peter will do?

Avoid crowds

Get flu shot

…

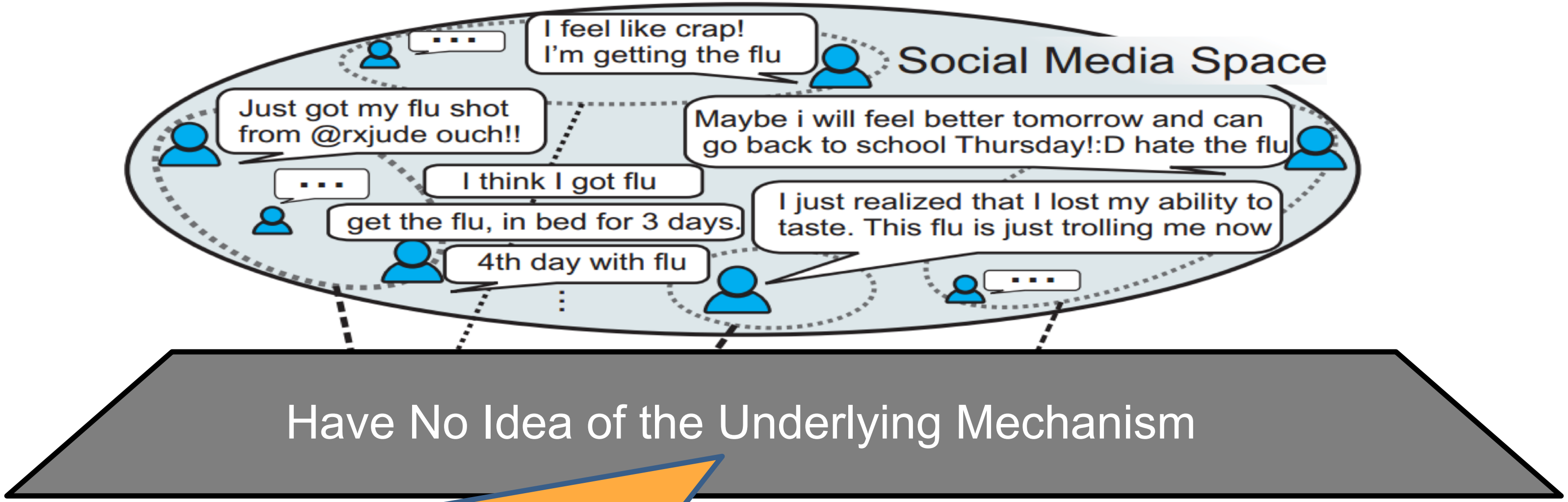


Feel I'm getting flu

3rd day in the bed

Maybe it's time back to work

Epidemics Modeling (Category 2): Data-driven Techniques on Social Media



Challenge: Real Mechanism is hidden to social media

What is the real disease contact network?

What is diffusion process of epidemics?

What is the consequence if someone took vaccine?

Any influence on infectivity if someone has summer holiday?

Motivations

Computational Epidemiology

- Advantages:
  - Mechanism on disease progression
  - Mechanism on disease diffusion
  - Consideration on interventions
- Drawbacks:
  - Temporally coarse-grained
  - Spatially coarse-grained
  - Poor dynamics in social contact network
  - One week delay

