

Analyse de la causalité pour l'aide à la décision dans un contexte de supervision pour l'industrie 4.0

Kenza AMZIL

Directeur de thèse :

M. Lionel ROUCOULES

Encadrantes :

Mme. Esma Yahia

Mme. Nathalie Klement

04/12/2020

LA SUPERVISION DES SYSTÈMES DE PRODUCTION POUR L'AIDE À LA DÉCISION



La **supervision** consiste à surveiller le fonctionnement des systèmes de production et leur environnement, alerter des éventuelles **dérives**, et pouvoir les **éviter**.



La **supervision** doit permettre une prise de décision :

- Dérives de processus ou système de production ;
- Optimisation du rendement de production ;
- Amélioration de la qualité ;
- Optimisation des coûts.

Contexte général et
problématique

Etat de l'art et
proposition

Cas d'étude

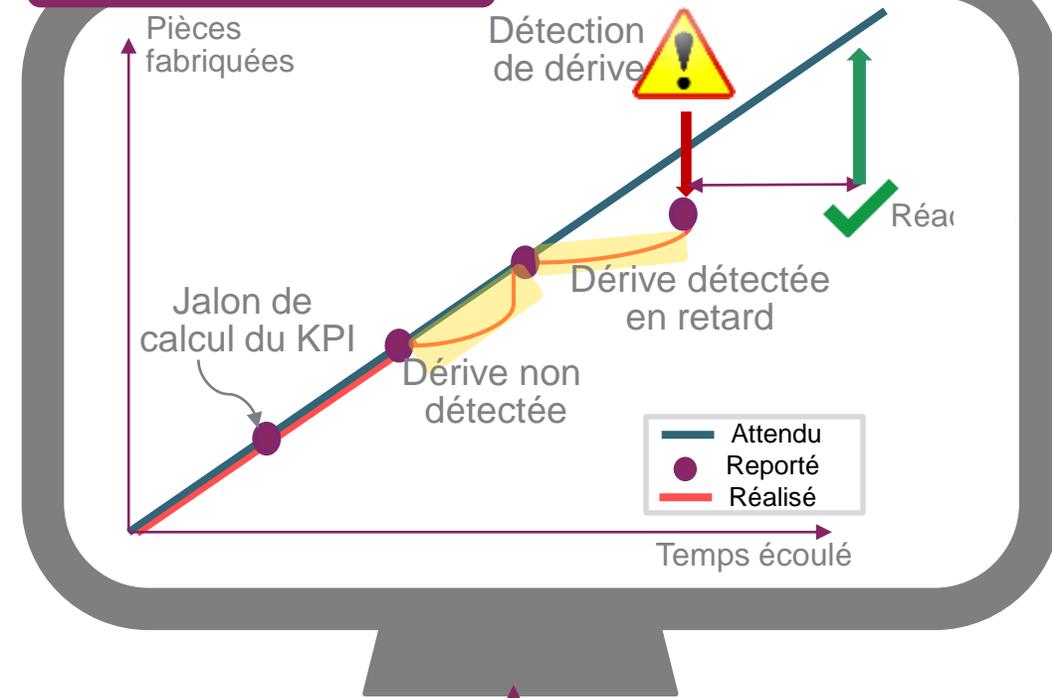
Conclusion et
perspectives

LA SUPERVISION DES SYSTÈMES DE PRODUCTION POUR L'AIDE À LA DÉCISION

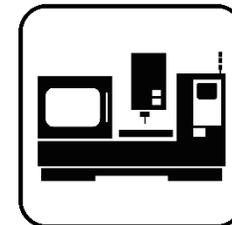
- Il faut **prévoir les dérives** et les détecter en temps réel.
- **Identifier les liens de causalité** entre les KPI et les données, de façon **exhaustive et objective**.
- **Prioriser causes identifiées** pour agir efficacement.

