

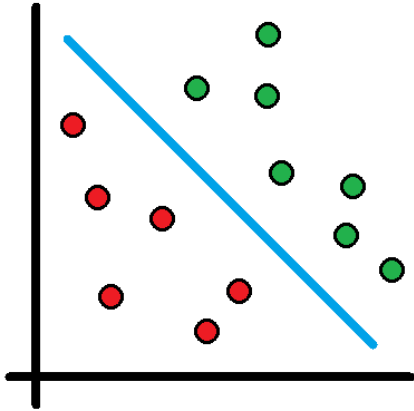
# Machine Learning and Optimization

## Adapting to Contemporary Applications

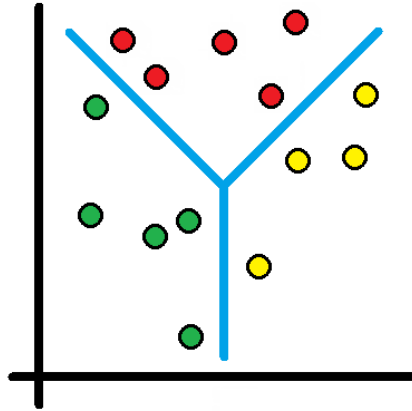
**Purushottam Kar**

Indian Institute of Technology Kanpur

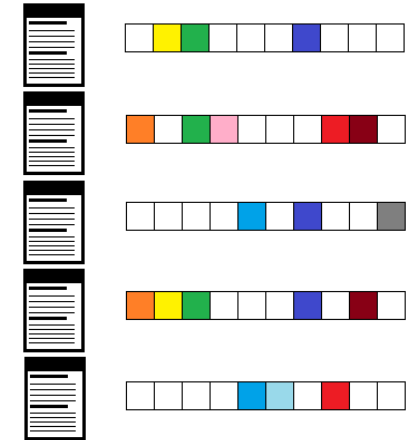
# Traditional Machine Learning Primitives



