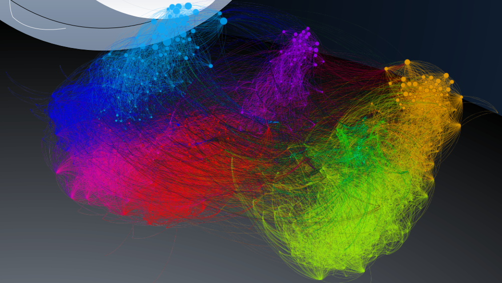
Graphs in Machine Learning

Michal Valko

Inria & ENS Paris-Saclay, MVA



Two (main) sources of graphs in ML

Natural graphs as models for networks

Constructed graphs as nonparametric basis

Two (main) sources of graphs in ML

Natural graphs as models for networks

- given as an input

Constructed graphs as nonparametric basis

Two (main) sources of graphs in ML

Natural graphs as models for networks

- given as an input
- discover interesting properties of the structure

Constructed graphs as nonparametric basis

Two (main) sources of graphs in ML

Natural graphs as models for networks

- given as an input
- discover interesting properties of the structure

Constructed graphs as nonparametric basis

- we create (learn) the similarity structure from flat data

Two (main) sources of graphs in ML

Natural graphs as models for networks

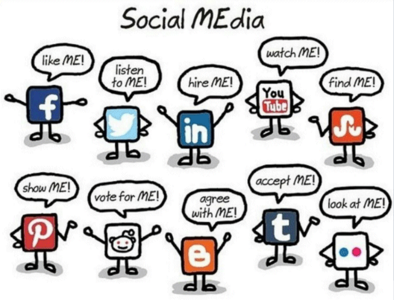
- given as an input
- discover interesting properties of the structure

Constructed graphs as nonparametric basis

- we create (learn) the similarity structure from flat data
- it's a tool (e.g., nonparametric regularizer) to encode structural properties (e.g., independence, ...)

Natural graphs from social networks

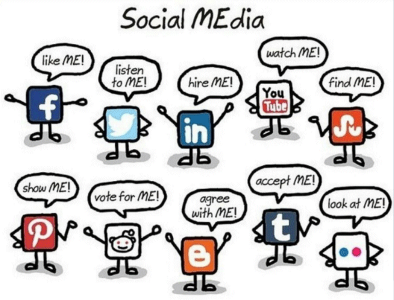
- people and their interactions



Natural graphs from social networks

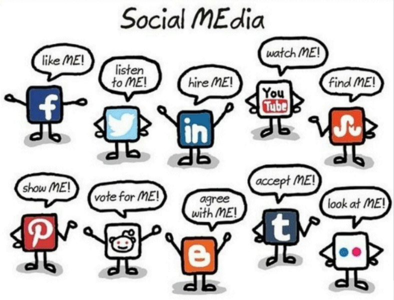
- people and their interactions

- structure is rather a *phenomena*



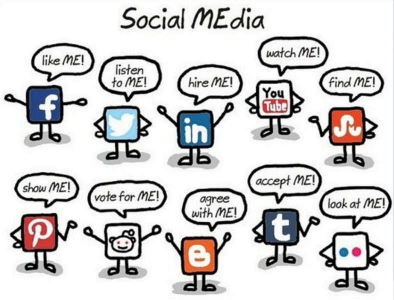
Natural graphs from social networks

- people and their interactions
- structure is rather a *phenomena*
- typical ML tasks



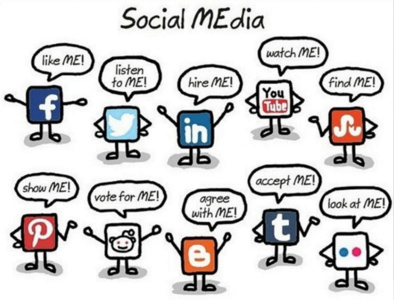
Natural graphs from social networks

- people and their interactions
- structure is rather a *phenomena*
- typical ML tasks
  - advertising



