Event Outlier Detection in Continuous Time

Siqi Liu[1,2] Milos Hauskrecht[1]

[1]University of Pittsburgh

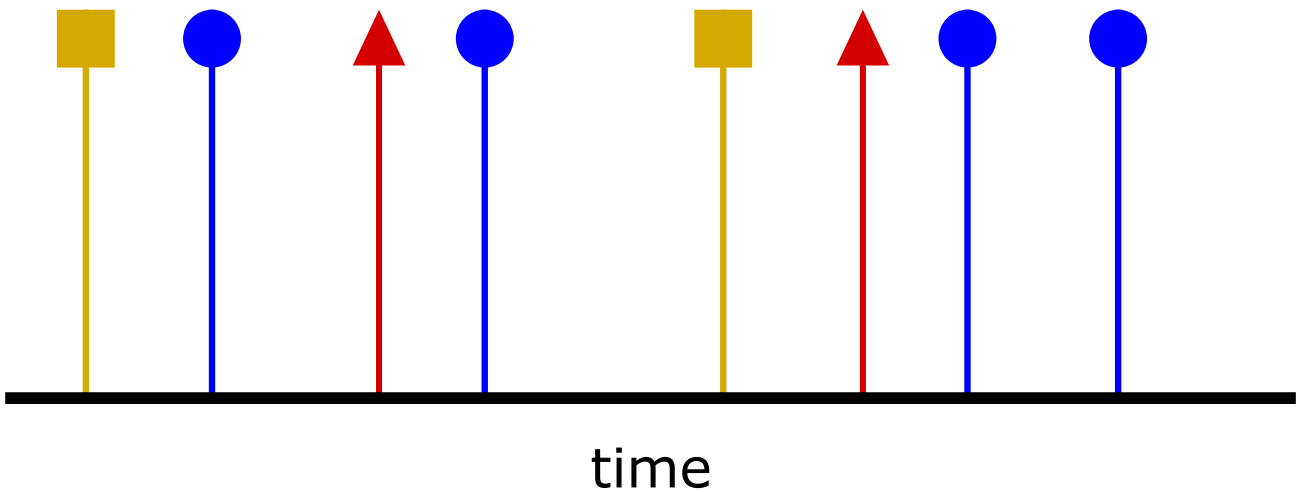
[2]Borealis AI

Event Sequences

- **Event sequences** represent discrete events happening in continuous time, e.g.
  - User actions when using a website
  - Earthquakes in a region
  - Financial transactions in a stock market
- Denoted as

$$\{(t_i, u_i) : t_i \in \mathbb{R}, u_i \in \mathbb{Z}\}_{i=1}^{N}$$

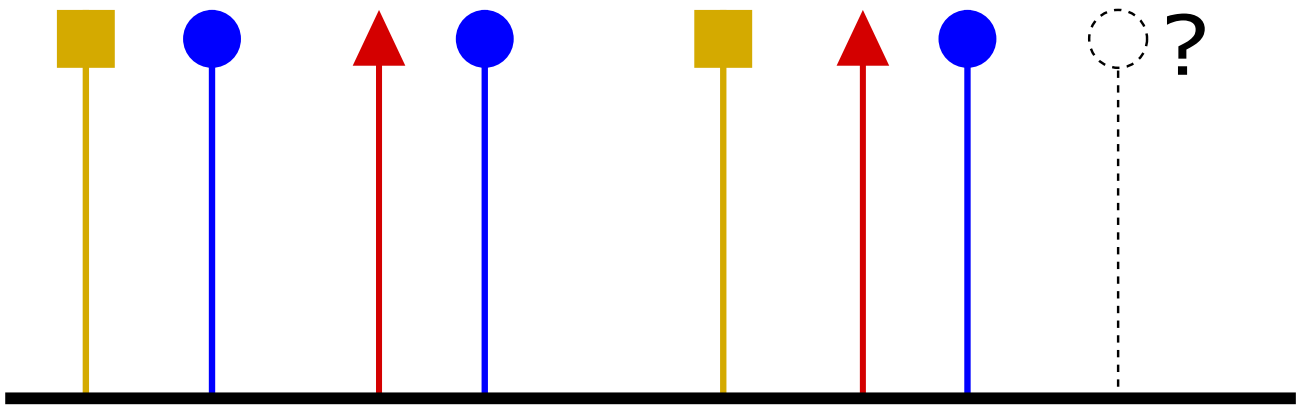
where t_i and u_i is the time and type of the i-th event



Prediction vs. Outlier Detection

Most previous works focus on solving the prediction task

Prediction: Given the history, what is the time and type of the next event?

