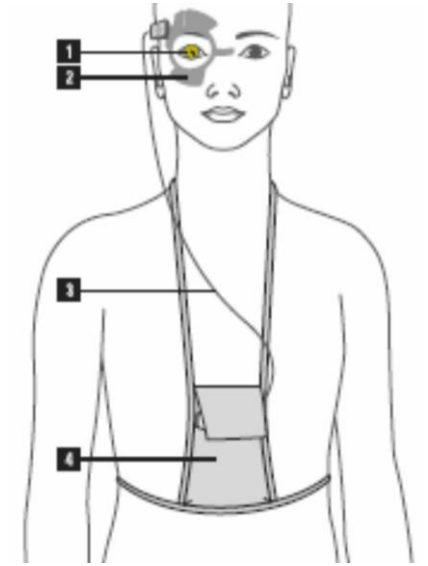
Machine Learning in MedTech

Glaucoma Prognostication Platform

Aristotelis Agianniotis, PhD EPFL
Senior academic associate, HES-SO VS

31.01.2017

Applied Machine Learning Days - SwissTech Convention Center, EPFL



To put you in context…

- Contact lens embedding a micro-sensor. (1)
- Antenna receives wirelessly the info from the contact lens. (2)
- Data transmitted to portable recorder. (3)
- Data stored locally in portable recorder during the session. (4) Later transferred to the software.

What kind of data?

- Continuous measurements
- 30s of measurements every 5min for 24h
- Metadata anonymized (e.g., health status, glaucoma type, glaucoma speed progression, others)
- More than 2000 individuals in total

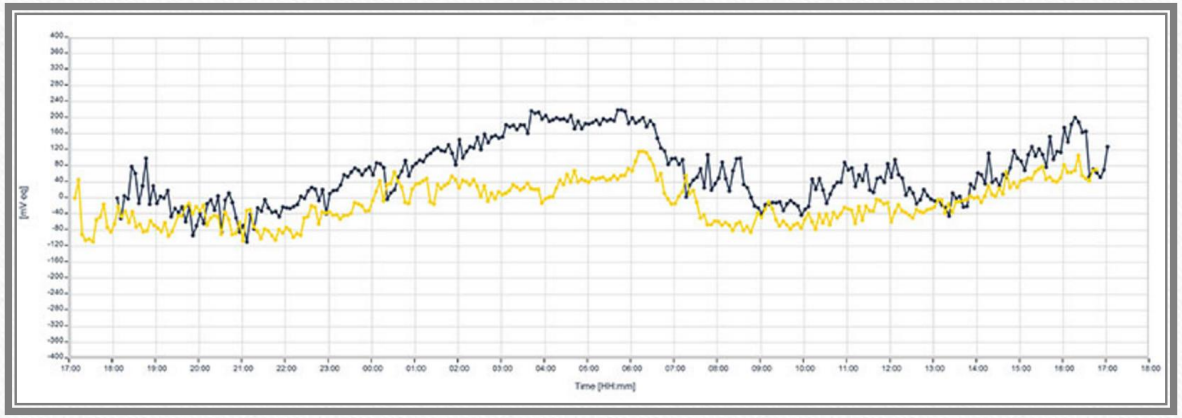
Why and How

- Why?
- 80 million people worldwide suffer from glaucoma leading to blindness
- Intraocular pressure follows individual nychthemeral patterns
- Need for continuous monitoring
- Help medical doctors in decision making

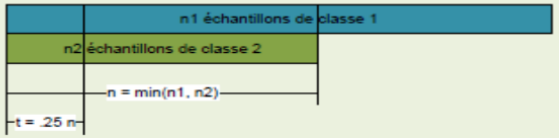
- How?
- Provide a **machine learning platform** for 24h profiles of ocular dimensional changes to predict the **progression of glaucoma**.