Reporting

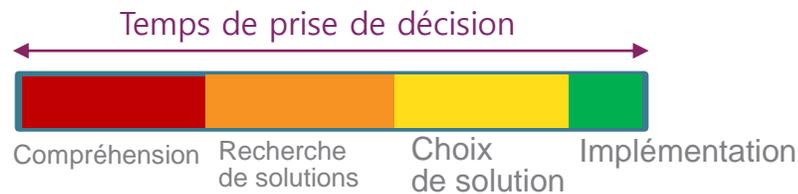
KPI : quantité fabriquée



Récupération des données



Sources de données



Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives

LA SUPERVISION DES SYSTÈMES DE PRODUCTION POUR L'AIDE À LA DÉCISION

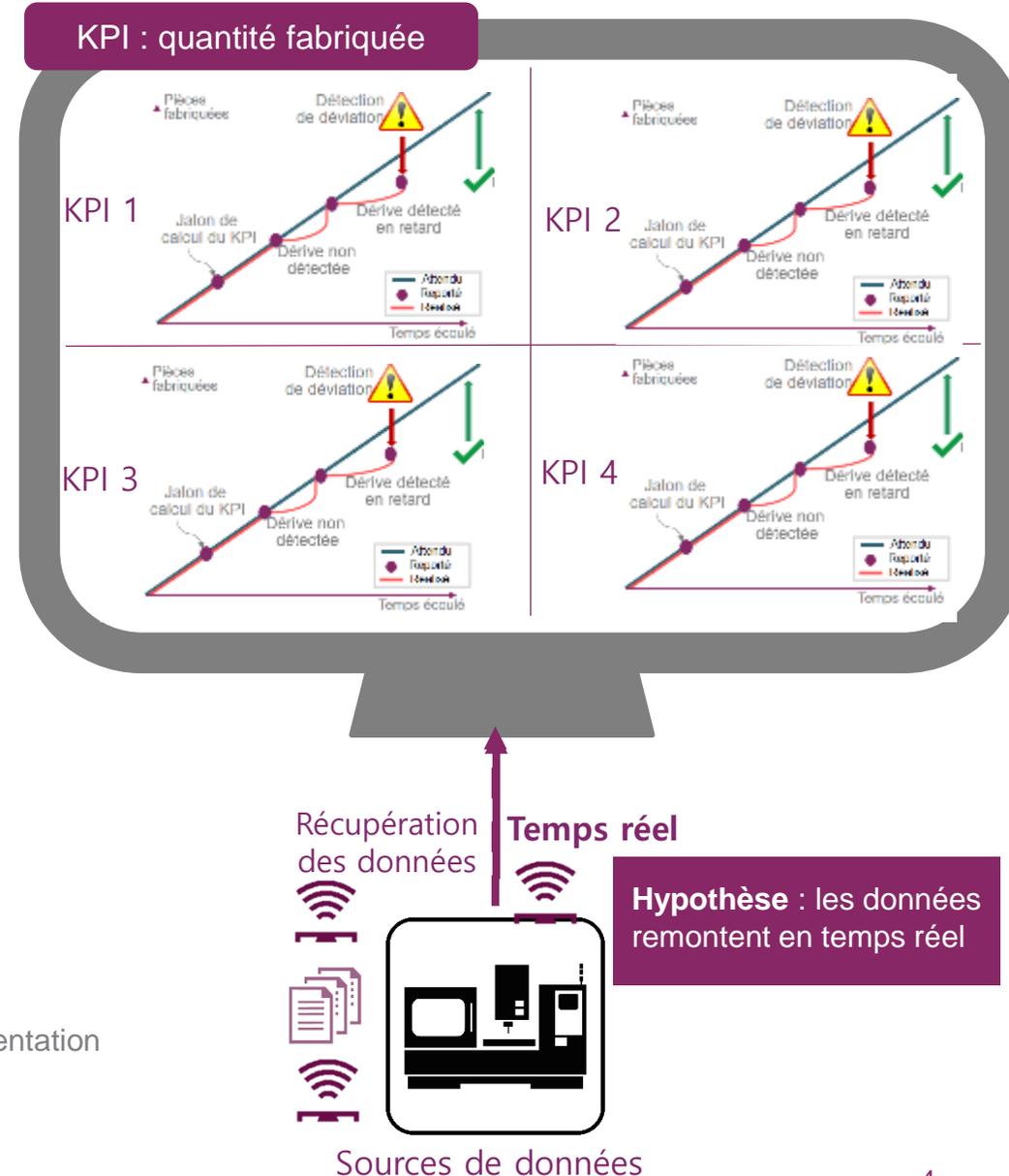
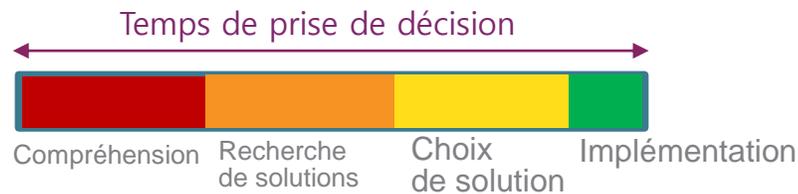
- Il faut **prévoir les dérives** et les détecter en temps réel.
- **Identifier les liens de causalité** entre les KPI et les données, de façon **exhaustive et objective**.
- **Prioriser causes identifiées** pour agir efficacement.
- Il faut que l'approche soit **générique**

Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

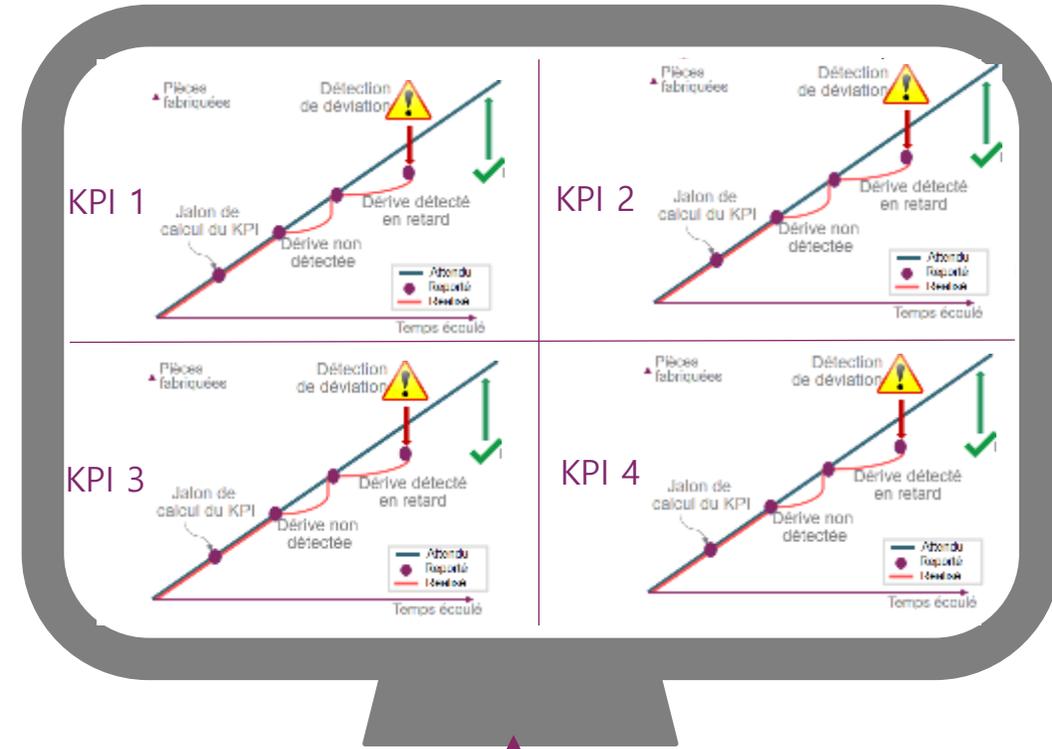
Conclusion et perspectives



LA SUPERVISION DES SYSTÈMES DE PRODUCTION POUR L'AIDE À LA DÉCISION

- Il faut **prévoir les dérives** et les détecter en temps réel.
- **Identifier les liens de causalité** entre les KPI et les données, de façon **exhaustive et objective**.
- **Prioriser causes identifiées** pour agir efficacement.
- Il faut que l'approche soit **générique**