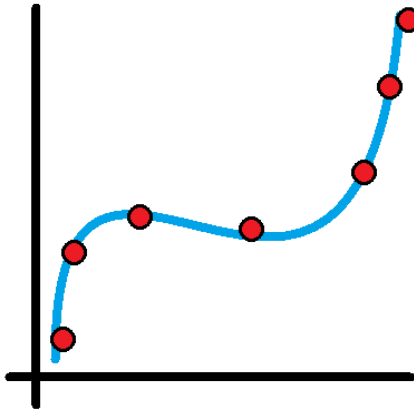
Binary Classification



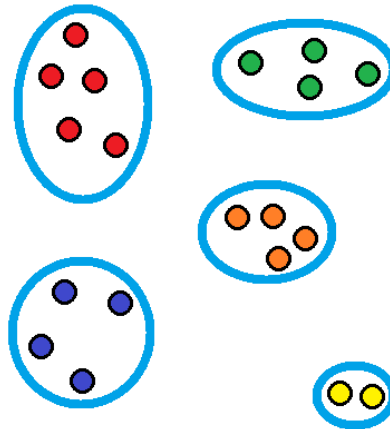
Multi Classification



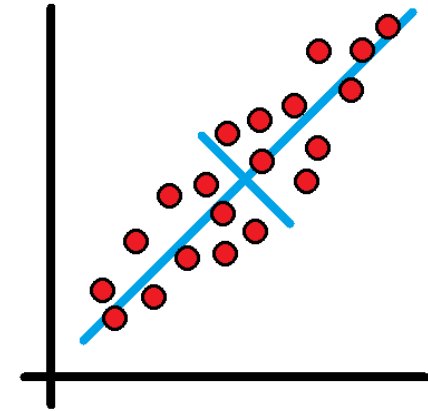
Multi-label Classfn



Regression

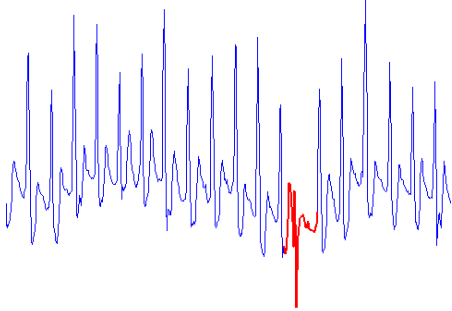


Clustering



Component Analysis

# Contemporary Applications



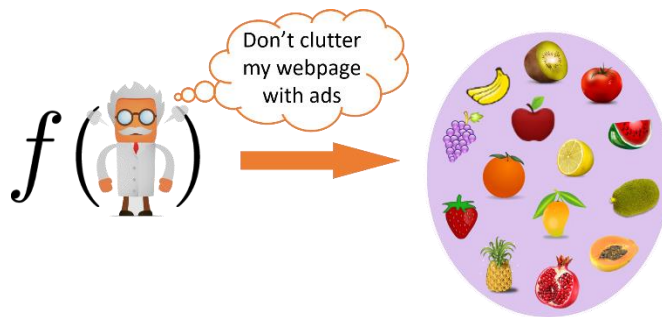
Anomaly Detection



Drug Discovery



Data Analytics



Recommender Systems



Medical Diagnostics

# Machine Learning in Action

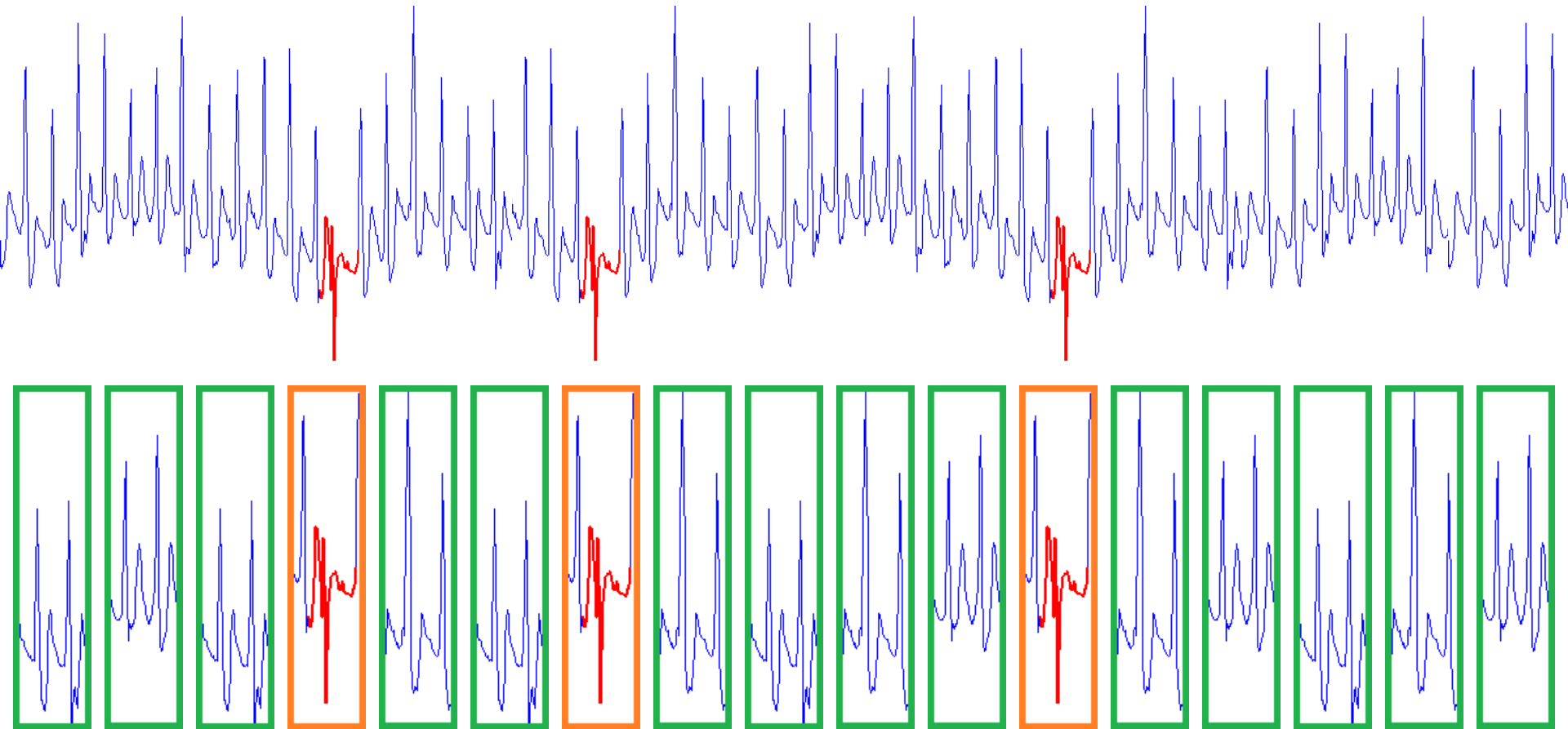
- **Predictive** modeling
  - Use system descriptions to predict anomalous behavior
  - Detecting non-authorized flow of control/information
  - Use access patterns to assess threat levels
- **Analytic** modeling
  - Building models of anomalies, attacks, failures
  - Identifying points of failure in a system
  - Differentiating stochasticity from anomaly