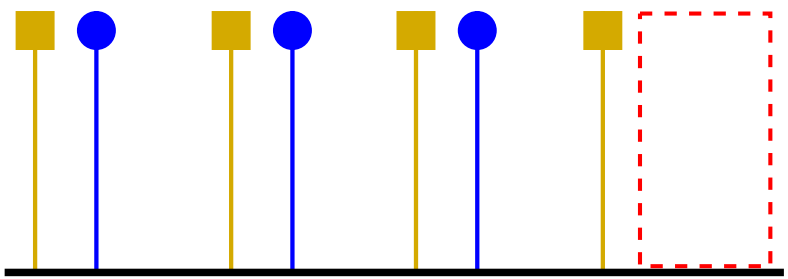


Prediction vs. Outlier Detection

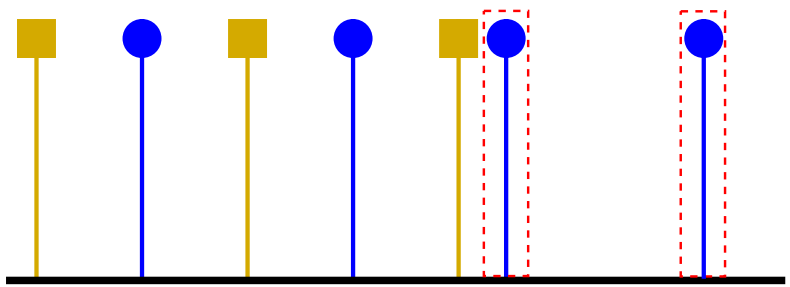
We formulate and solve the outlier (or anomaly) detection task

Outlier detection: Given the history, is the recent absence or occurrence of events abnormal?

Abnormal absence



Abnormal occurrence



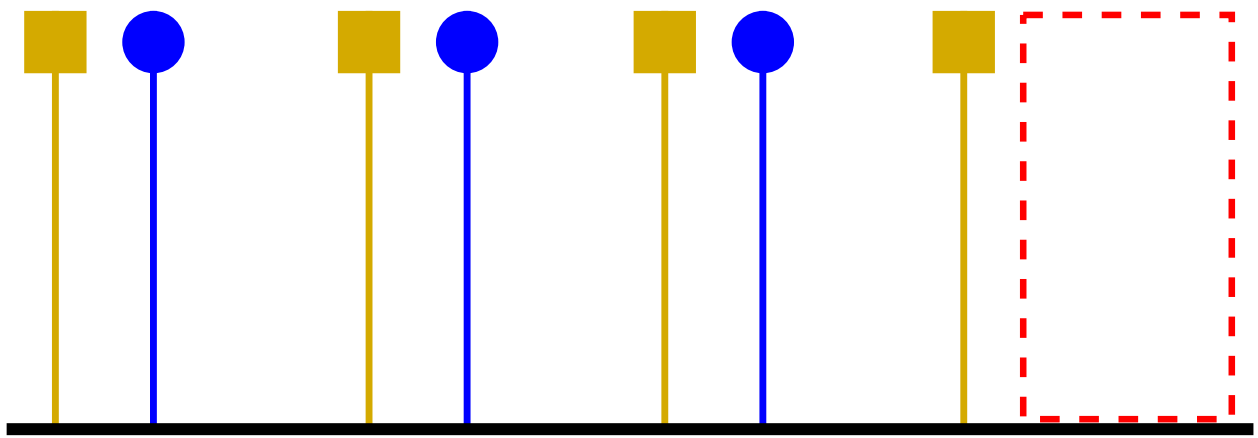
Abnormal Absence of Events

Hospitalized patients need to be given different medications at different times to treat their conditions

- Event (t, u): medication u was given at time t

Example 1

A patient needed to take a specific medication an hour ago but still has not received it