+

Social Media Mining

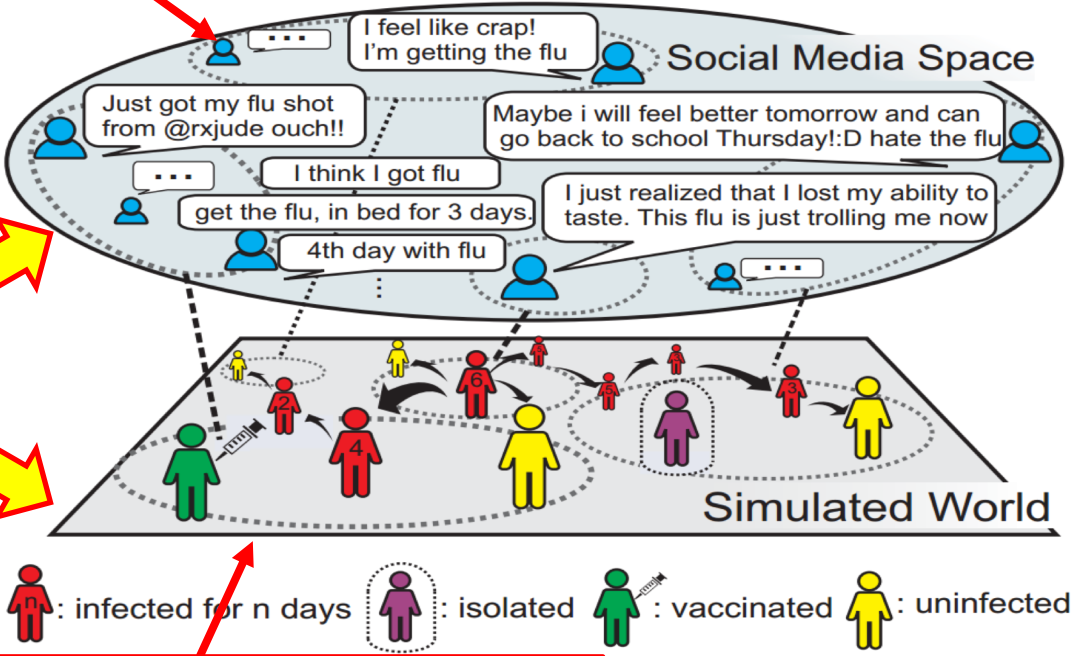
- Drawbacks:
  - No mechanism on disease progression
  - No mechanism on disease diffusion
  - No consideration on interventions
- Advantages:
  - Temporally fine-grained
  - Spatially fine-grained
  - Change in social contact network is observable in real time
  - No time delay