## So business as usual? Nope!

- Changes in the **nature of data**
  - Volume of data – large to huge
  - Data access – online or streaming
  - Data distributions – heavy tailed, skewed
  - Noise levels – high, malicious corruptions
- Changes in **application requirements**
  - Extreme precision
  - Cost sensitivity – allowing fraudulent access vs false alarm
  - Scalability, ease of use and modification

# Machine Learning for Critical Applications

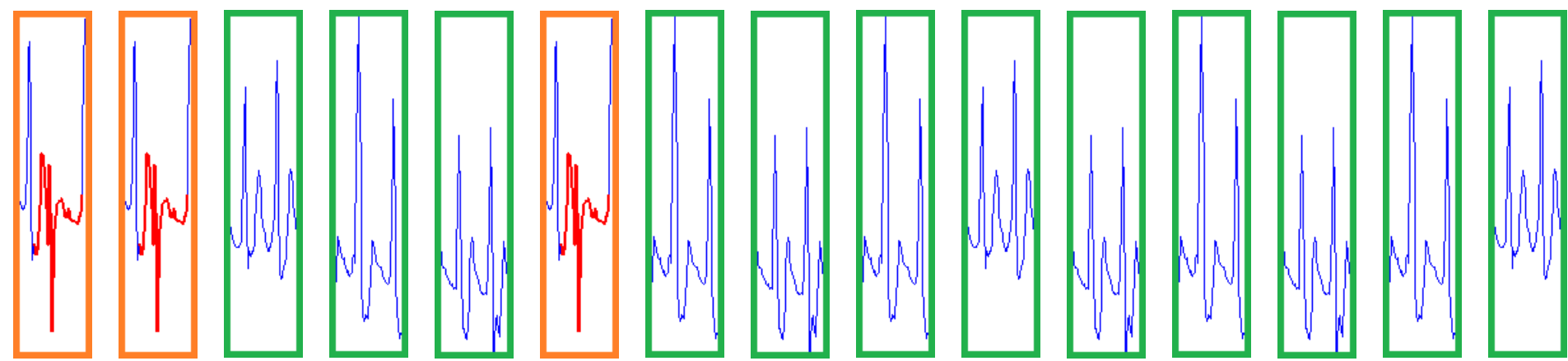
## Learning on Streams



**Online learning, Stochastic Optimization**

# Machine Learning