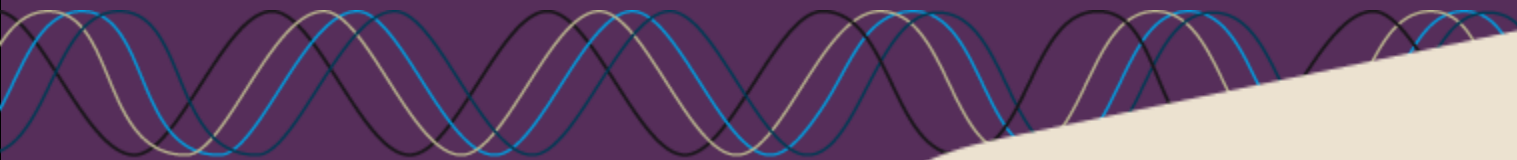


« Classification de graphes de connectivité du cerveau »

Romain Chion

encadré par: S. Achard, M. Desvignes, F. Forbes



gipsa-lab

Grenoble | images | parole | signal | automatique | laboratoire

SOMMAIRE



PRESENTATION DU CONTEXTE



METHODES USUELLES

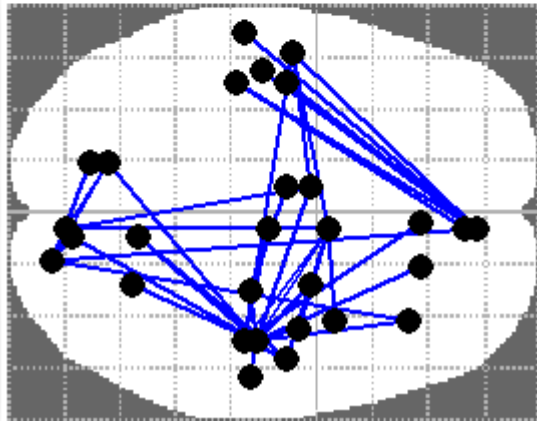
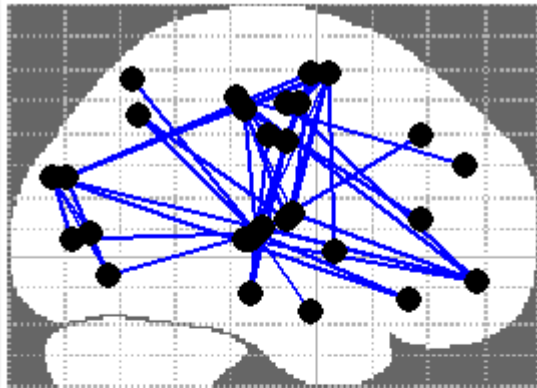


MESURE LOCALE



COMPARAISON DES RESULTATS

CONTEXTE



- Comment comparer les graphes entre eux?
- Est-il possible de modéliser les graphes de connectivité cérébrale (GCC)?
- A quel point peut-on catégoriser les GCC?

3 CONTEXTE

METHODES



MODELES GENERATIFS

- *Erdos-Renyi*
- *Forest Fire*
- *Kronecker*
- *Preferential Attachment*
- *Random k -regular*
- *Random Power Law*
- *Random Typing*
- *Small-World*

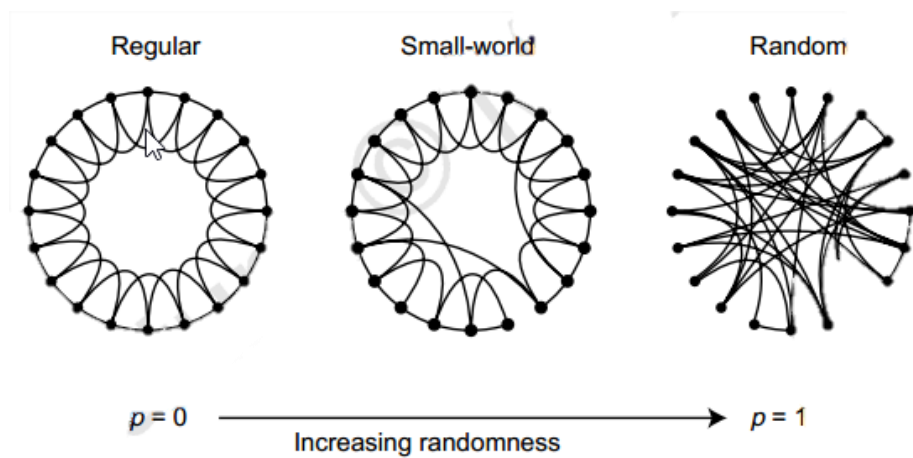


Illustration « Small World », *Collective dynamics of 'small-world' networks*, D. J. Watts & S. H. Strogatz

