ARAMIS LAB
BRAIN DATA SCIENCE

UKRAINIAN CATHOLIC UNIVERSITY
FACULTY OF APPLIED SCIENCES

23 – 25 July 2018

Lviv Data Science Summer School 2018

Machine Learning for Medical Applications:

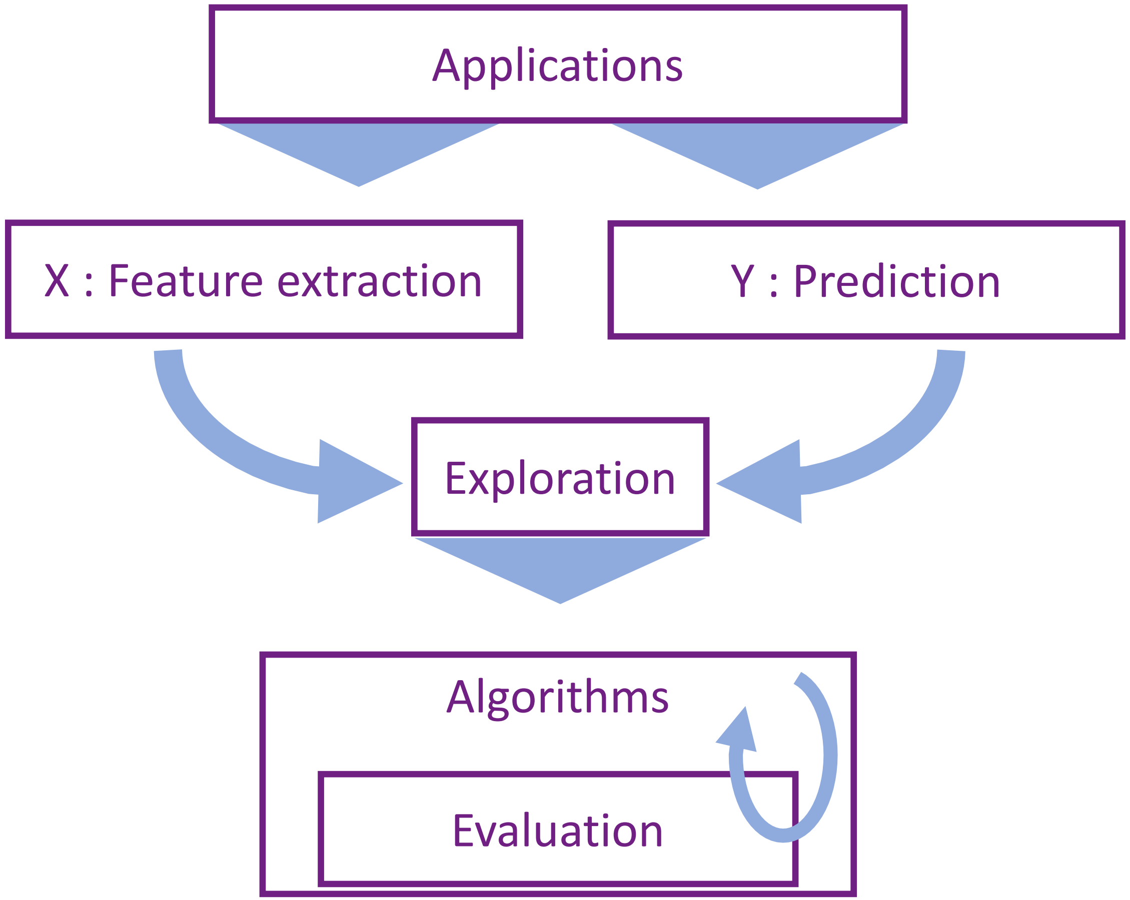
Igor Koval

PhD Student in Applied Mathematics

Brain and Spine Institute, Pitié Salpétrière Hospital, Paris, France
& Mathematical Laboratory of Ecole Polytechnique

igor.koval@icm-institute.org

Plan


Applications
X : Feature extraction
Y : Prediction
Exploration
Algorithms
Evaluation

Miscellaneous - non-exhaustive

ARAMIS LAB BRAIN DATA SCIENCE

FACULTY OF APPLIED SCIENCES

I. Current focus in research
   I. Disease progression & Longitudinal data
   II. Deep Learning
   III. Mixed-effects & Generative models
   IV. Not Euclidean data
   V. AramisLab-related model