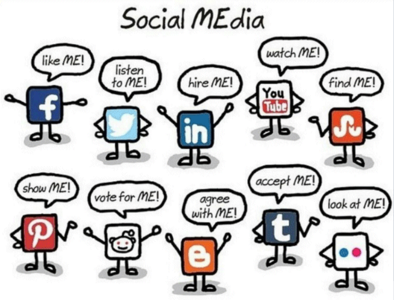
Natural graphs from social networks

- people and their interactions
- structure is rather a *phenomena*
- typical ML tasks
  - advertising
  - link prediction (PYMK)



Natural graphs from social networks

- people and their interactions
- structure is rather a *phenomena*
- typical ML tasks
  - advertising
  - link prediction (PYMK)
  - find influential sources



Natural graphs from utility and technology networks

- power grids, roads, Internet, sensor networks

Berkeley's Floating Sensor Network

Natural graphs from utility and technology networks

- power grids, roads, Internet, sensor networks
- structure is either *hand designed* or not

Berkeley's Floating Sensor Network

Natural graphs from utility and technology networks

- power grids, roads, Internet, sensor networks
- structure is either *hand designed* or not
- typical ML tasks

Berkeley's Floating Sensor Network

Natural graphs from utility and technology networks

- power grids, roads, Internet, sensor networks
- structure is either *hand designed* or not
- typical ML tasks
  - best routing under unknown or variable costs

Berkeley's Floating Sensor Network

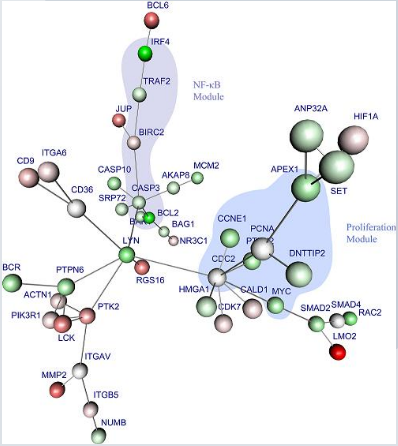
Natural graphs from utility and technology networks

- power grids, roads, Internet, sensor networks
- structure is either *hand designed* or not
- typical ML tasks
  - best routing under unknown or variable costs
  - identify the node of interest

Berkeley's Floating Sensor Network

Natural graphs from biological networks

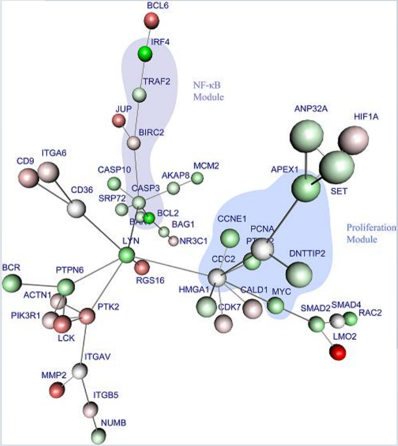
- protein-protein interactions



Diffuse large B-cell lymphomas - Dittrich et al. (2008)

Natural graphs from biological networks

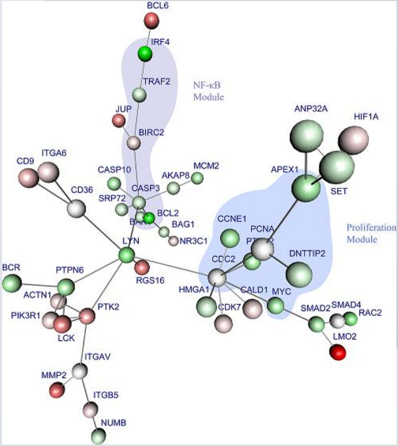
- protein-protein interactions
- gene regulatory networks



Diffuse large B-cell lymphomas - Dittrich et al. (2008)

Natural graphs from biological networks

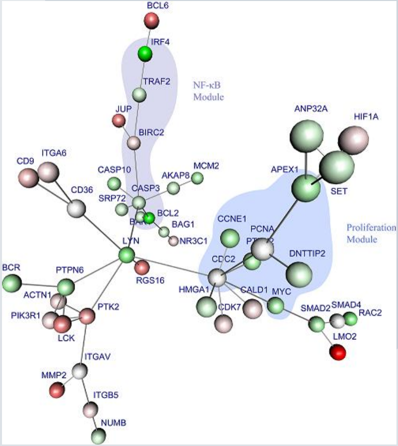
- protein-protein interactions
- gene regulatory networks
- typical ML tasks