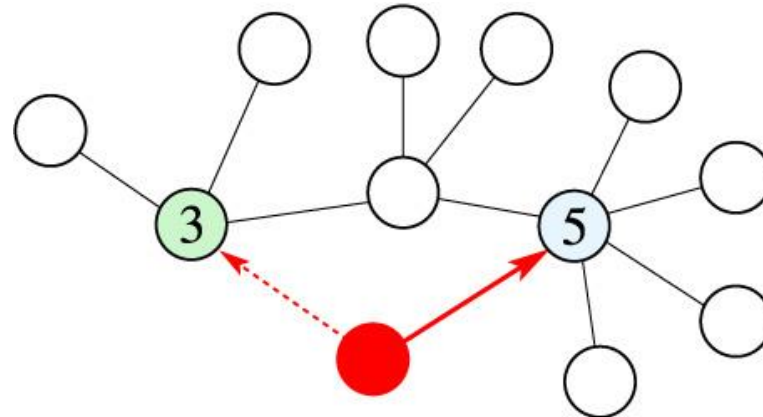


Illustration « Preferential Attachement », *Choice-driven phase transition in complex networks*, P. L. Krapivsky and S. Redner

4 CONTEXTE

METHODES



COMPARAISON DE GRAPHES

- Transformation d'un graphe vers un autre
ex : *Distance d'édition*

MESURES
STRUCTURELLES

- Tendance des nœuds à se regrouper, distribution des degrés, chemins entre nœuds
ex : *Clustering, Plus Court Chemin*

MESURES
LOCALES
(pour chaque nœud)

- Mesures locales moyennées, formation de noyaux et de communautés
ex : *Assortativité, Centralité, Modularité, Diamètre*

MESURES
GLOBALES

5 CONTEXTE

METHODES



ETAT DE L'ART : JANSSEN *et al.* 2012

Comptage de Graphlets

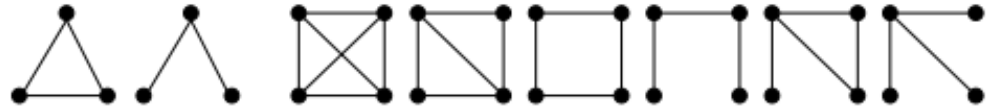
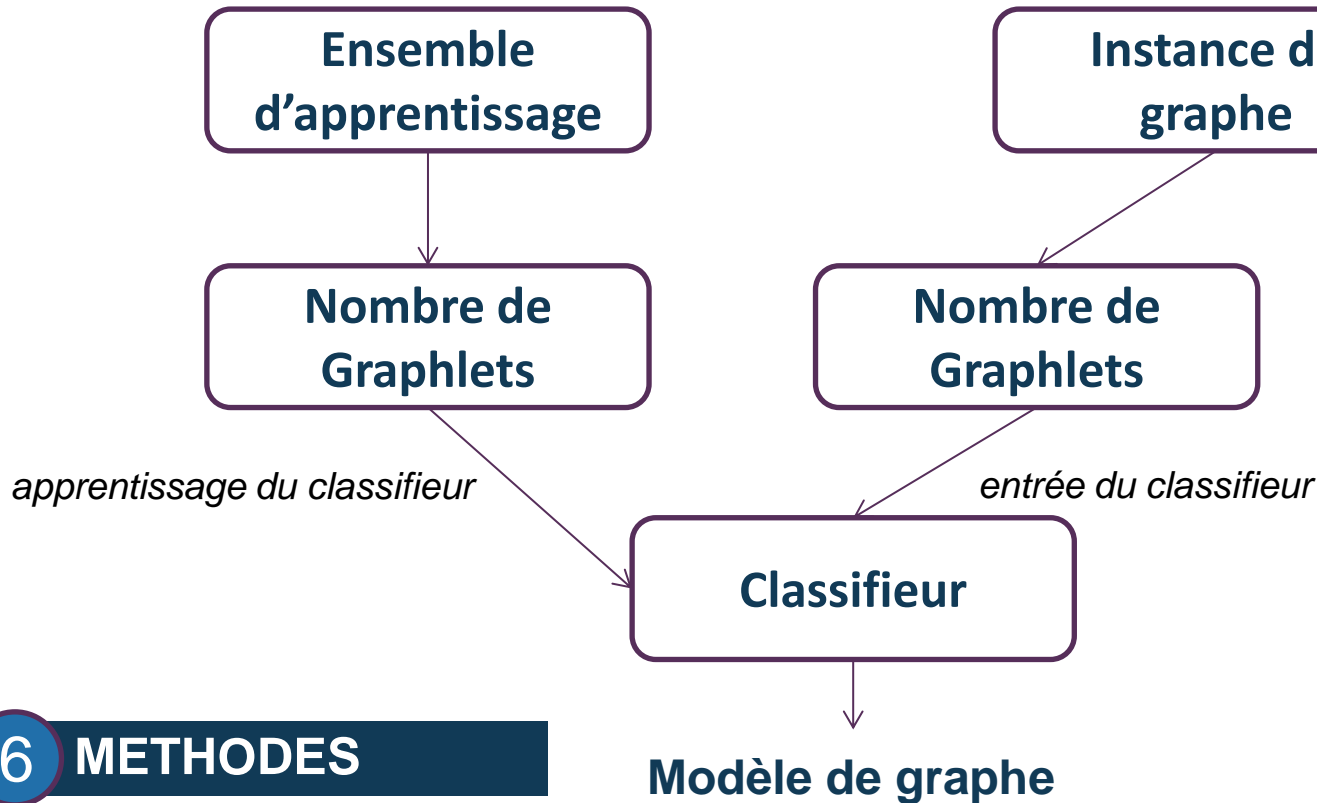