II. Challenges in Medicine
   I. What question(s) to answer
   II. Labels, variability and legal aspects
   III. Companies

I. Current focus in research

I. Disease progression & Longitudinal data
II. Deep Learning
III. Mixed-effects & Generative models
IV. Non-Euclidean data
V. AramisLab-related model

Disease Progression & Longitudinal data

ARAMIS LAB
BRAIN DATA SCIENCE

UKRAINIAN CATHOLIC UNIVERSITY
FACULTY OF APPLIED SCIENCES

Disease Progression:
Progression of features, from a normal to a abnormal state

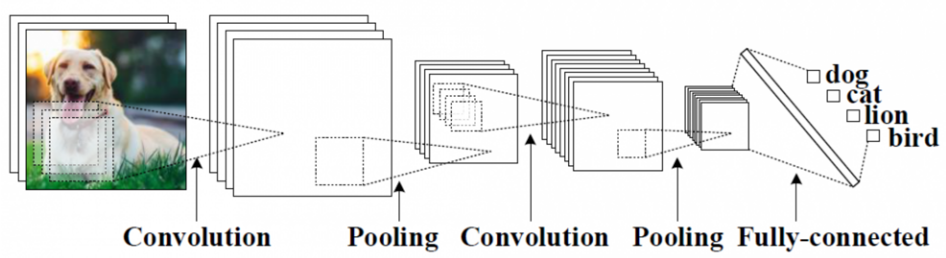
- Group-average level
- Individual level

Longitudinal data:
Repeated observations of individuals at different time-points

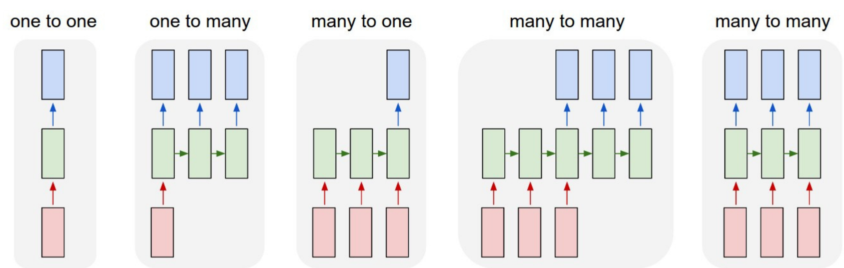
Each row is a patient	Each row is a visit
How to handle patients with different number of visits?	How to deal with multiple prediction per patient
Regress the longitudinal data and put ratio values : slopes, coefficients, …	Predict for each visit, and reprocess to average : not easy -> what kind of average, more weight to the last visit?