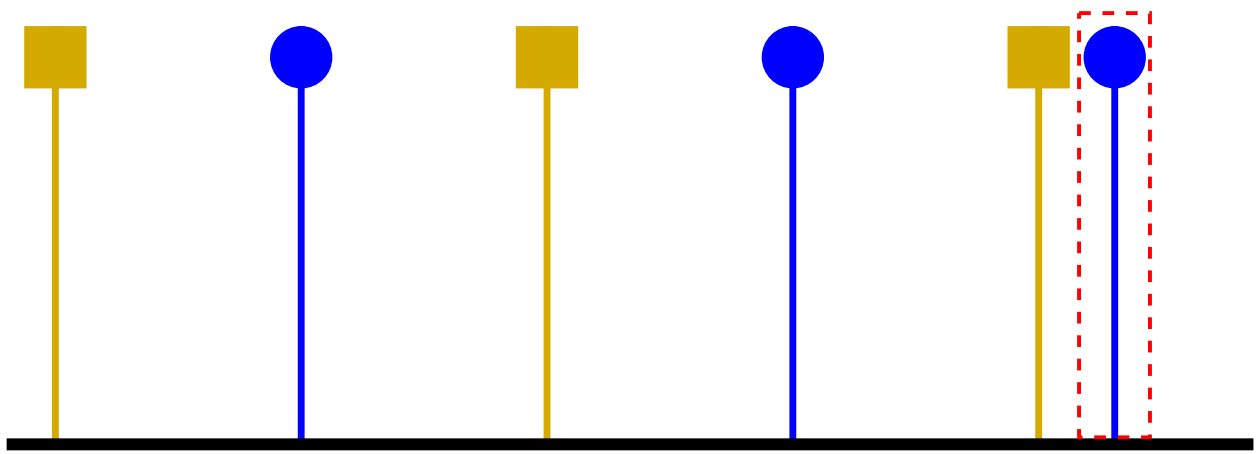
Abnormal Occurrence of Events

Hospitalized patients need to be given different medications at different times to treat their conditions

- Event (t, u): medication u was given at time t

Example 2

A patient was given a medication **too soon** from the previous dose



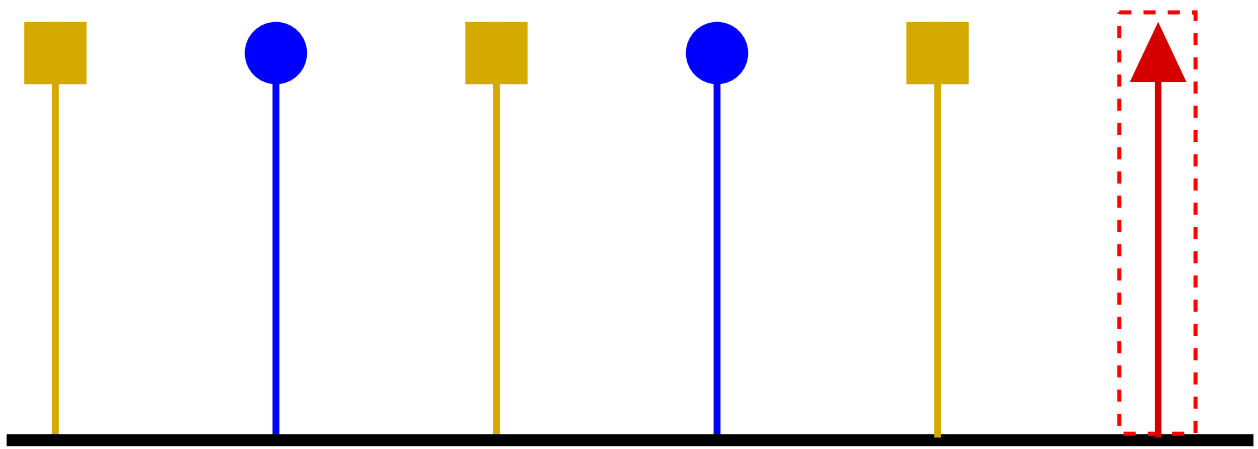
Abnormal Occurrence of Events

Hospitalized patients need to be given different medications at different times to treat their conditions