for Critical Applications

## Learning with imbalanced, heavy tailed data



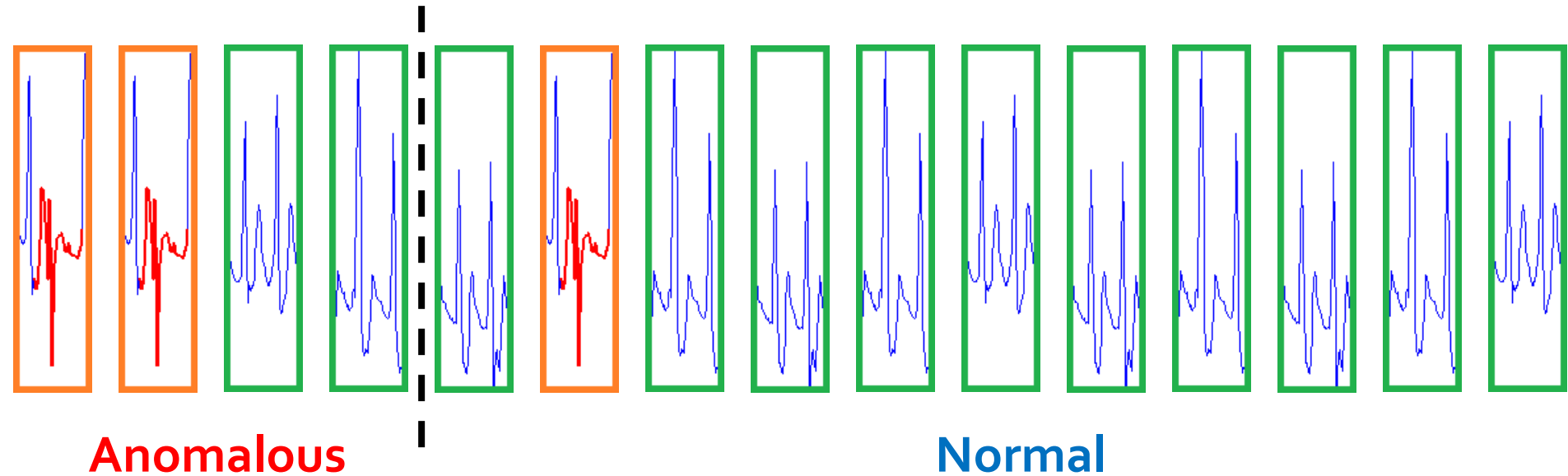
Promote objects of interest at the top of the ranked list

**Precision at the Top, Area under the ROC Curve,  
Precision-Recall Break-even Point, partial AUC,  
concentrated AUC**

**Learning to Rank**

# Machine Learning for Critical Applications

Learning with imbalanced, heavy tailed data



Take into account risk models, asymmetries, data skew

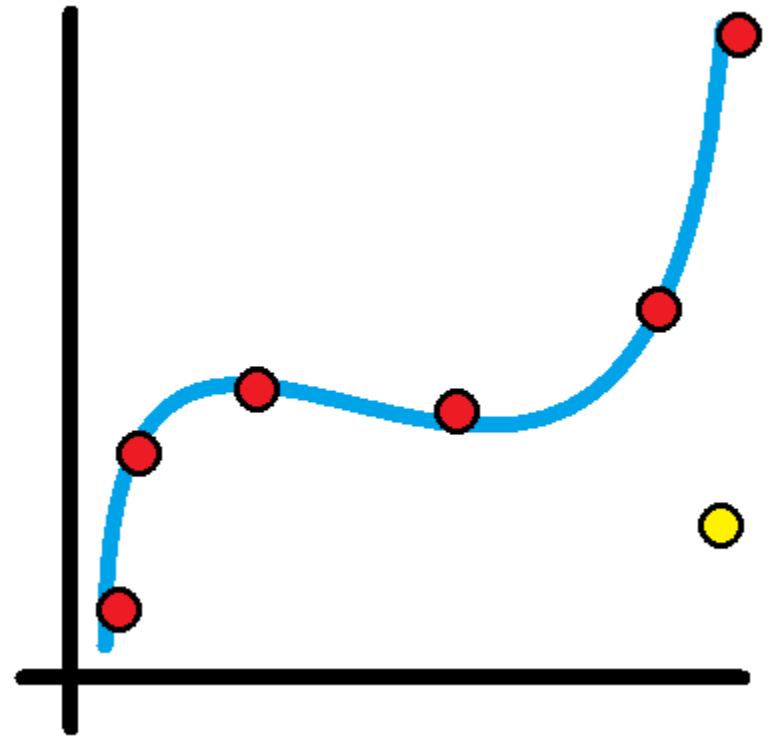
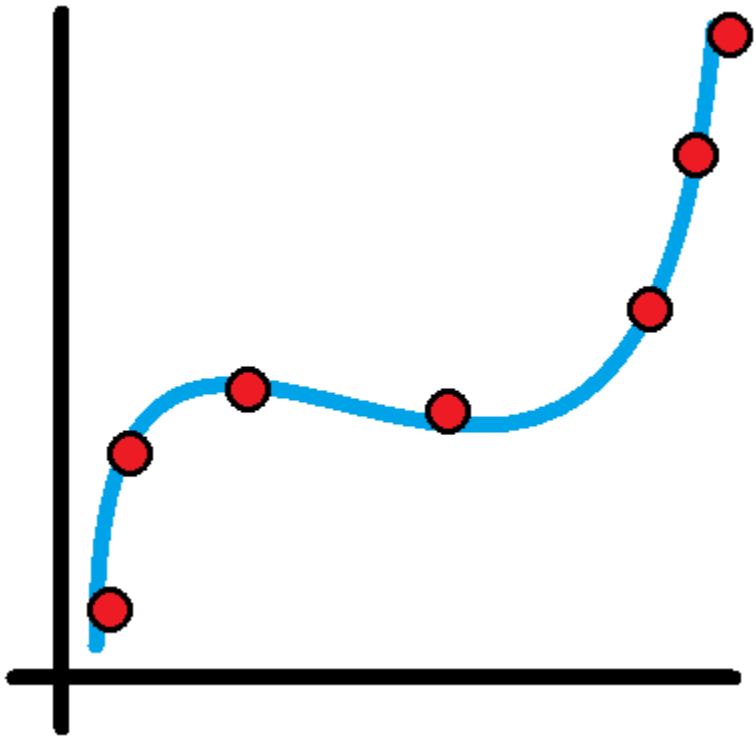
$$\Psi (\text{TPR}, \text{TNR})$$