Deep Learning

Convolutional Neural Network - CNN



Reccurent Neural Network – RNN
Especially Long-Short Term Memory - LSTM



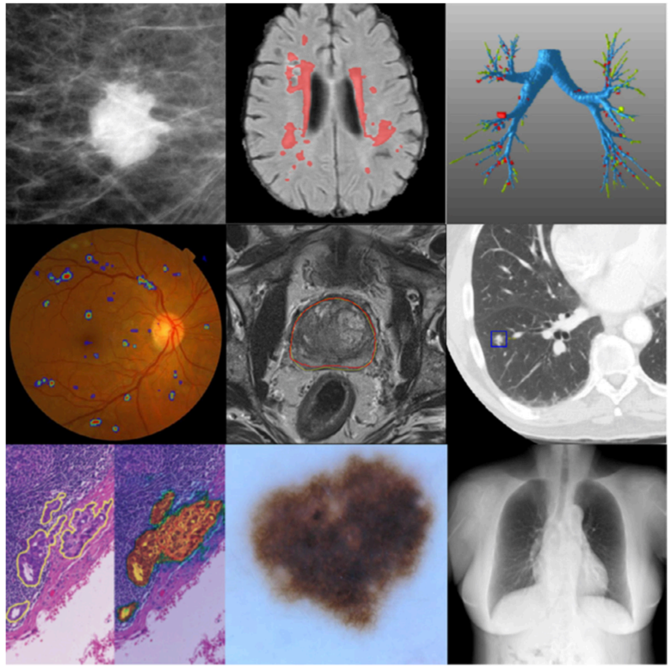


Fig. 3. Collage of some medical imaging applications in which deep learning has achieved state-of-the-art results. From top-left to bottom-right: mammographic mass classification (Kooi et al., 2016), segmentation of lesions in the brain (top ranking in BRATS, ISLES and MRBrains challenges, image from Ghafoorian et al. (2016b), leak detection in airway tree segmentation (Charbonnier et al., 2017), diabetic retinopathy classification (Kaggle Diabetic Retinopathy challenge 2015, image from van Grinsven et al. (2016), prostate segmentation (top rank in PROMISE12 challenge), nodule classification (top ranking in LUNA16 challenge), breast cancer metastases detection in lymph nodes (top ranking and human expert performance in CAMELYON16), human expert performance in skin lesion classification (Esteva et al., 2017), and state-of-the-art bone suppression in x-rays, image from Yang et al. (2016c).

A survey on deep learning in medical image analysis
Litjens et al. 2017

One drawback of Deep Learning is its need of large datasets

- Data augmentation
- Transfer Learning

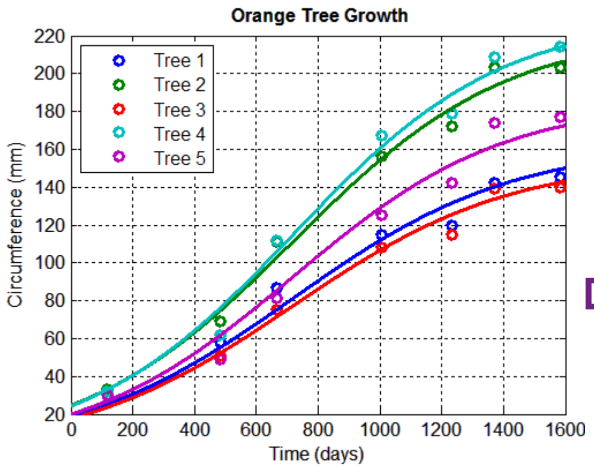
Mixed-effects & Generative Models

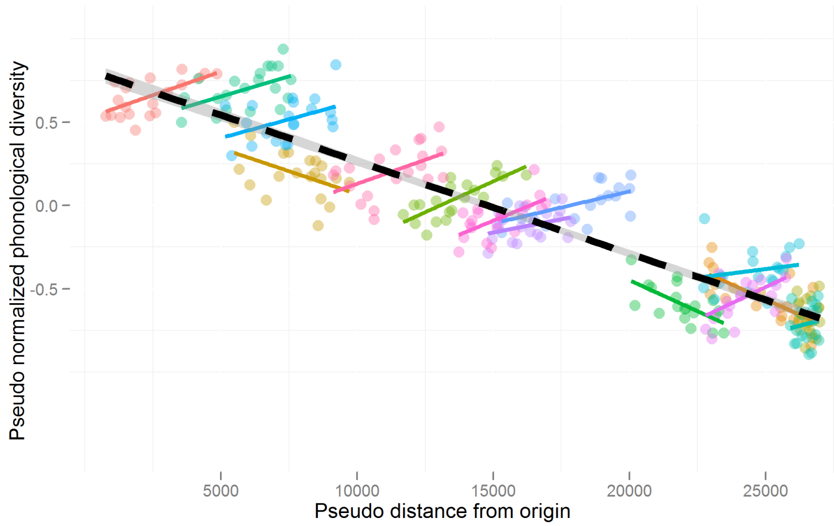
ARAMIS LAB BRAIN DATA SCIENCE — FACULTY OF APPLIED SCIENCES

Mixed-effects models

Part of you is general & Part of you is specific

$$y_i = \alpha_{pop} X^T + \alpha_i X_i^T + \epsilon$$





Derive individual patterns from other subjects

Generative models

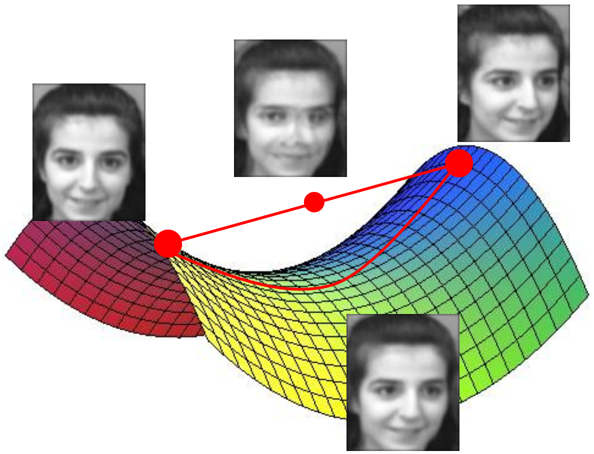
Draw some new samples from the model $p(y\,;\theta)$