Combine

Idea

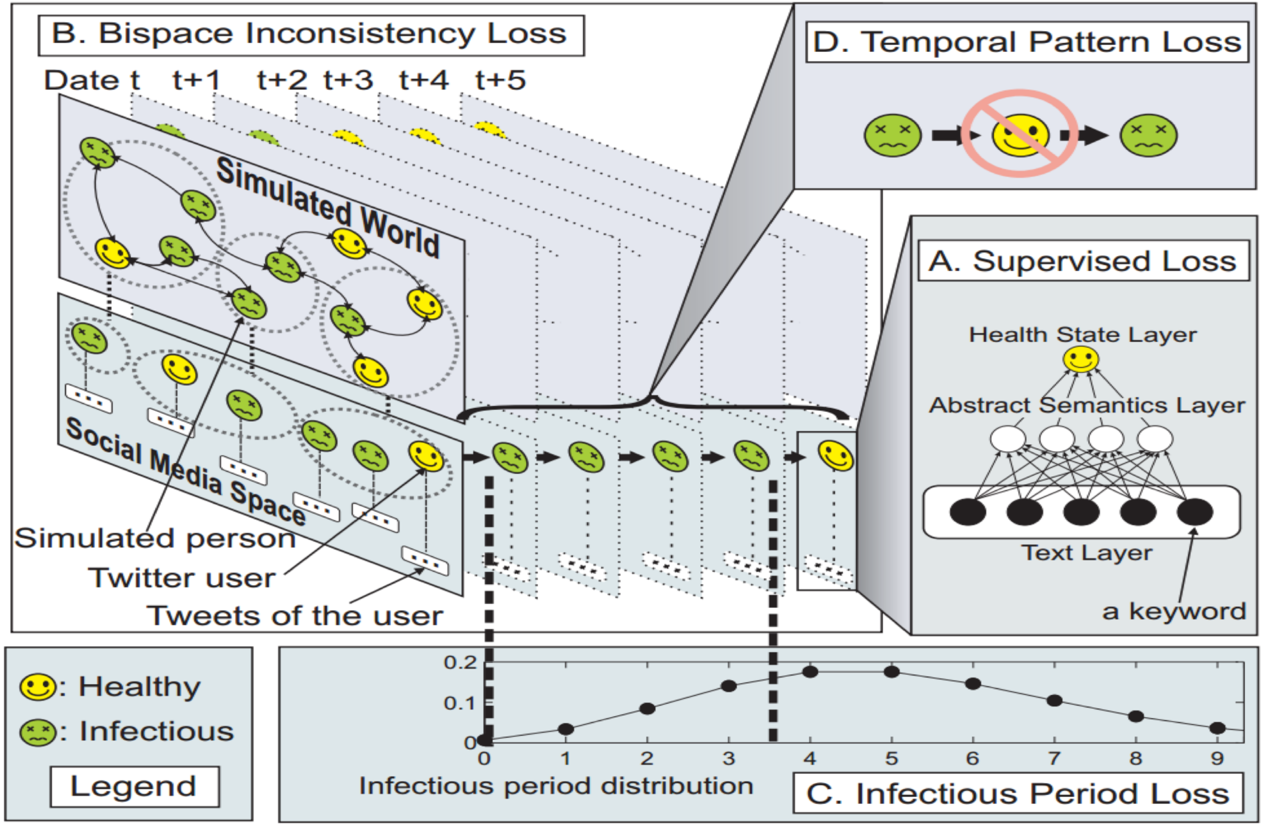
Timely and fine-grained observations

Combine



Mechanisms of epidemics diffusion

Model: Overview



Our objective:
Minimize loss

$$\min \mathcal{L} = \min \mathcal{L}_A + \mathcal{L}_B + \mathcal{L}_C + \mathcal{L}_D$$

The Proposed Model

Tweets of User u at Time t

- Learn a mapping: $f_W(X_{u,t}) : \boxed{X_{u,t}} \rightarrow \boxed{Y_{u,t}}$ Infectious (1) or not (0)
- Minimize supervised Loss: $$\mathcal{L}_{\mathrm{A}} = \min_W \sum_u^{\mathcal{U}_1} \sum_t^{\mathcal{T}} \| f_W(X_{u,t}) - Y_{u,t} \|^2$$

Deep neural networks

- Online training by alternating optimization

Health stage of Person v at Time t in simulated world

- Maximize the likelihood of infectious period distribution:

Infectious period is Gaussian distributed

$$[\sum_t^{\mathcal{T}} f_W(X_{u,t})] = d_u \sim p_I(u) = \mathcal{N}(u|\mu_I, \sigma_I)$$

- Health stage should be consecutive:

$$\mathcal{L}_{\mathrm{D}} = \min_W \sum_u^{\mathcal{U}_2} \sum_t^{\mathcal{T}} \| f_W(X_{u,t}) - f_W(X_{u,t+1}) \|^2$$

Experiments: Dataset

- Dataset:
  - Twitter: Year 2011 ~ Year 2014 in the US.
  - Training set: Aug 1 2011 ~ Jul 31 2012.
  - Test set: Aug 1 2012 ~ Jul 31 2014.

Table I: Twitter data set and demographics

	Demographics		Twitter	
state	population size	#connections	#tweets	#users
CT	3,518,288	175,866,264	9,513,741	10,257
DC	599,657	19,984,180	12,148,925	7,015
MA	6,593,587	332,194,314	19,785,147	15,005
MD	5,699,478	285,159,648	20,754,218	19,758
VA	7,882,590	407,976,012	15,899,713	14,302

Connecticut (CT), Massachusetts (MA), Maryland (MD), and Virginia (VA), and the District of Columbia (DC)

Experiments: Label and Metrics

- Label:
  - influenza statistics reported by the Centers for Disease Control and Prevention (CDC).
  - The CDC weekly publishes the percentage of the number of physician visits related to influenza-like illness (ILI) within each major region in the United States.
- Metrics:
  - Lead time: How much time the output is ahead of the input.
  - Mean squared error (MSE)
  - Pearson correlation
  - P-value
  - Peak time error: Error of the predicted time of peak value