- Event (t, u): medication u was given at time t

Example 3

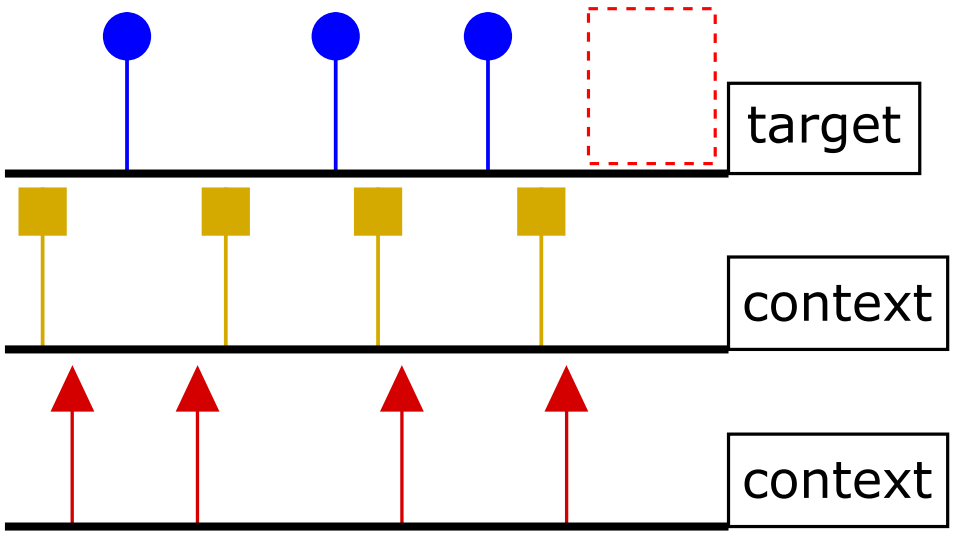
A patient was given a medication that is **unrelated** to their condition



Problem Formulation

Contextual outlier detection [1, 2]

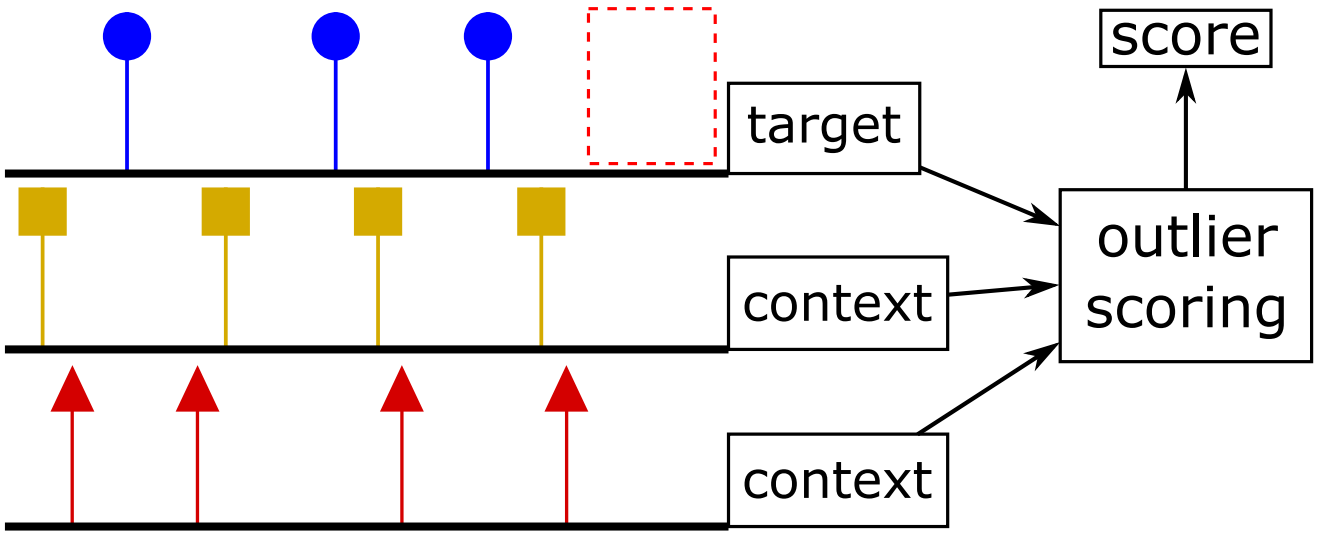
- Whether there is an outlier in a specific (target) type of events can depend on other (context) types of events