Examples :

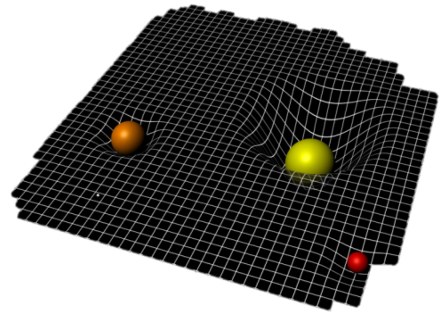
- Bayesian statistics
- Generative Adverserial Networks
- Hidden Markov Chain
- Gaussian Mixture Model

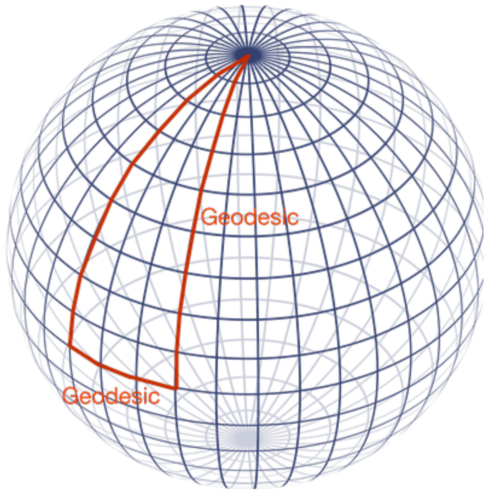
- Better understand the distribution of the observations
- Draw new samples for other algorithms (classif/regression)

Non-Euclidean data

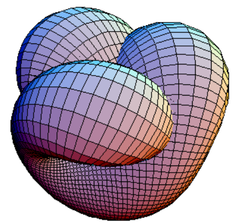


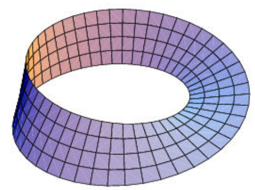
Introduce a distance between the features that are non-Euclidean