What

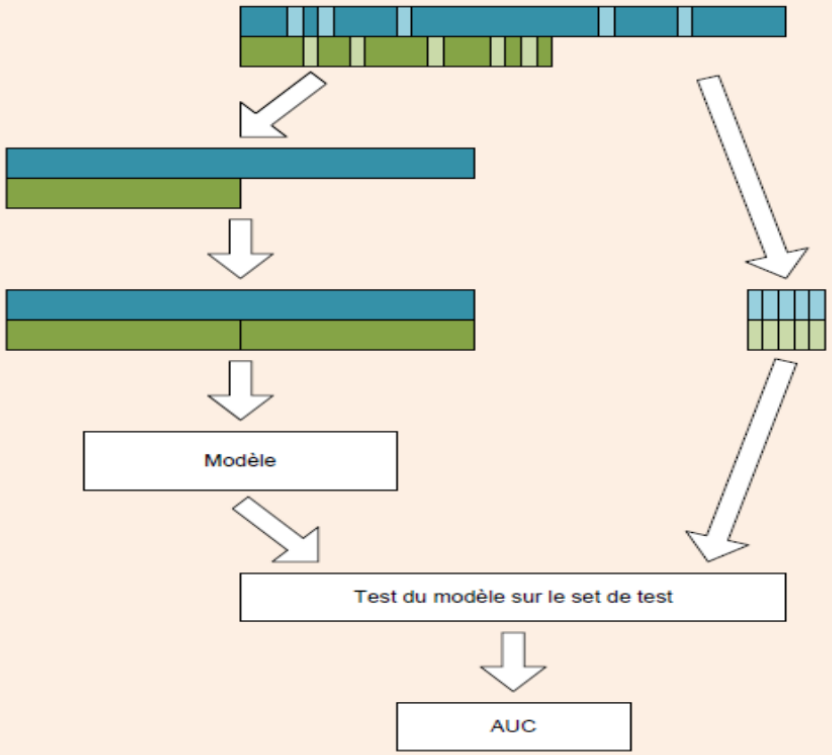
- From business needs formulate machine learning problems
- Data preparation, outliers management, etc.
- Features definition and extraction
- Dimension reduction
- Leading parameters in a profile
- Pattern recognition
- Supervised learning
  - Prediction of glaucoma visual field loss / glaucoma progression
  - Classification of different states of glaucoma
- Unsupervised learning
  - Clustering curves to glaucoma varieties



Example of curve comparison:
Two 24 hour profiles of the same patient before and one month after treatment.



n1 échantillons de la classe 1
n2 échantillons de la classe 2
Taille t de chaque classe du set de test :
t = 0.25 * min(n1, n2)



Sélection aléatoire de t échantillons de classe 1 et t échantillons de classe 2 pour faire le set de test de cette itération.
Les échantillons restants vont dans le set d'entrainement.

Sélection des caractéristiques et optimisation des hyper paramètres du modèle sur le **set d'entrainement uniquement**.

Si le set d'entrainement n'est pas balancé, répliquer les éléments du set le plus petit pour arriver à la parité.

Entrainement du modèle sur tout le set d'entrainement en utilisant les meilleurs hyper paramètres et les meilleures caractéristiques.

Test du modèle sur le set de test.

