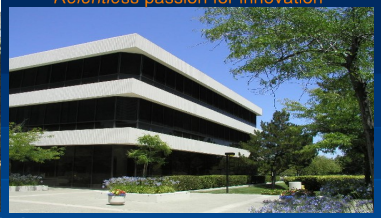
Orchestrating a brighter world NEC

【NEC Group Internal Use Only】

# Efficient Multiple Instance Learning

Dongkuan Xu
The Pennsylvania State University

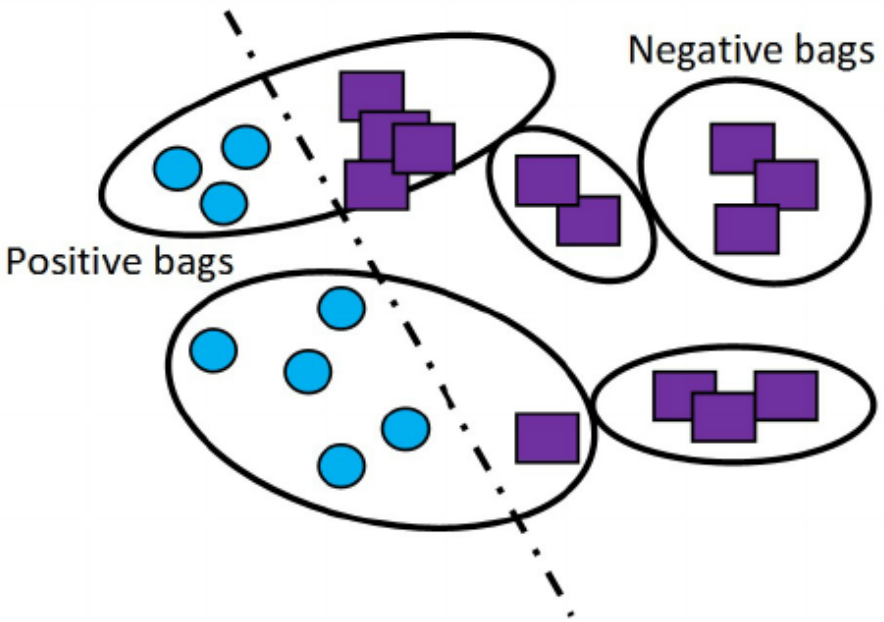
NEC Laboratories America
*Relentless* passion for innovation



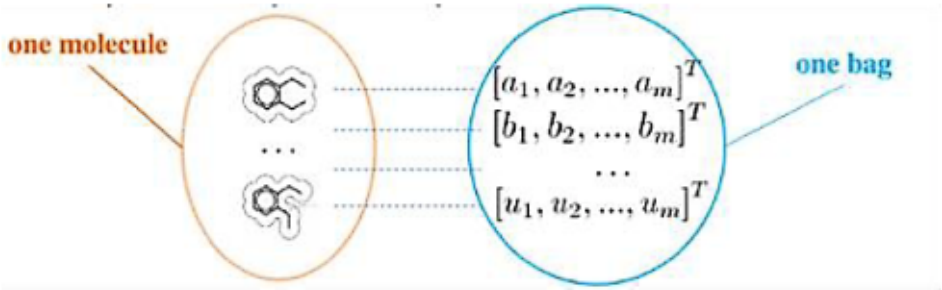
www.nec-labs.com

# Background

- Multiple Instance Learning (MIL)
- Diagram
  - Weakly-Supervised Classification
  - ***Bag*** v.s. ***Instance***
  - **Uncertainty**: Label available only for bags