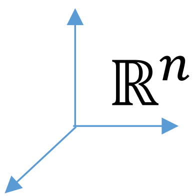


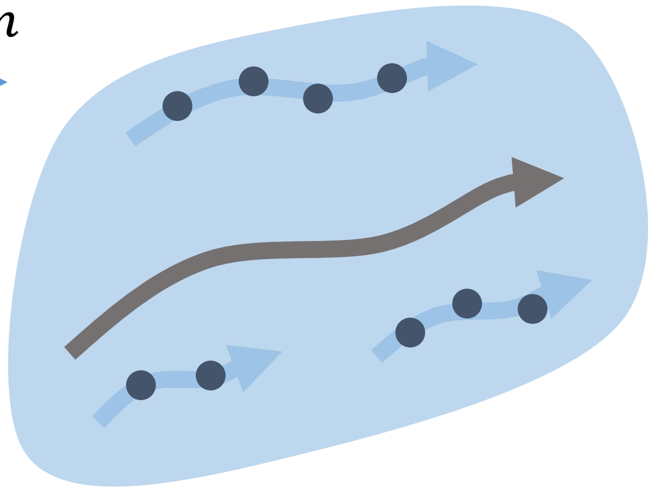


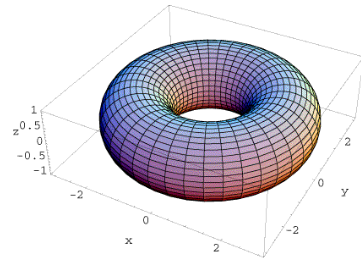
Riemannian geometry









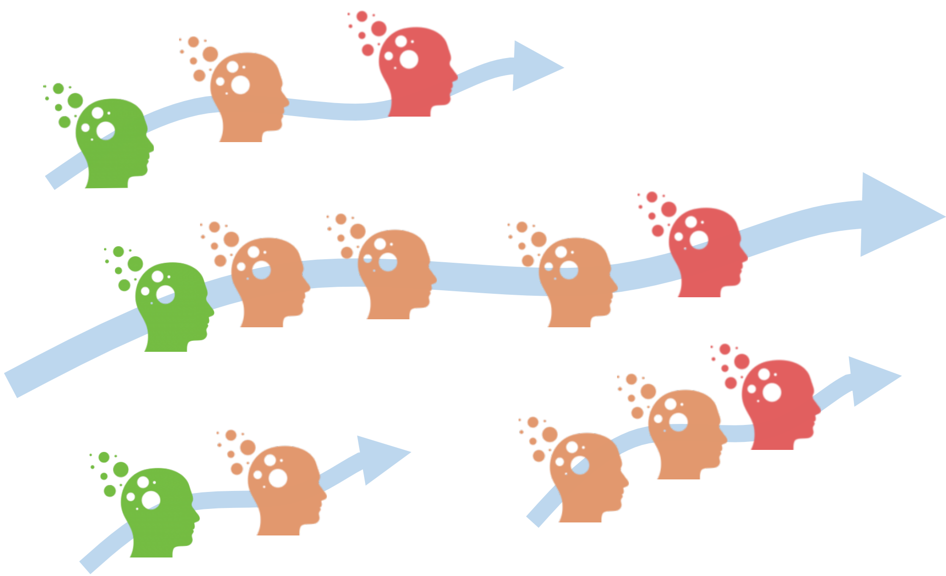


AramisLab related model

ARAMIS LAB
BRAIN DATA SCIENCE

UKRAINIAN CATHOLIC UNIVERSITY
FACULTY OF APPLIED SCIENCES

From short-term observations to long-term history

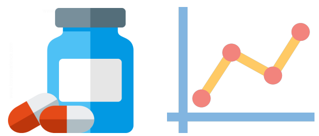


Various data:
- Biomarkers
- MRI & PET
- Meshes
- ...

Inter-individual variability
- Temporal
- Spatial

Better understand the mean disease progression

Personalize the model to charaterize individual trajectories



Predict individual disease stage and future outcomes

- Unsupervised learning
- Mixed-effects model
- Manifold Learning
- Generative model

Results : Paraview & online

II. Challenges

I. What question(s) to answer
II. Labels, variability and legal aspects
III. Companies

Challenges in Medicine

ARAMIS LAB
BRAIN DATA SCIENCE

FACULTY OF APPLIED SCIENCES

What question(s) to answer

- What does it mean to have a cancer at 92.1 %?
- It is the same to have a false positive or a false negative?
- Two paradigms : Accuracy Vs Interpretability
- Answering a clinical question or optimizing some exotic metric? (Be careful with "state-of-the-art" scores)

Labels, variability and legal aspects

Labels

- Not always « true »
- Definition may change over time
- Are they really what one ultimately wants?

Variability

- Inter individual variability
- Intra-individual variability : no one is closer to you than yourself
- Variability in the scan machine, procedures, ...

Legal aspects

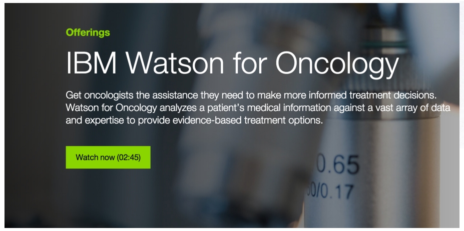
- Hard to get the data : anonymization
- The data of interest are to be asked before the clinical study

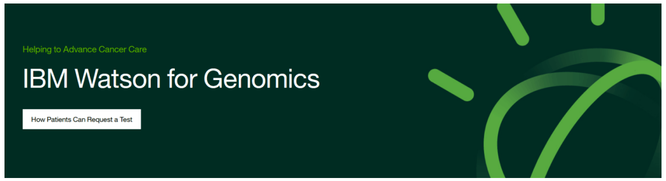
Challenges in Medicine

ARAMIS LAB
BRAIN DATA SCIENCE

FACULTY OF APPLIED SCIENCES

Companies





Growing field

Perfect fit between intellectual interest and social meaning

Still many challenges at a research level

... but be careful of the hype,
especially «Machine are better than doctors »

https://lukeoakdenrayner.wordpress.com/2016/11/27/do-computers-already-outperform-doctors/

Few words about the summer school project

Medecine is a sexy field !

Thanks