FIG. 2. The graphlets with three and four nodes



6 METHODES

MESURE LOCALE



ETAT DE L'ART : *MOTALLEBI et al. 2013*

Classifieur de Réseaux Complexes

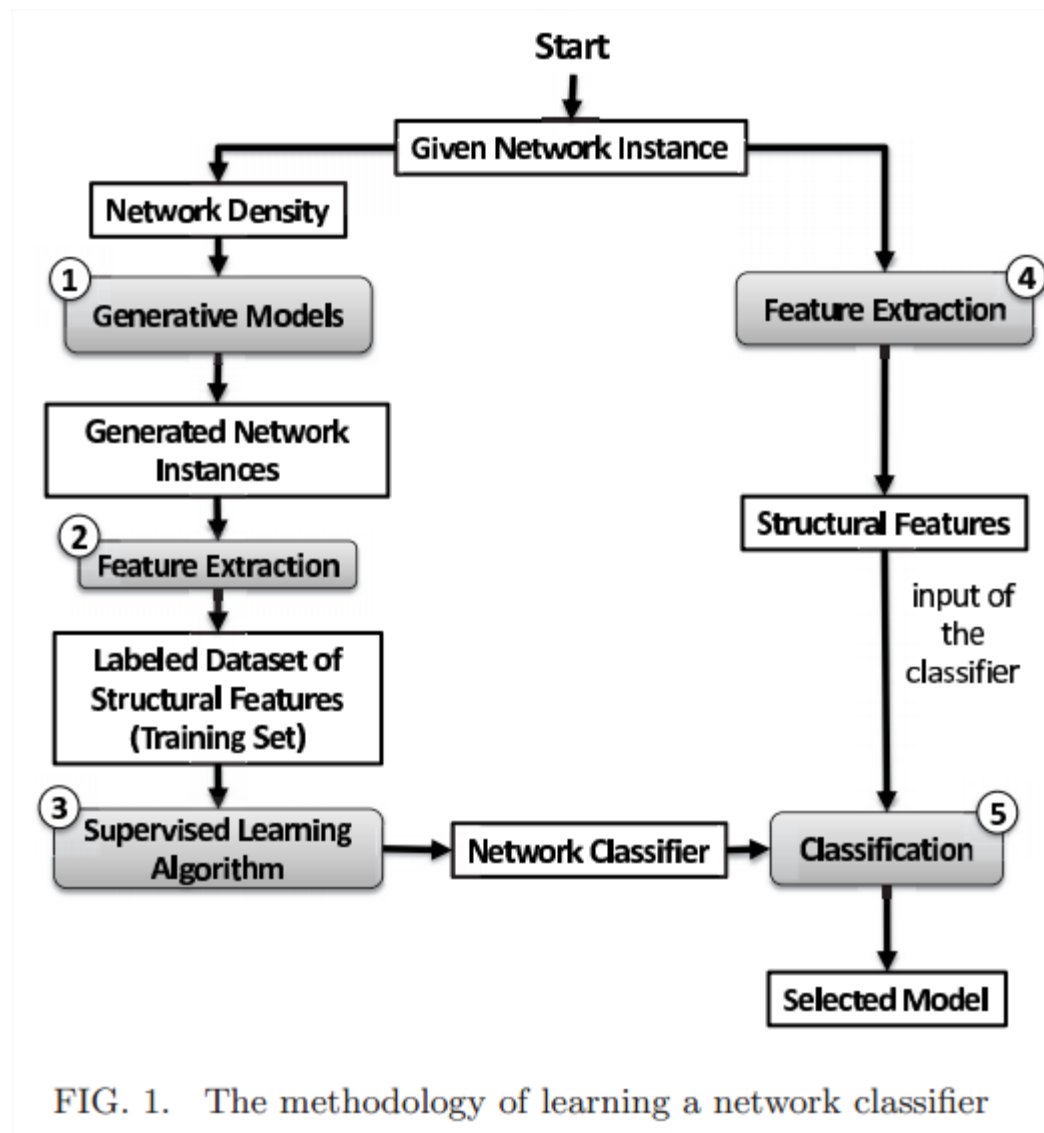


FIG. 1. The methodology of learning a network classifier



MODELISATION DES GCC

Caractérisation des GCC vers 4 modèles

(Erdos-Renyi, Preferential Attachment, Random k-regular, Small-World)

Intervalle de confiance ~25%

Classe	Prédiction	E-R	P A	R k-R	S-W
Control	Small-World	0.2502	0.2501	0.2492	0.2505
Patient	Small-World	0.2502	0.2501	0.2492	0.2505

➤ Résultat de la caractérisation avec mesures globales et classifieur SVM

8 METHODES

MESURE LOCALE



IDENTIFICATION DES GCC

Précision de la méthode à 50.16%, aléatoire à 50%

	true Control	true Patient	class precision
pred. Control	13	11	54.17%
pred. Patient	7	6	46.15%
class recall	65.00%	35.29%	50.16%