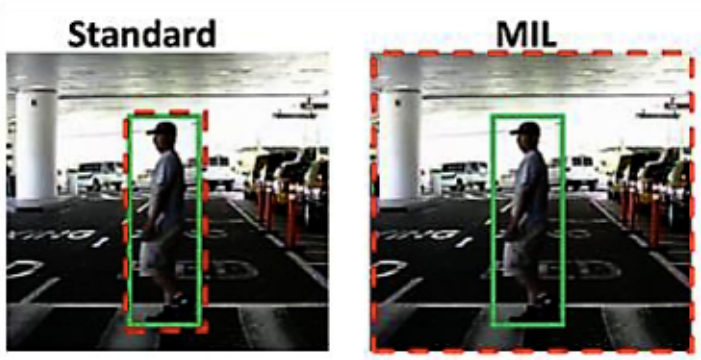
# Various Applications




Drug activity prediction

Image categorization


Text categorization




Pedestrian detection

# Time-Series Application

## Modeling Precursors for Event Forecasting via Nested Multi-Instance Learning[C]//(KDD)2016

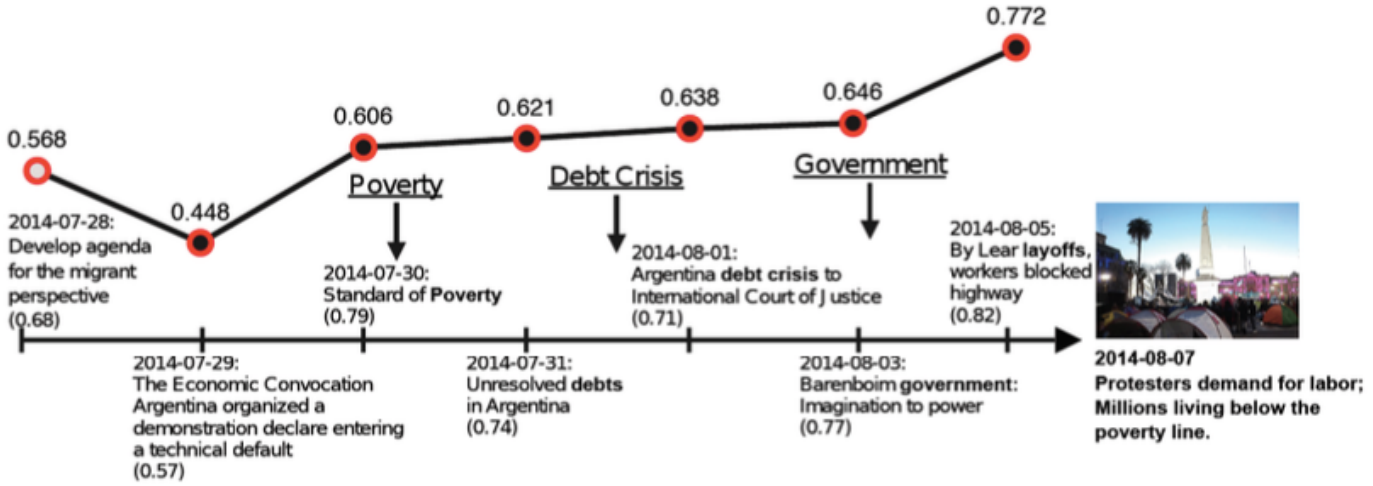



Figure 1: Precursor story line for a protest event in Argentina. The x-axis is the timeline. The dots above with numbers are the probabilities for each day that the model generated for the target event. Each precursor document is titled in the timeline.