Problem Formulation

Outlier scoring

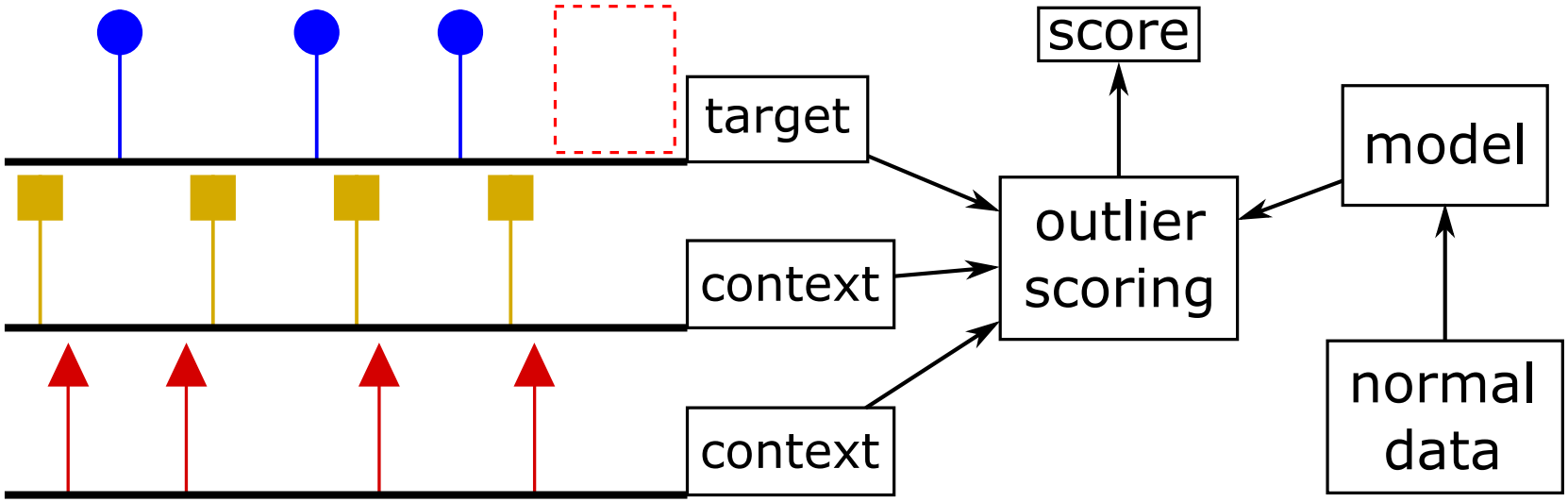
- A score is assigned to an event or blank interval to indicate how likely it is to be an outlier



Problem Formulation

Semi-supervised outlier detection [3]

- A model trained on normal data is available



Outlier Scoring Methods

- We propose outlier scoring methods based on the *conditional intensity function* $\lambda_0(t)$ of the underlying point-process model
  - Score for occurrence of an event at t: $-\lambda_0(t)$
  - Score for absence of events in a blank interval B: $\int_B \lambda_0(s)ds$
- Our methods can be combined with any point-process model
  - In this work, we use a model adapted from the continuous-time LSTM [4]
- Our methods have theoretical justifications and guarantees

Experiments

- We conduct experiments on synthetic data and real-world clinical data
- We simulate outliers using different types of outlier generating processes
- The results show the effectiveness of the proposed methods

Thank you

References I


Milos Hauskrecht, Michal Valko, Branislav Kveton, Shyam Visweswaram, and Gregory Cooper.
Evidence-based anomaly detection.
In *Annual American Medical Informatics Association Symposium*, pages 319–324, November 2007.


Xiuyao Song, Mingxi Wu, Christopher Jermaine, and Sanjay Ranka.
Conditional anomaly detection.
Knowledge and Data Engineering, IEEE Transactions on, 19(5):631–645, 2007.


Varun Chandola, Arindam Banerjee, and Vipin Kumar.
Anomaly detection: A survey.
ACM Comput. Surv., 41(3):15:1–15:58, July 2009.

References II


Hongyuan Mei and Jason M. Eisner.
The neural Hawkes process: A neurally self-modulating multivariate point process.
In *Advances in Neural Information Processing Systems*, pages 6757–6767, 2017.