Diffuse large B-cell lymphomas - Dittrich et al. (2008)

Natural graphs from biological networks

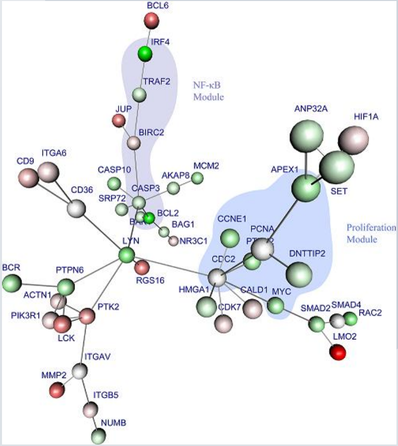
- protein-protein interactions
- gene regulatory networks
- typical ML tasks
  - discover unexplored interactions



Diffuse large B-cell lymphomas - Dittrich et al. (2008)

Natural graphs from biological networks

- protein-protein interactions
- gene regulatory networks
- typical ML tasks
  - discover unexplored interactions
  - learn or reconstruct the structure



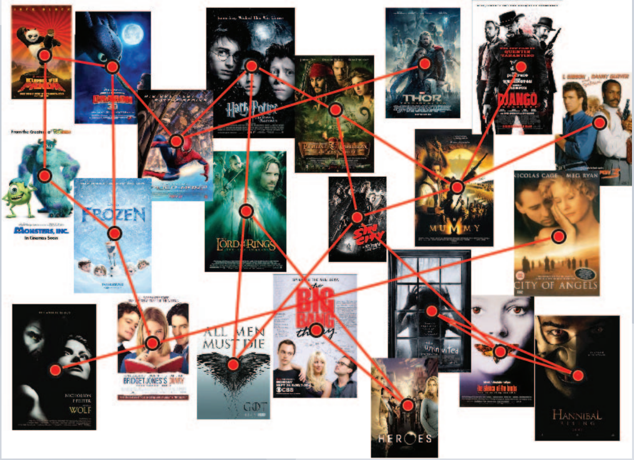
Diffuse large B-cell lymphomas - Dittrich et al. (2008)

Constructed graphs from similarity networks

graph is not naturally given

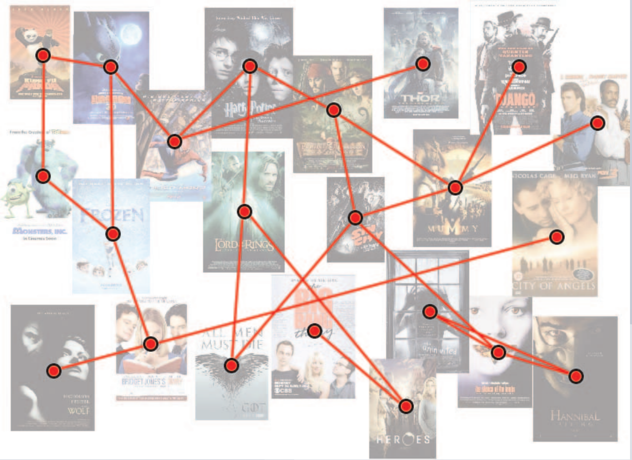
Constructed graphs from similarity networks

but we can construct it



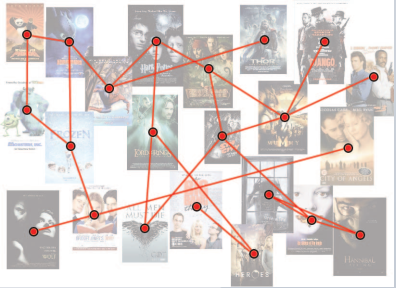
Constructed graphs from similarity networks

and use it as an abstraction



Constructed graphs from similarity networks

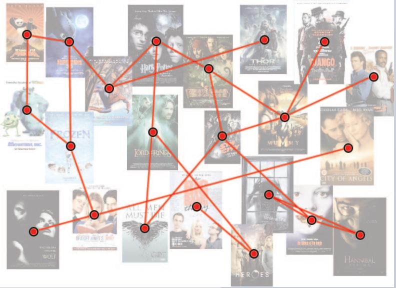
- vision
- audio
- text



Movie similarity

Constructed graphs from similarity networks

- vision
- audio
- text
- typical ML tasks
  - semi-supervised learning
  - spectral clustering
  - manifold learning



Movie similarity

What will you learn in the Graphs in ML course?