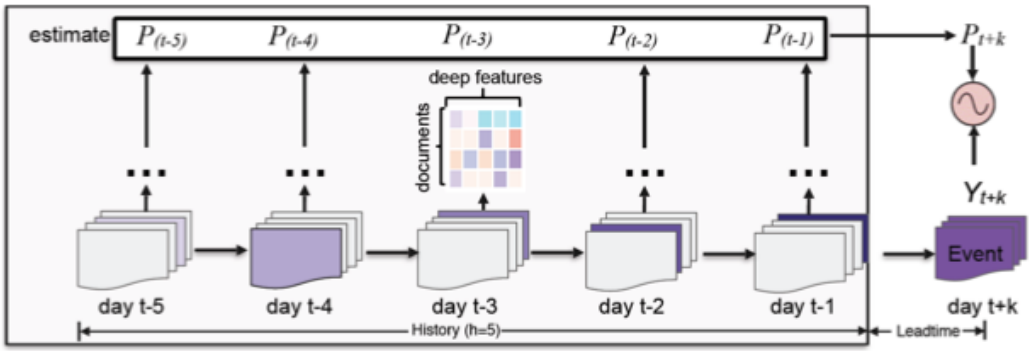
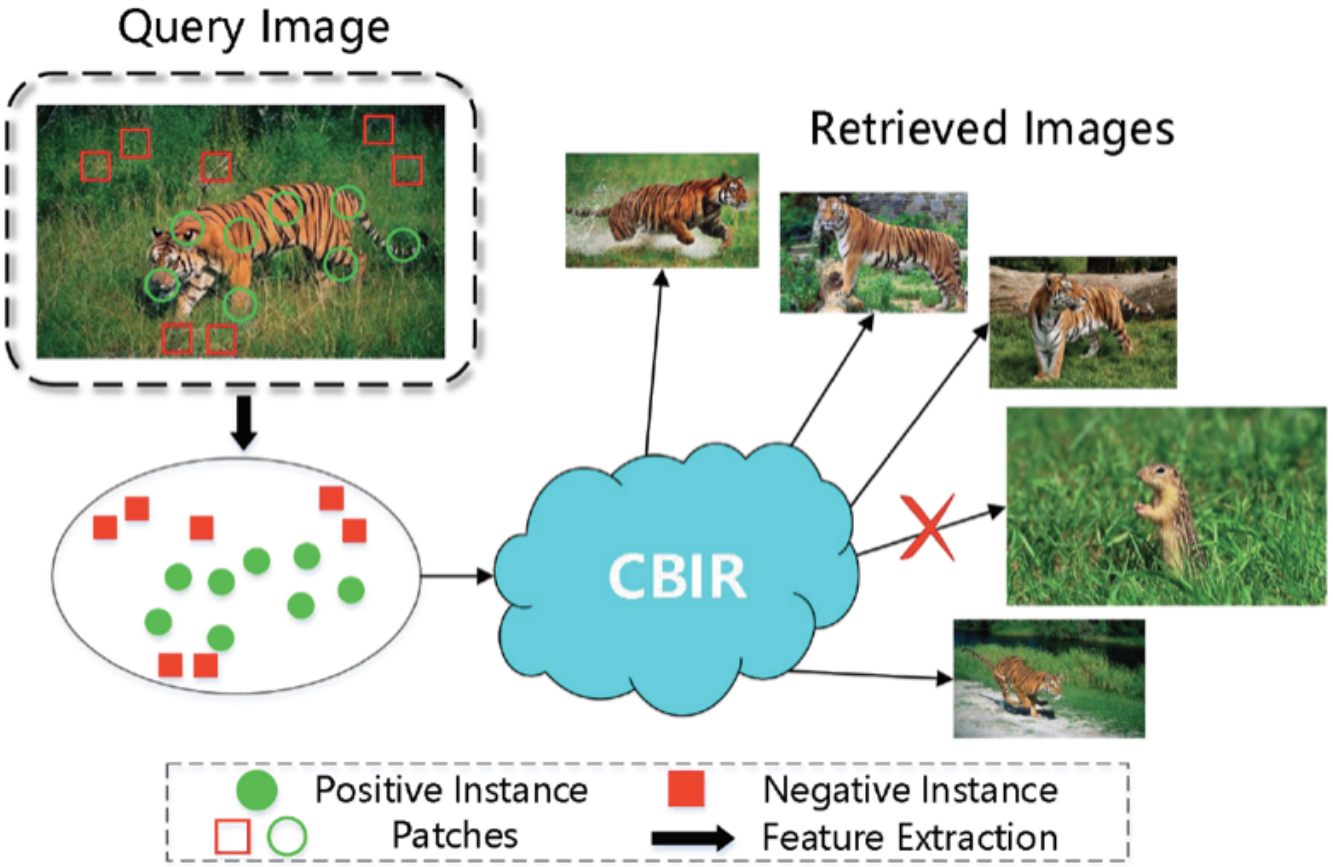



Figure 2: Overview of proposed approach to forecasting and precursor discovery.

