Human-Centered ML
Part II

Human-centered Machine Learning
http://courses.mpi-sws.org/hcml-ws18/

MAX PLANCK INSTITUTE
FOR SOFTWARE SYSTEMS
Lectures for Part II

Five lectures on fundamentals

→ Marked temporal Point Processes
→ Optimal control
  Reinforcement learning

Five lectures on applications

Information propagation
Viral marketing
Opinion mining
Information integrity
Human learning
Evaluation for Part II

Paper reviewing assignments:
- Only for lectures on applications
- Due just before the lecture

Two coding assignments:
- Information propagation
  From Dec 20 to Jan 17
- Viral marketing
  From Jan 24 to Feb 5

Final exam:  On Feb 7 (review on Feb 5)
Introduction to Temporal Point Processes (I)
Many discrete events in continuous time

Online actions

Financial trading

Disease dynamics

Mobility dynamics
Variety of processes behind these events

Events are (noisy) observations of a variety of complex dynamic processes...

- Stock trading
- Flu spreading
- Article creation in Wikipedia
- News spread in Twitter
- Reviews and sales in Amazon
- Ride-sharing requests
- A user’s reputation in Quora

...in a wide range of temporal scales.
Example I: Information propagation

They can have an impact in the off-line world
Example II: Knowledge creation
Example III: Human learning

1st year computer science student

- Introduction to programming
- Discrete math
- Project presentation

For/do-while loops
Set theory
If ... else
Private functions
Class destructor
Graph Theory
Powerpoint vs. Keynote
Export pptx to pdf
Logic
Plot library
Class
templates
PP
Geometry
Powerpoint
Graph
theory
Inheritance
Set
theory
Private functions
Class destructor
Logic
Plot library
Export pptx to pdf
Powerpoint vs. Keynote
Graph Theory
Inheritance
Detailed event traces
The availability of event traces boosts a new generation of data-driven models and algorithms.
Aren’t these event traces just time series?

Discrete and continuous times series

Discrete events in continuous time

What about aggregating events in epochs?

How long is each epoch?
How to aggregate events per epoch?
What if no event in one epoch?
What about time-related queries?
Aren’t these event traces just time series?

Discrete and continuous times series

Discrete events in continuous time

What about aggregating events in epochs?

How long is each epoch?

How to aggregate events per epoch?

What if no event in one epoch?

What about time-related queries?

The framework of temporal point processes provides a **native representation**
Temporal Point Processes: Intensity function
Temporal point process:
A random process whose realization consists of discrete events localized in time

Formally:
\[ N(t) = \int_0^t dN(s) \]
\[ dN(t) = \sum_{t_i \in \mathcal{H}(t)} \delta(t - t_i) \, dt \]
Model time as a random variable

History, $\mathcal{H}(t)$

Prob. between $[t, t+dt)$

$\int_{t}^{t+dt} f^*(t) \, dt$

Density

$f^*(t) := f(t | \mathcal{H}(t))$

Prob. not before $t$

$S^*(t)$

Likelihood of a timeline:

$f^*(t_1) f^*(t_2) f^*(t_3) f^*(t) S^*(T)$
Problems of density parametrization (I)

It is difficult for model design and interpretability:

1. Densities need to integrate to 1 (i.e., partition function)
2. Difficult to combine timelines
Problems of density parametrization (II)

Difficult to combine timelines:

\[ f_1^*(t) + f_2^*(t) \neq f^*(t) \]

Sum of random processes:

\[ f^*(t) \neq f_1^*(t) \star f_2^*(t) \]

\[ f^*(t) \neq f_1^*(t) + f_2^*(t) \]
Intensity function

**Intensity:**
Probability between \([t, t+dt)\) but not before \(t\)

\[
\lambda^*(t)dt = \frac{f^*(t)dt}{S^*(t)} \geq 0 \implies \lambda^*(t)dt = \mathbb{E}[dN(t)|\mathcal{H}(t)]
\]

**Observation:** \(\lambda^*(t)\) It is a rate = # of events / unit of time
Advantages of intensity parametrization (I)

Suitable for model design and interpretable:

1. Intensities only need to be nonnegative
2. Easy to combine timelines
Advantages of intensity parametrization (II)

Easy to combine timeline:

\[ \lambda^*_1(t) + \lambda^*_2(t) = \lambda^*_1(t) \star \lambda^*_2(t) \]

Sum of random processes
Relation between $f^*$, $F^*$, $S^*$, $\lambda^*$

Central quantity we will use!