**Multivariate Optimization**



# Machine Learning for Critical Applications

## Learning with (adversarially) Corrupted Data



**Robust Classification, Regression, Ranking**

# Machine Learning for Critical Applications

## Recommendation, ranking



- [Main page](#)
- [Contents](#)
- [Featured content](#)
- [Current events](#)
- [Random article](#)
- [Donate to Wikipedia](#)
- [Wikipedia store](#)

### Interaction

- [Help](#)
- [About Wikipedia](#)
- [Community portal](#)
- [Recent changes](#)
- [Contact page](#)

### Tools

- [What links here](#)

[Create account](#) [Log in](#)

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)

## Bharata Natyam

From Wikipedia, the free encyclopedia

**Bharathanatyam** (**Tamil:** பரதநாட்டியம்) is a form of **Indian classical dance** that originated in the temples of **Tamil Nadu**.<sup>[1][2][3][4][5]</sup> It was described in the treatise *Natya Shastra* by **Bharata** around the beginning of the common era. Bharata Natyam is known for its grace, purity, tenderness, expression and sculpturesque poses. **Lord Shiva** is considered the God of this dance form. Today, it is one of the most popular and widely performed dance styles and is practiced by male and female dancers all over the world, although it is more commonly danced by women.<sup>[6]</sup>

### Contents [\[hide\]](#)

- [1 Etymology](#)
- [2 Dance tradition](#)
- [3 Essential ideas](#)
  - [3.1 Spiritual symbolism](#)
- [4 Medieval decline](#)
- [5 Modern rebirth](#)

**Bharathanatyam**



American Football, Architecture, Artificial Intelligence, Aviation, Billionaires, Bodies of water, Broadcasting, Card Games, Cardiology, Celebrity, Censorship in the arts, Comics, Continents, Countries, Crafts, Crime, Critical theory, Cultural anthropology, **Dances by name**, Deserts, Design, Drawing, Earth, Epidemiology, Family, Film, Folklore, Food and drink, Food culture, Fishing, Geometry, Humor, **Indian culture**, Internet, Lacrosse, Lakes, Landforms, Languages, Literature, Magazines, Mountains, Museums, Mythology, Navigation, Newspapers, Nootropics, Oceanography, Opera, Oral hygiene, **Performing arts in India**, Philosophers, Plumbing, Pilat  
Suffixes,

**Extreme Multi-label classification**

, Subterranea, s, World

# Practical Applications

- **Reality**: long stream of corrupted, imbalanced data
- **Critical**
  - proper modeling of data, feature
  - appropriate choice of performance measure
- **Desirable**
  - balance between scalability and accuracy
  - modularity, extendibility

## Works in Progress

- Online optimization for ranking tasks
  - **Learn to rank** objects in a stream [NIPS 2014 ICML 2015]
- Online optimization for learning with imbalanced data
  - **Learn to identify needles** in a stream of hay [ICML 2015]
- Scalable robust optimization
  - **Learn to regress** in the presence of an adversary [NIPS 2015]
- Scalable optimization for extreme classification
  - **Multi-label classification** with a million labels [NIPS 2015]

**Thank you**