Postulat : L'environnement de production peut être équipé de capteurs et de sources de données dans un contexte industrie 4.0

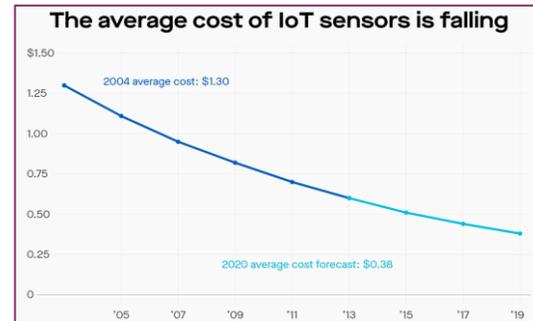
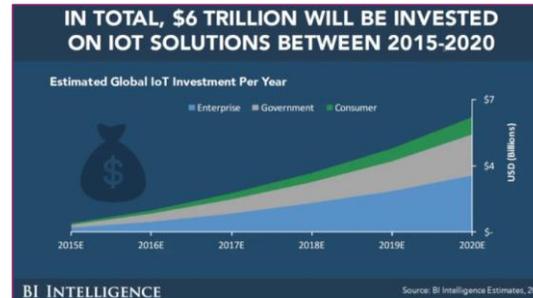


Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives



Récupération des données Temps réel

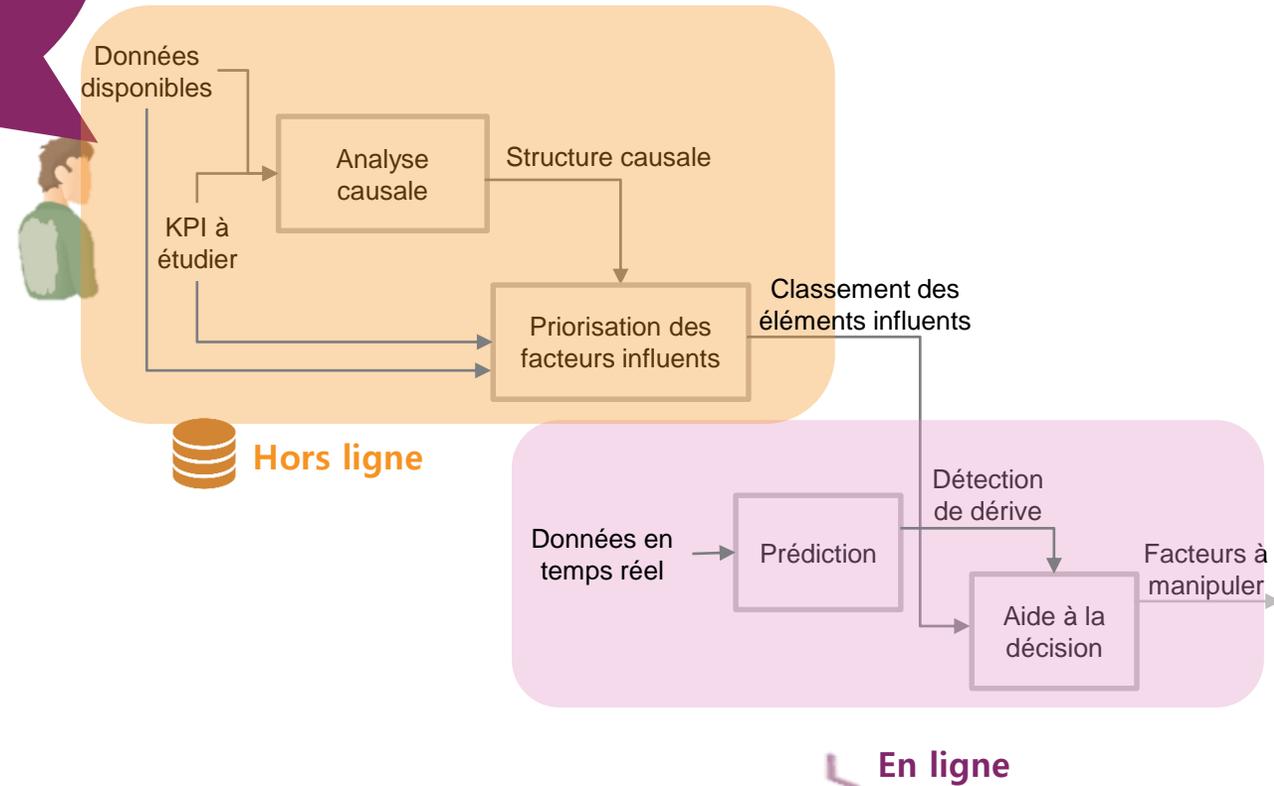


Sources de données

Hypothèse : les données remontent en temps réel

LA SUPERVISION DES SYSTÈMES DE PRODUCTION POUR L'AIDE À LA DÉCISION

- Il faut **prévoir les dérives** et les détecter en temps réel.
- **Identifier les liens de causalité** entre les KPI et les données, de façon **exhaustive et objective**.
- **Prioriser causes identifiées** pour agir efficacement.
- Il faut que l'approche soit **générique**



Contexte général et problématique

Etat de l'art et proposition

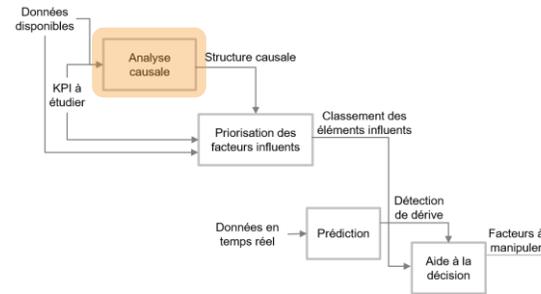
Cas d'étude

Conclusion et perspectives

ANALYSE CAUSALE : CARACTÉRISTIQUES DE LA CAUSALITÉ



La notion de causalité, pourtant abordée par plusieurs domaines de recherche, n'a pas encore de définition figée et vérifiée.