Concepts, tools, and methods to work with graphs in ML.

What will you learn in the Graphs in ML course?

Concepts, tools, and methods to work with graphs in ML.

Specific applications of graphs in ML.

What will you learn in the Graphs in ML course?

Concepts, tools, and methods to work with graphs in ML.

Specific applications of graphs in ML.

Theoretical toolbox to analyze graph-based algorithms.

What will you learn in the Graphs in ML course?

Concepts, tools, and methods to work with graphs in ML.

Specific applications of graphs in ML.

Theoretical toolbox to analyze graph-based algorithms.

How to tackle: *large graphs, online setting, graph construction* ...

What will you learn in the Graphs in ML course?

Concepts, tools, and methods to work with graphs in ML.

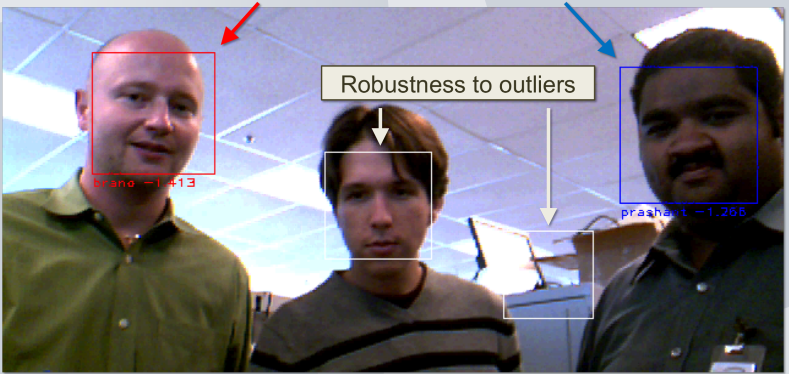
Specific applications of graphs in ML.

Theoretical toolbox to analyze graph-based algorithms.

How to tackle: *large graphs, online setting, graph construction* ...

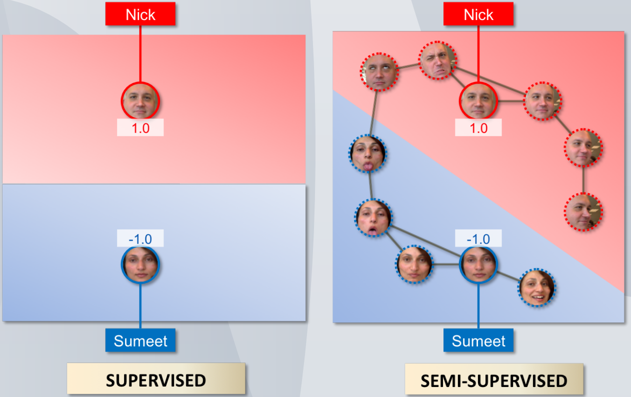
One example: Online Semi-Supervised Face Recognition