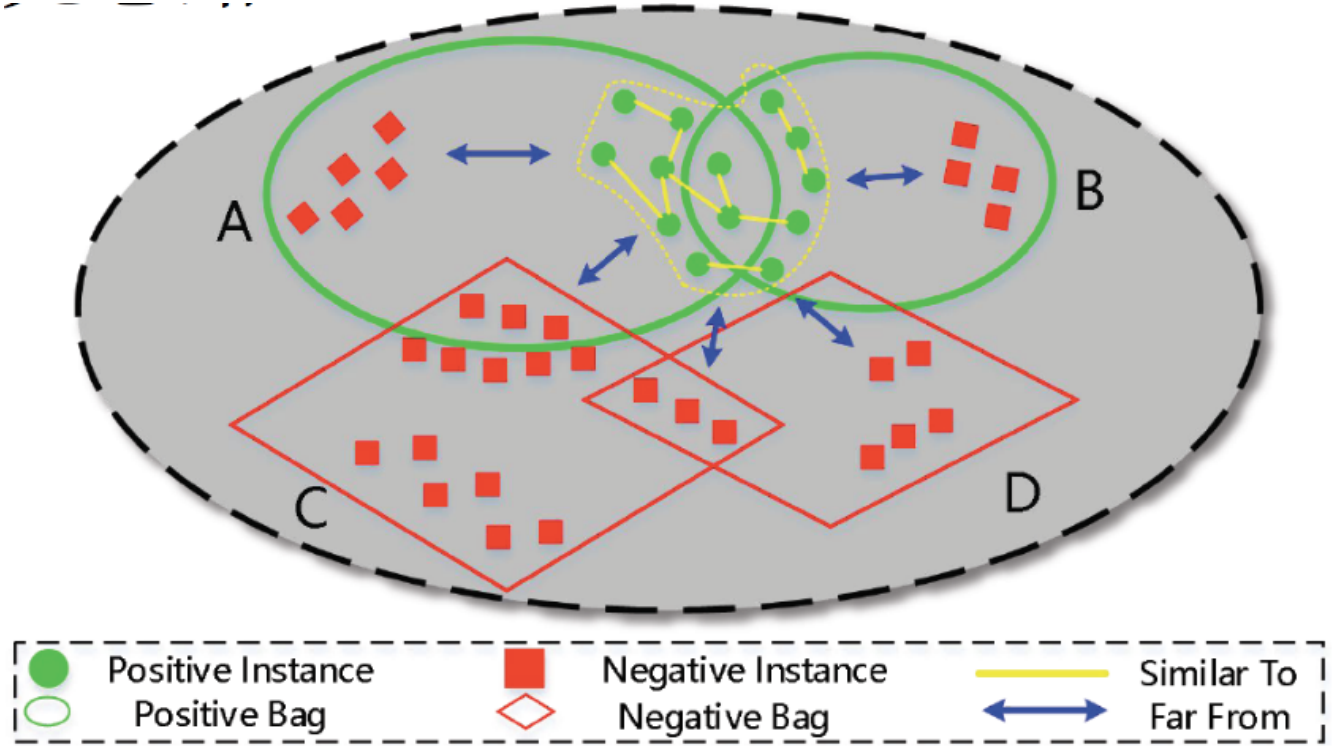
# (1) MIL via Positive Instance Detection

- Motivation: “True” positive instance
- Content-based image retrieval (CBIR)



# (1) MIL via Positive Instance Detection

- Key Idea of Solution
  - **Similarity**
  - **Consistency**
  - **Discrimination**



# (1) MIL via Positive Instance Detection

- Objective Function

$$\max_{\mathcal{L}} \sum_{(x_i,x_j),x_k \in PCP} \alpha \mathcal{S}(x_i,x_j) + \mathcal{C}(x_i,x_j) + \beta \mathcal{D}(x_k)$$

$$\mathcal{S}(x_i,x_j) = \begin{cases} \frac{1}{\sum_{k=1}^{K} |\varphi(k,i) - \varphi(k,j)|} & if\ \Upsilon_{k,i} > \tau\ ,\ \Upsilon_{k,j} > \tau \\ 0 & otherwise \end{cases}$$