➤ Résultat de l'identification avec mesures globales et classifieur SVM

9 METHODES

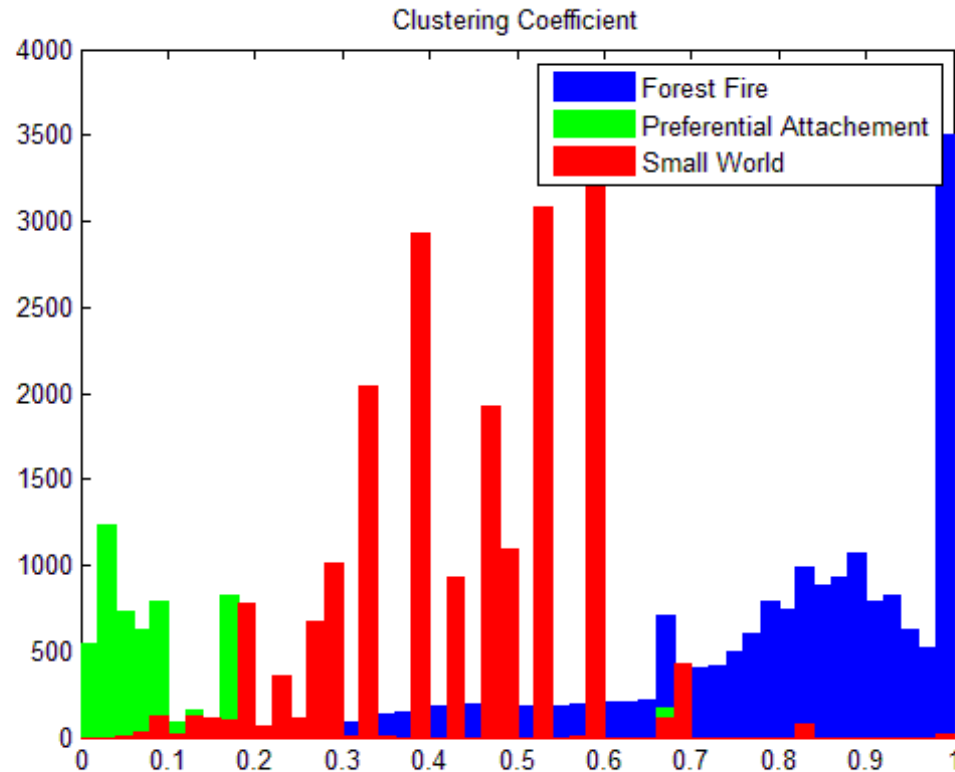
MESURE LOCALE



PROBLEMATIQUE