Caractéristiques de la causalité :

- Aspect probabiliste ;
- Ordre temporel ;
- Dépendance de l'espace de recherche
- Représentation graphique;
- Taille de l'effet ;
- Stabilité ;
- Plausibilité ;
- Vérificabilité;

- Acyclicité ;
- Consistance ;
- Régularité ;
- Relation dose-effet ;
- Spécificité

Caractéristiques non retenues dans notre contexte

Ces caractéristiques, en cas de leur vérification, représentent une série d'arguments; qui, ajoutés les uns aux autres, nous permettrons d'approcher la présomption de relation causale.

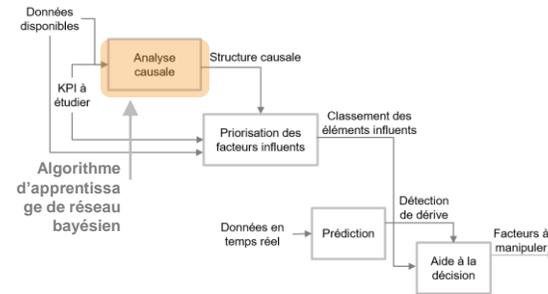
Contexte général et problématique

Etat de l'art et proposition

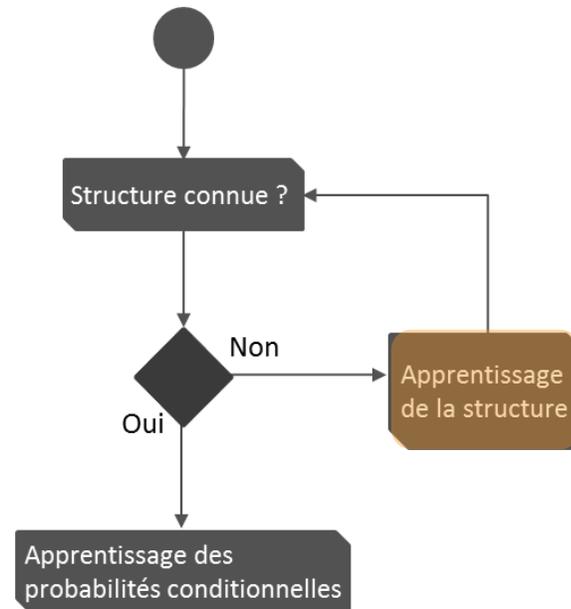
Cas d'étude

Conclusion et perspectives

ANALYSE CAUSALE : APPRENTISSAGE DE RÉSEAU BAYÉSIEN



Apprentissage de réseau bayésien



Prétraitement des données

Apprentissage de la structure

Recherche d'indépendances conditionnelles (Pearl et Verma)

Méthodes basées sur un score

- Explosion du nombre de tests d'indépendance à effectuer.
- A quel ordre s'arrêter ?
- Sensibilité au seuil du test d'indépendance
- L'orientation des arcs est faite à posteriori