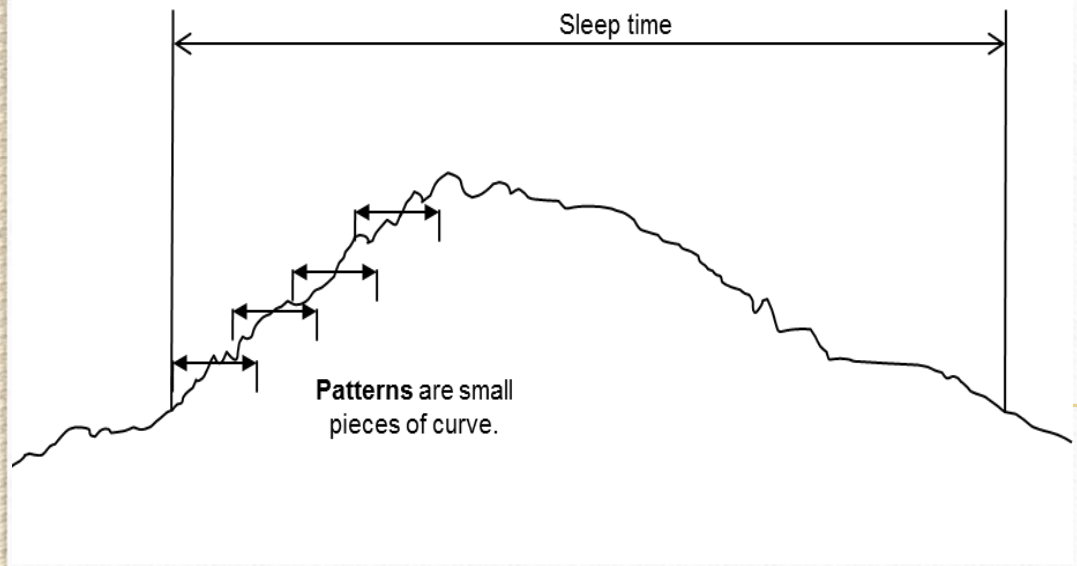
Résultat : AUC sur le set de test

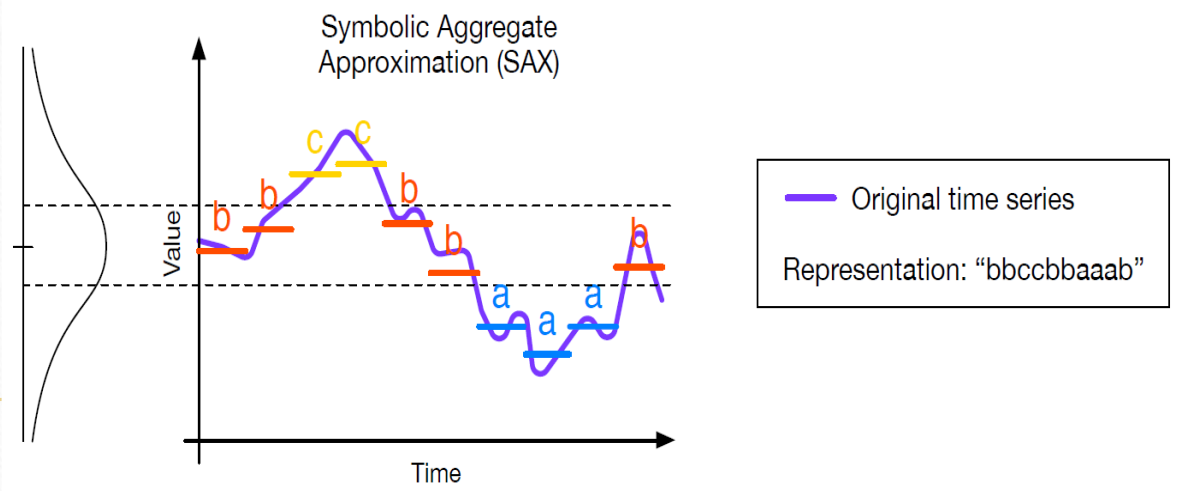
100 itérations

Distribution d'AUC

Calcul de la distribution des 100 AUC. La moyenne et l'intervalle de confiance correspondent au **score final**.

Prinipal Component Analyis, Dynamic Time Warping, clustering methods (**K-Means, Agglomerative, Birch, Gaussian Mixture Models**), **Symbolic Aggregate Approximation,** feature selection based on **Random Forest** and **Gaussian Mixture Models**, classification methods (**Random Forest, Support Vector Machine)**, regression (**linear model**, **K-Nearest Neighbors**, **Kernel Ridge Regression**, **Support Vector Regression**)





Slow vs. Fast

Initial dataset

- POAG + Horus
- Size : 142
- dB range - inf. to + 2 dB

According to manufacturer

- Normal + 2 dB to - 2 dB
- Early - 2 dB to - 6 dB
- Moderate - 6 dB to - 12 dB
- Advanced - 12 dB to - inf.

ProgressionStart \ ProgressionEnd	Normal	Early	Moderate	Advanced
Normal	5	9	10	14
Early	2	10	3	15
Moderate	10	17	2	7
Advanced	28	8	0	2

Fast?

Slow?

65 | 19 | 58

Added value

- Better understanding of glaucoma disease and glaucoma progression (visual field loss)
- New business opportunities
- Medical doctors increased interest and involvement
- Clinical trials to come
- FDA De Novo (Diurnal Pattern Recorder System)

Partners

Hes·so // VALAIS WALLIS

Hes·so // FRIBOURG FREIBURG

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Commission pour la technologie et l'innovation CTI

SENSIMED

innovation in medical micro-technology

Contact info https://ch.linkedin.com/in/aristotelisagianniotis