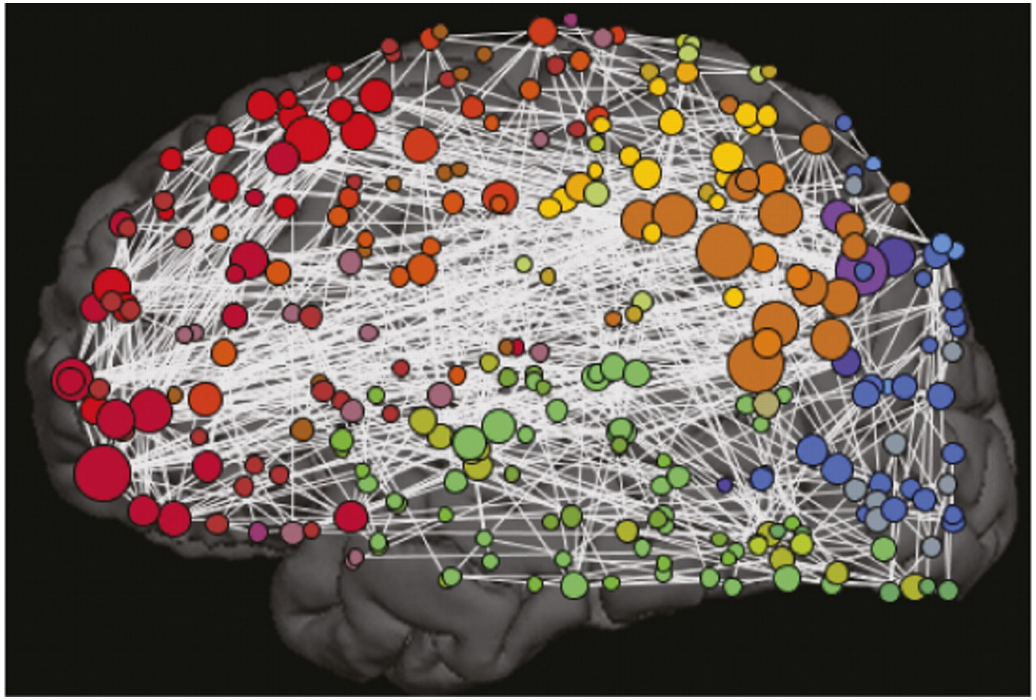
$$\mathcal{C}(v_i,v_j) = \begin{cases} \max_k \{|Q_k|\} & \forall k\ v_i, v_j \in Q_k \\ 0 & \nexists k\ v_i, v_j \in Q_k \end{cases}$$

$$\mathcal{D}(x) = \frac{1}{Z \sum_{j=1}^{N^-} n_j} \sum_{L_j=-1} \sum_{i=1}^{n_j} d(\Delta(x, x_{ji}))$$

- ***Paper is Under Review***

# (2) MIL via Hash Learning

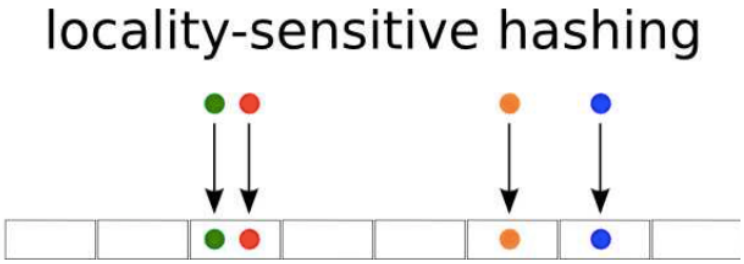
- Motivation: Scalability
- Issue: large number of instance in bags



- **Bag**: A whole brain v.s. **Instance**: A smaller region

# (2) MIL via Hash Learning

- Key Idea of Solution
- Locality-Sensitive Hashing to transfer feature space



# (2) MIL via Hash Learning

- Key Idea of Solution
- Framework: New representation of bags


- ***Xu, D., Wu, J., Li, D., Tian, Y., Zhu, X., & Wu, X. (2017). SALE: Self-adaptive LSH encoding for multi-instance learning. Pattern Recognition, 71, 460-482.***

# Thank you