Online Semi-Supervised Face Recognition



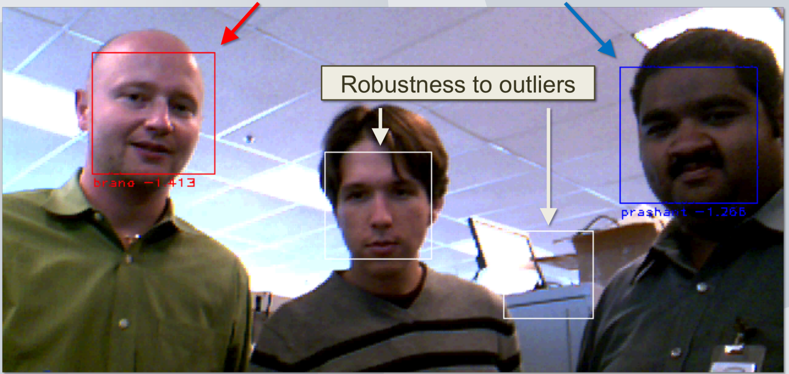
Online Semi-Supervised Face Recognition

graph-based semi-supervised learning



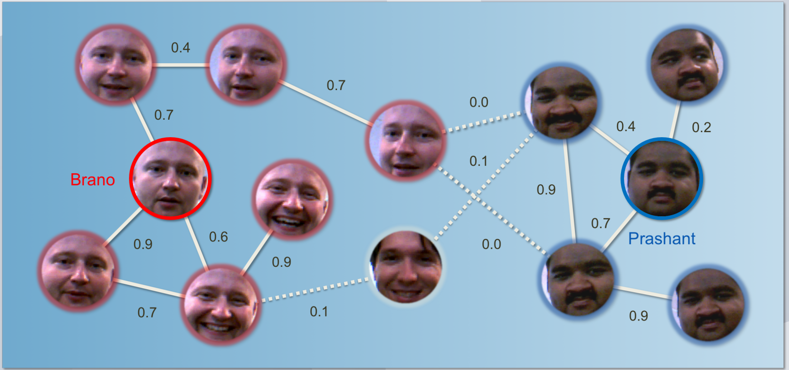
Online Semi-Supervised Face Recognition

graph is not given



Online Semi-Supervised Face Recognition

we will construct it!

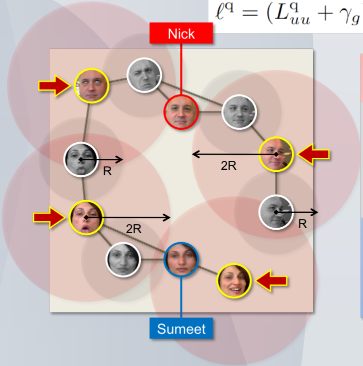


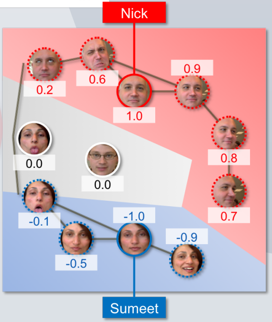
An example of a similarity graph over faces. The faces are vertices of the graph. The edges of the graph connect similar faces. Labeled faces are outlined by thick solid lines.

Online Semi-Supervised Face Recognition

online learning - graph sparsification

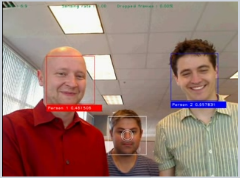
$$\ell^{\mathrm{q}} = (L^{\mathrm{q}}_{uu} + \gamma_g V)^{-1} W^{\mathrm{q}}_{ul} \ell_l$$

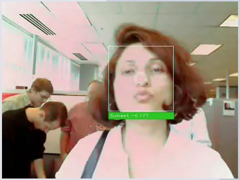




DEMO

second TD





see the demo: `https://misovalko.github.io/projects/courses/graphsML/lectures/videos/kveton2009nipsdemo.officespace.mov`

OSS FaceReco: Analysis

$$\frac{1}{n}\sum_t (\ell_t^{q}[t]-y_t)^2 \leq \frac{3}{n}\sum_t (\ell_t^{*}-y_t)^2 + \frac{3}{n}\sum_t (\ell_t^{o}[t]-\ell_t^{*})^2 + \frac{3}{n}\sum_t (\ell_t^{q}[t]-\ell_t^{o}[t])^2$$

Error of our solution	Offline learning error	Online learning error	Quantization error

Claim: When the regularization parameter is set as $\gamma_g = \Omega(n_l^{3/2})$, the difference between the risks on labeled and all vertices decreases at the rate of $O(n_l^{-1/2})$ (with a high probability)

$$\frac{1}{n}\sum_t (\ell_t^{*}-y_t)^2 \leq \frac{1}{n_l}\sum_{i\in l} (\ell_i^{*}-y_i)^2 + \beta + \sqrt{\frac{2\ln(2/\delta)}{n_l}}(n_l\beta+4)$$

$$\beta \leq \left[\frac{\sqrt{2}}{\gamma_g+1} + \sqrt{2n_l}\,\frac{1-\sqrt{c_u}}{\sqrt{c_u}}\,\frac{\lambda_M(L)+\gamma_g}{\gamma_g^2+1}\right]$$