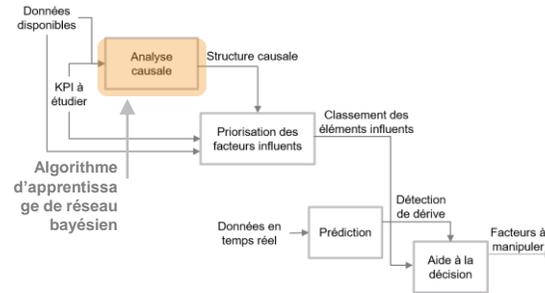
Contexte général et problématique

Etat de l'art et proposition

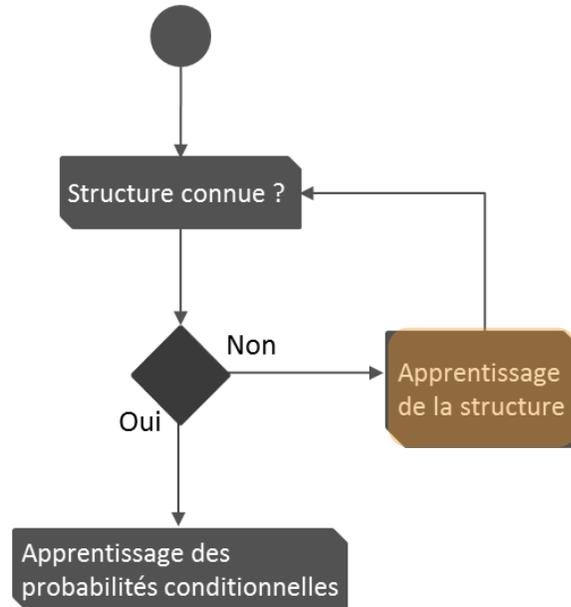
Cas d'étude

Conclusion et perspectives

ANALYSE CAUSALE : APPRENTISSAGE DE RÉSEAU BAYÉSIEN



Apprentissage de réseau bayésien



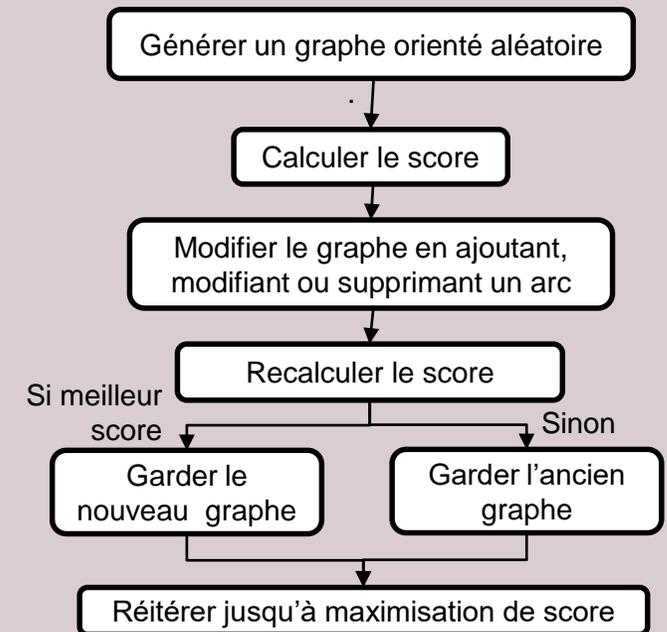
Prétraitement des données

Apprentissage de la structure

Recherche d'indépendances conditionnelles (Pearl et Verma)

Méthodes basées sur un score

Fonction de score basée sur la vraisemblance que les données observées soient générées à partir du graphe évalué



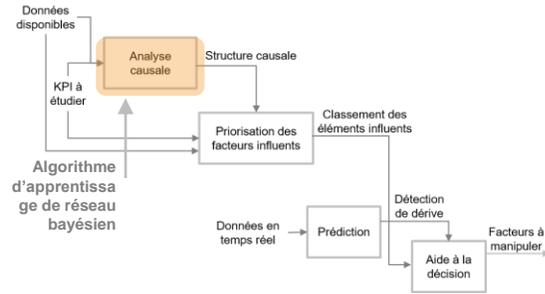
Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives

ANALYSE CAUSALE : APPRENTISSAGE DE RÉSEAU BAYÉSIEN



Prétraitement des données

Apprentissage de la structure

Méthodes basées sur un score

Apprentissage de la structure

Algorithme Hill climbing

Générer un graphe orienter aléatoire

Calculer le score

Modifier le graphe en ajoutant, modifiant ou supprimant un arc

Recalculer le score

Si meilleur score

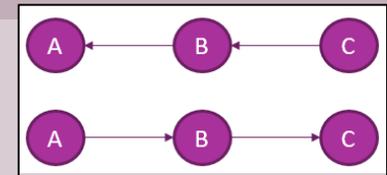
Garder le nouveau graphe

Sinon

Garder l'ancien graphe

Réitérer jusqu'à maximisation de score

Sélection des meilleurs graphes



Éliminer les graphes contenant des relations impossibles

Garder le graphe contenant des relations causales déjà connues

Algorithme Hill climbing

Algorithme Hill climbing

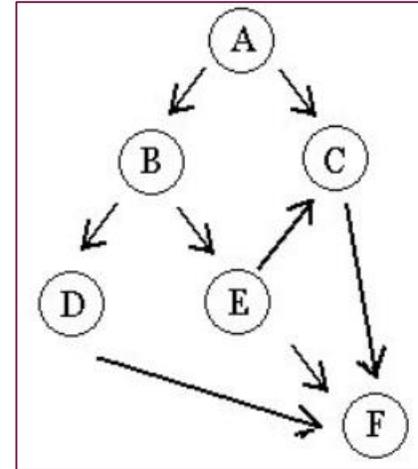
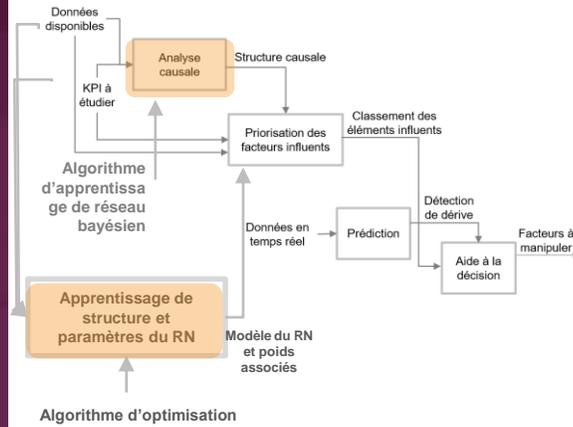
Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives

PRIORISATION DES ÉLÉMENTS INFLUENTS



C	E	F=5	F=10
1	1	0,95	0,05
1	0	0,64	0,36
0	1	0,35	0,65
0	0	0,20	0,80

Les probabilités ne sont pas faciles à interpréter quand il s'agit de classer les variables indépendamment des valeurs qu'elles peuvent prendre.