« *Les mesures globales ne sont pas représentatives du comportement local* »

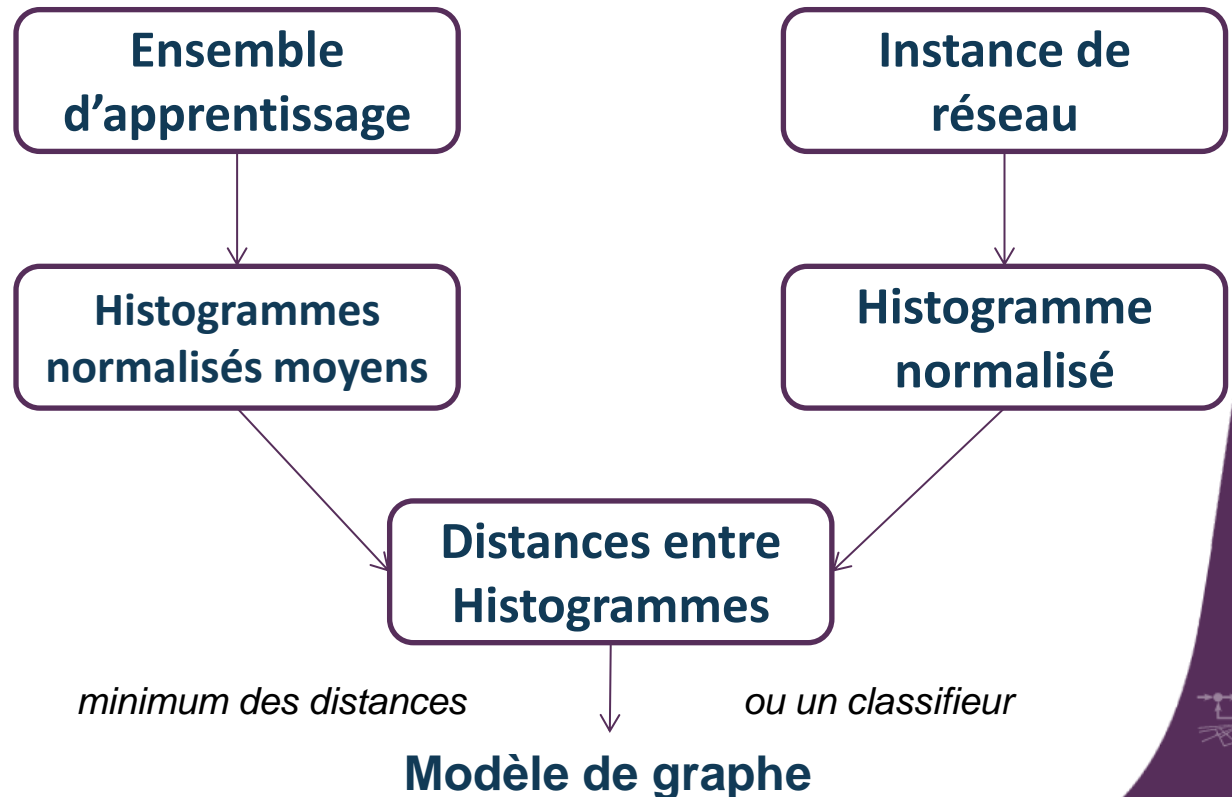
- Histogrammes du coefficient de clustering local pour 3 modèles