OSS FaceReco: Analysis

$$\frac{1}{n}\sum_t (\ell_t^{\mathrm{q}}[t] - y_t)^2 \le \frac{3}{n}\sum_t (\ell_t^* - y_t)^2 + \frac{3}{n}\sum_t (\ell_t^{\mathrm{o}}[t] - \ell_t^*)^2 + \frac{3}{n}\sum_t (\ell_t^{\mathrm{q}}[t] - \ell_t^{\mathrm{o}}[t])^2$$

Error of our solution | Offline learning error | Online learning error | Quantization error

Claim: When the regularization parameter is set as $\gamma_g = \Omega(n^{1/4})$, the average error between the offline and online HFS predictions decreases at the rate of $O(n^{-1/2})$

$$\frac{1}{n}\sum_t (\ell_t^{\mathrm{o}}[t] - \ell_t^*)^2 \le \frac{1}{n}\sum_t \left\| \ell^{\mathrm{o}}[t] - \ell^* \right\|_2^2 \le \frac{4n_l}{(\gamma_g + 1)^2}$$

$$\|\ell\|_2 \le \frac{\|y\|_2}{\lambda_m(C^{-1}K + I)} = \frac{\|y\|_2}{\lambda_m(K)\lambda_M^{-1}(C) + 1} \le \frac{\sqrt{n_l}}{\gamma_g + 1}$$

OSS FaceReco: Analysis

$$\frac{1}{n}\sum_t(\ell_t^{\mathrm{q}}[t]-y_t)^2 \le \frac{3}{n}\sum_t(\ell_t^* - y_t)^2 + \frac{3}{n}\sum_t(\ell_t^{\mathrm{o}}[t]-\ell_t^*)^2 + \frac{3}{n}\sum_t(\ell_t^{\mathrm{q}}[t]-\ell_t^{\mathrm{o}}[t])^2$$

Error of our solution	Offline learning error	Online learning error	Quantization error

Claim: When the regularization parameter is set as $\gamma_g = \Omega(n^{1/8})$, and the Laplacians L^q and L^o and normalized, the average error between the online and online quantized HFS predictions decreases at the rate of $O(n^{-1/2})$

$$\frac{1}{n}\sum_t(\ell_t^{\mathrm{q}}[t]-\ell_t^{\mathrm{o}}[t])^2 \le \frac{1}{n}\sum_t\left\|\ell^{\mathrm{q}}[t]-\ell^{\mathrm{o}}[t]\right\|_2^2 \le \frac{n_l}{c_u^2\gamma_g^4}\left\|L^{\mathrm{q}}-L^{\mathrm{o}}\right\|_F^2$$

$$\left\|L^{\mathrm{q}}-L^{\mathrm{o}}\right\|_F^2 \propto O(k^{-2/d})$$

The distortion rate of online k-center clustering is $O(k^{-1/d})$, where d is dimension of the manifold and k is the number of representative vertices

MVA and Graphs: 2 courses

The two MVA graph courses offer complementary material.

Fall: **Graphs in ML**

this class

- focus on learning
- spectral clustering
- random walks
- graph Laplacian
- semi-supervised learning
- theoretical analyses
- online learning
- recommender systems
- **graph neural networks**

MVA and Graphs: 2 courses

The two MVA graph courses offer complementary material.

Fall: **Graphs in ML**

this class

- focus on learning
- spectral clustering
- random walks
- graph Laplacian
- semi-supervised learning
- theoretical analyses
- online learning
- recommender systems
- **graph neural networks**

Late Fall: **ALTeGraD**

by Michalis Vazirgiannis

- dimensionality reduction
- feature selection
- text mining
- graph mining
- community mining
- graph generators
- graph-evaluation measures
- privacy in graph mining
- big data

Administrivia

8 lectures + 3 recitations (TDs)

Validation: grades from TDs (40%) + class project (60%)

Prerequisites: linear algebra, basic statistics

Language: English

Course website:
`https://misovalko.github.io/mva-ml-graphs.html`

Michal Valko

michal.valko@inria.fr

Inria & ENS Paris-Saclay, MVA

https://misovalko.github.io/mva-ml-graphs.html