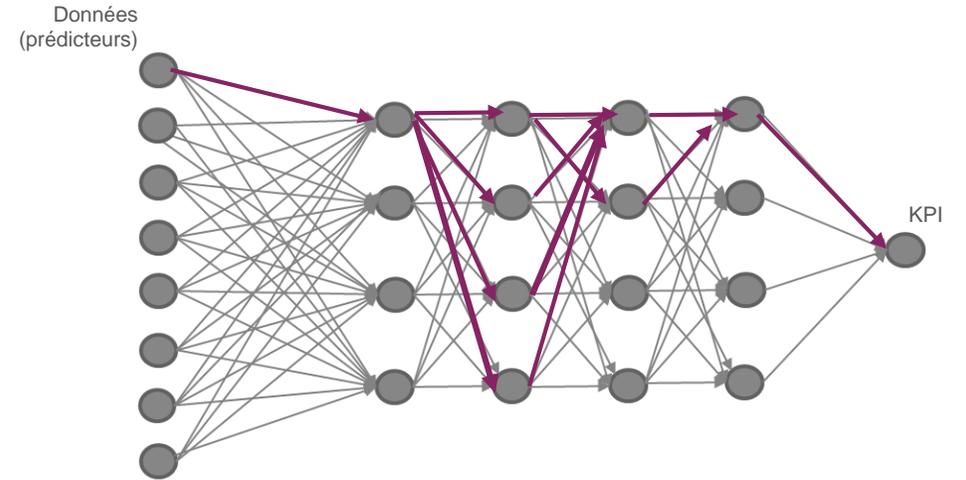
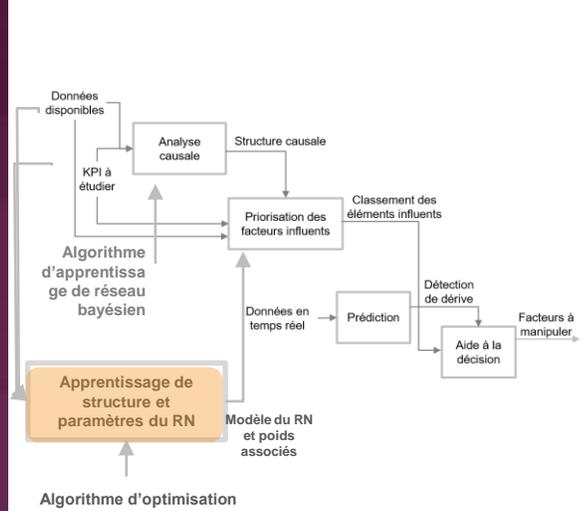
Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives

PRIORISATION DES ÉLÉMENTS INFLUENTS

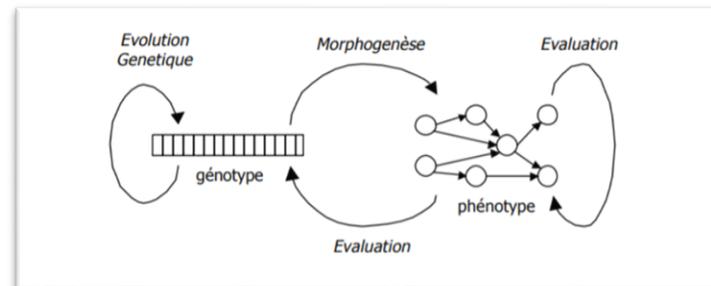
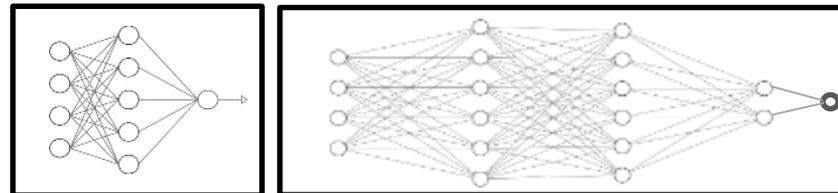


Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives



Algorithme génétique

Généricité

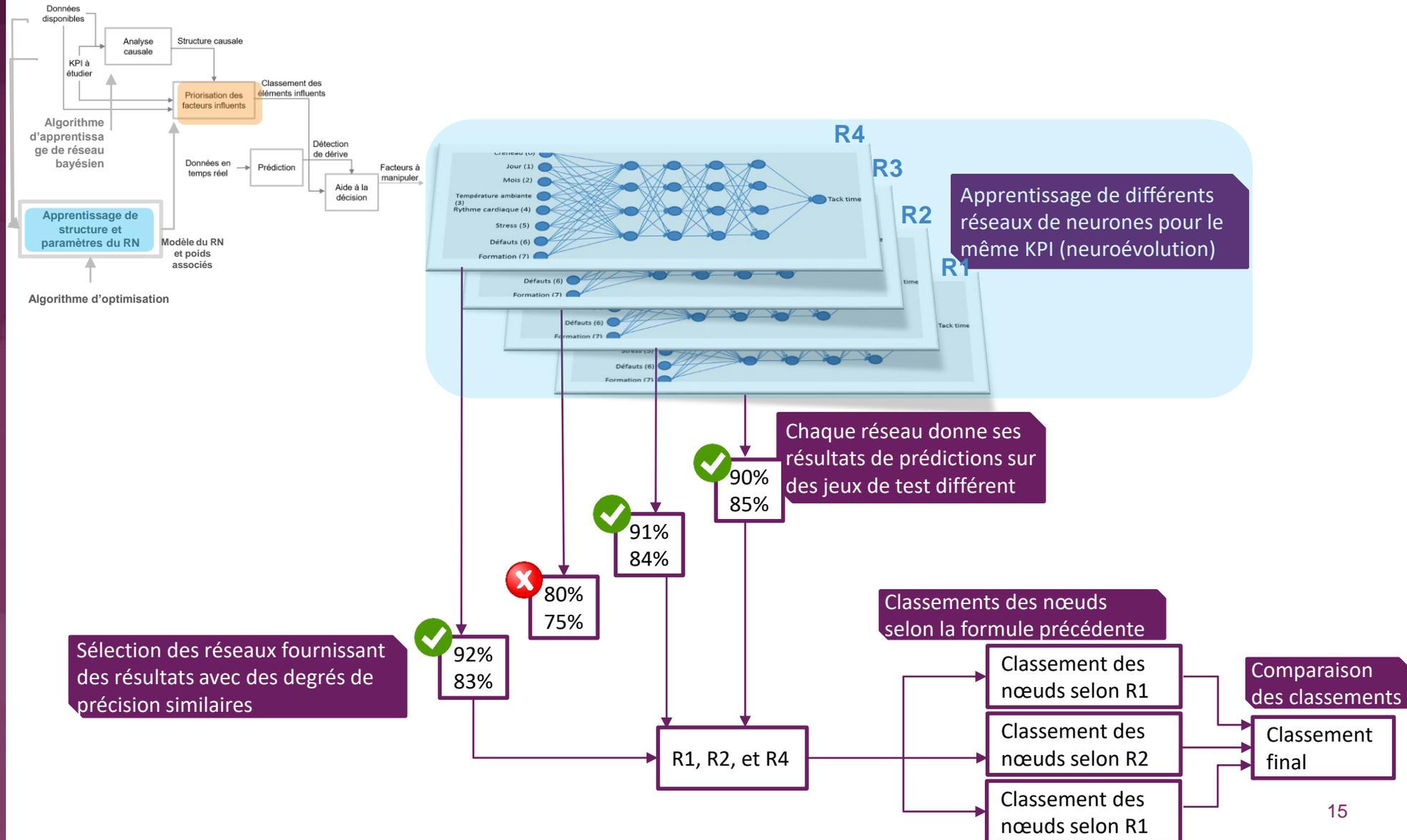
Classer les éléments influents en se basant sur les poids finaux d'un réseau de neurones avec un bon pouvoir de prédiction