HISTOGRAMME NORMALISE

- **Clustering Coefficient**
- *Characteristic Path Length*
- *Degrees Distribution*
- *Efficiency*

Histogramme des mesures locales



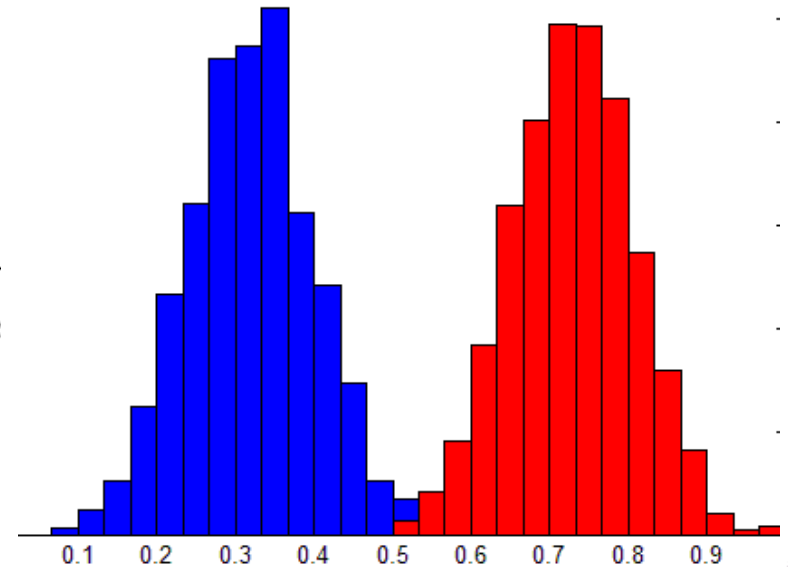
DISTANCE ENTRE HISTOGRAMMES

- Mesure de dissemblance bin à bin :

➤ **Battacharyya** : $BC(p, q) = \sum_{x \in X} \sqrt{p(x)q(x)}$

➤ **Chi²**

➤ **Hellinger** : $H(P, Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^k (\sqrt{p_i} - \sqrt{q_i})^2}$



- Mesure de dissemblance avec conservation de la silhouette :

➤ **Earth Mover Distance** : Optimisation du travail minimum qu'un cantonnier doit fournir pour transporter un tas de terre en un autre

➤ **Match** : Comparaison des histogrammes cumulés

12 MESURE LOCALE

RESULTATS



DONNEES DE SYNTHESSES

Performances

- *graphlets* : **78%**
- *mesures globales* : **88% à 97.3%** 6 mesures voire plus
- *mesures locales* : **86% ou 100%** 1 seule mesure

	Precision
<i>SW</i>	100%
<i>RPL</i>	100%
<i>RkR</i>	100%
<i>PA</i>	100%
<i>KG</i>	100%
<i>FF</i>	100%
<i>ER</i>	100%
	100%

mesures locales

Résultat de la classification

	Precision
<i>SW</i>	100%
<i>RTG</i>	96%
<i>RPL</i>	98%
<i>PA</i>	99%
<i>KG</i>	96%
<i>FF</i>	98%
<i>ER</i>	93%
	97.2%

mesures globales



GRAPHES DE CONNECTIVITES

mesures globales 63% v.s. 83% histogrammes

		GLOBALES		
		<i>C</i>	<i>P</i>	
A.N.N.	<i>C</i>	11	9	55%
	<i>P</i>	5	12	71%
		69%	57%	63%

		HISTOGRAMME		
		<i>C</i>	<i>P</i>	
CLUSTERING ET CHI ²	<i>C</i>	18	2	90%
	<i>P</i>	4	13	76%
		82%	87%	83%

➤ Matrice de confusion de l'identification Control / Patient



MODELISATION DES GCC

7	Clustering Degrés	
ER	0,418	0,133
FF	0,207	0,074
KG	0,112	0,211
RPL	0,156	0,088
PA	0,437	0,242
RkR	0,459	0,183
SW	0,103	0,238

- Distance EMD entre GCC et modèles pour deux mesures locales



CONCLUSION

- De bonnes performances sur les graphes de synthèse
- Les histogrammes locaux sont importants
- Le clustering local est particulièrement intéressant
- Dépendant du nombre et du choix des modèles

- Les résultats sur les données réelles sont à approfondir
- Une combinaison des modèles est à envisager