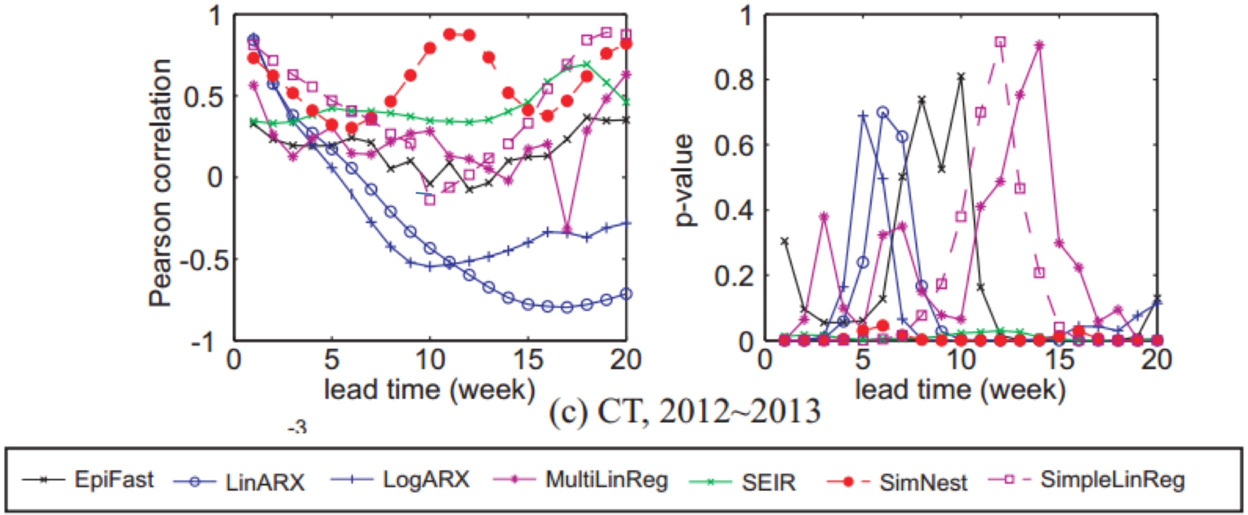
Experiments: Comparison Methods

- social media mining methods:
  - Linear Autoregressive Exogenous model (LinARX)
  - Logistic Autoregressive Exogenous model (LogARX)
  - Simple Linear Regression model (simpleLinReg)
  - Multi-variable linear regression model (multiLinReg)
- computational epidemiology methods:
  - SEIR
  - EpiFast
- Detailed parameter settings:
  - See here: http://people.cs.vt.edu/liangz8/materials/papers/SimNestAddon.pdf

Influenza Epidemic Forecasting Performance

Training set: Tweets in Aug 2011 ~ Jul 2012 in the US.
Test set: Tweets Aug 2012 ~ Jul 2014 in the US.

Label set: CDC surveillance data



P-value: likelihood that the null hypothesis is true.

Pearson correlation: Strength of linear relation

Lead time: How much time the output is ahead of the input.

Conclusion and Future Directions – Spatio-Temporal Event Forecasting

- Spatial-temporal event forecasting methods are typically designed based on the modeling of complex relationships of past and future events from both the geographical and temporal dimensions.
- Future directions
  - Spatial dependencies among the events
  - Bridge the event forecasting and decision making
    - Interpretability, uncertainty, robustness
  - Bridge the communities between data scientists and social scientists.
  - World common sense model that build a unified world surrogate model for event synthesis.

References

- Liang Zhao, Jiangzhuo Chen, Feng Chen, Wei Wang, Chang-Tien Lu, and Naren Ramakrishnan. "SimNest: Social Media Nested Epidemic Simulation via Online Semi-supervised Deep Learning." in Proceedings of the IEEE International Conference on Data Mining (ICDM 2015), Atlantic City, NJ, pp. 639-648, Nov 2015.
- Naren Ramakrishnan, Patrick Butler, Sathappan Muthiah, et al.. "'Beating the news' with EMBERS:forecasting civil unrest using open source indicators." In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD 2014), industrial track, pp. 1799-1808. ACM, 2014.
- Sathappan Muthiah, Patrick Butler, Rupinder Paul Khandpur, et al. "EMBERS at 4 years:Experiences operating an Open Source Indicators Forecasting System." in Proceedings of the 22st ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2016), pp. 205-214, San Francisco, California, Aug 2016.
- Liang Zhao, Jiangzhuo Chen, Feng Chen, Fang Jin, Wei Wang, Chang-Tien Lu, and Naren Ramakrishnan. Online Flu Epidemiological Deep Modeling on Disease Contact Network. GeoInformatica (impact factor: 2.392), to appear, 2019.

Acknowledgements

Liang Zhao
Assistant Professor
Department of Information Science and Technology
George Mason University

Chang-Tien Lu
Professor
Department of Computer Science
Virginia Tech

Naren Ramakrishnan
Thomas L. Phillips Professor of Engineering
Department of Computer Science
Virginia Tech

Thank you
Q&A

Feel free to email questions or suggestions to
yue.ning@stevens.edu or lzhao9@gmu.edu