Trouver la structure et les paramètres RN

- Approche expérimentale composée d'une **série d'affinages successifs** des paramètres pour avoir une prédiction optimale
- ⇒ **Pour chaque KPI, il faut trouver la structure permettant la meilleure prédiction**

Neuroévolution : Algorithme d'optimisation permettant d'automatiser le processus de construction des réseaux de neurones.

PRIORISATION DES ÉLÉMENTS INFLUENTS : VALIDATION DU CLASSEMENT



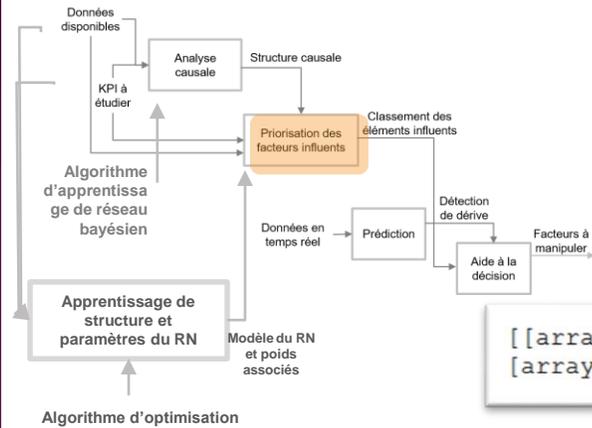
Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives

PRIORISATION DES ÉLÉMENTS INFLUENTS : VALIDATION DU CLASSEMENT



```
[[array([2021.85]), 4], [array([1967.68]), 0], [array([1934.16]), 3], [array([1861.07]), 1],
[array([1852.13]), 2], [array([1839.75]), 5], [array([1779.06]), 6], [array([1724.95]), 7]]
```

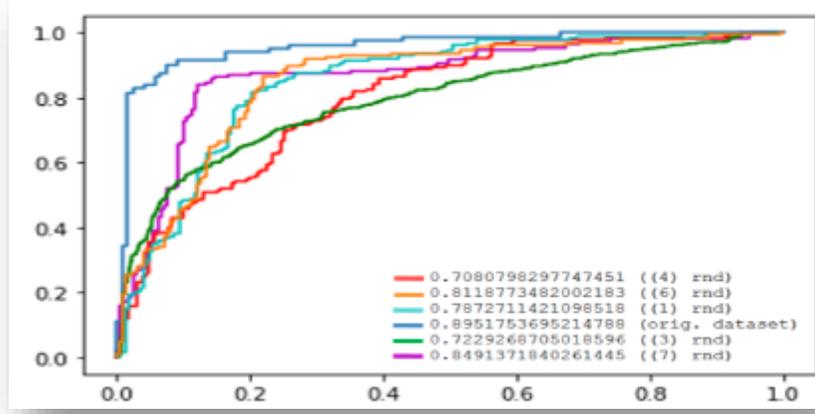
Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives

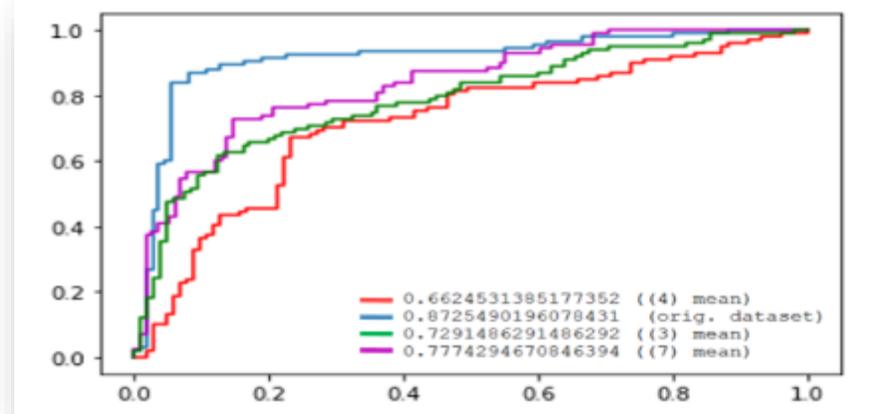
Sensibilité (Taux de vrais positifs)



1-spécificité (Taux de faux positifs)

Courbes ROC de prediction après remplacement de chaque variable par des valeurs aléatoires de son intervalle de variation.

Sensibilité



1-spécificité

Courbes ROC de prediction après remplacement de chaque variable par sa moyenne.

PRÉDICTION ET AIDE À LA DÉCISION

- Il faut **prévoir les dérives** et les détecter en temps réel.
- **Identifier les liens de causalité** entre les KPI et les données, de façon **exhaustive et objective**.
- **Prioriser causes identifiées** pour agir efficacement.
- Il faut que l'approche soit **générique**

Caractéristiques de la causalité :

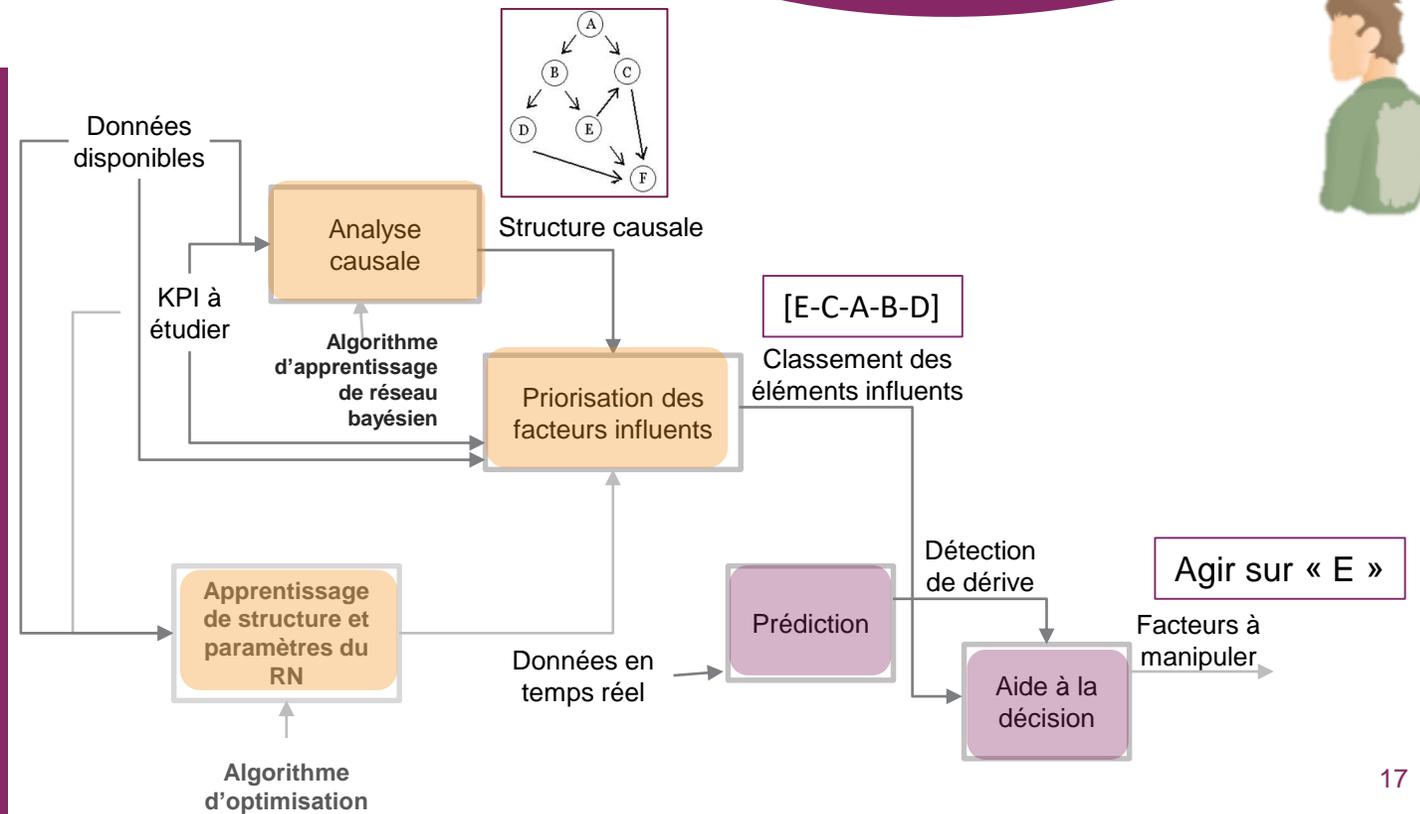
- ✓ - Aspect probabiliste ;
- ✓ - Ordre temporel ;
- ✓ - Dépendance de l'espace de recherche
- ✓ - Représentation graphique;
- ✓ - Taille de l'effet ;
- ✓ - Stabilité ;
- ✓ - Plausibilité ;
- ✓ - Vérificabilité;

Contexte général et problématique

Etat de l'art et proposition

Cas d'étude

Conclusion et perspectives



CONCLUSION ET PERSPECTIVES

Contexte général et problématique

Etat de l'art et proposition

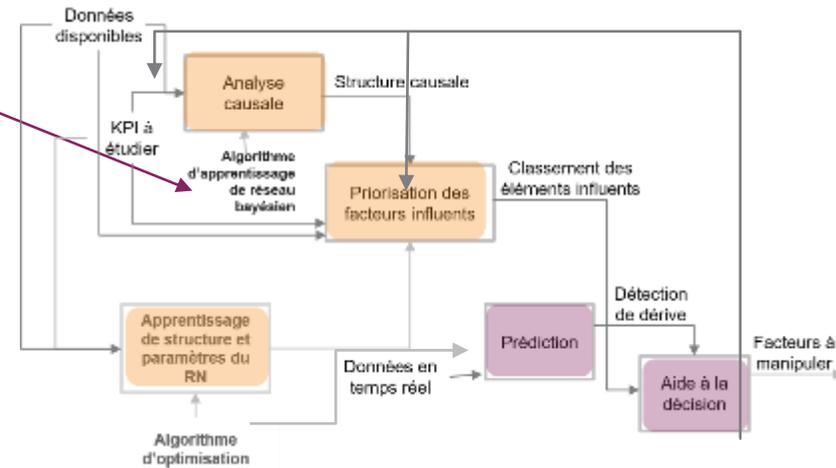
Cas d'étude

Conclusion et perspectives

Automatisation de sélection du réseau bayésien à retenir parmi plusieurs réseaux équivalents

Mise en lien des différentes briques

Application sur cas d'étude réel (En cours sur l'usine école à Aix, et la plateforme « Usine du futur » au campus de Lille)



MERCI
😊