



# **Apport et limitation de l'Intelligence Artificielle au diagnostic et au pronostic**

André Cabarbaye


**CAB INNOVATION**


3 rue de la coquille, 31500 Toulouse – Tel : 05 61 54 68 08

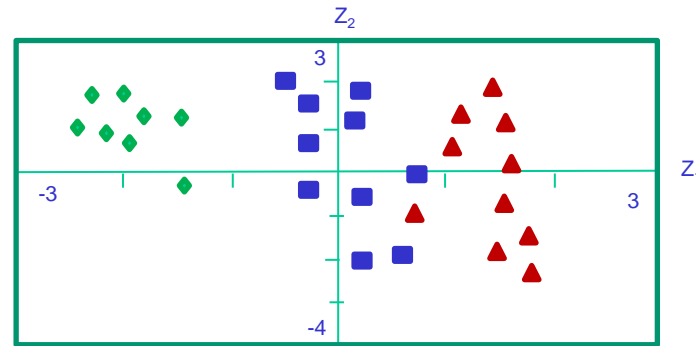
Courriel : [contact@cabinnovation.com](mailto:contact@cabinnovation.com) Site : [cabinnovation.com](http://cabinnovation.com)

- La maintenance prévisionnelle (ou prédictive) est subordonnée à la surveillance de l'évolution du niveau de dégradation d'une entité afin de retarder les interventions ou inversement de les avancer pour éviter des dysfonctionnements.
- Cette application du Health Monitoring (HM) constitue le Graal des opérateurs de systèmes car elle permet d'améliorer la disponibilité et la sécurité tout en réduisant les coûts d'exploitation, la consommation de ressources et la production de déchets.
- La maintenance prédictive est fondée sur un diagnostic (état de dégradation), un pronostic (potentiel restant) et un critère de décision d'intervention.
- L'intelligence artificielle constitue-t-elle la panacée promise par certains experts en mégadonnées (ou data scientists) ?
- Comprendre les principes de l'IA et des traitements de données permet d'en appréhender les capacités et limitations.

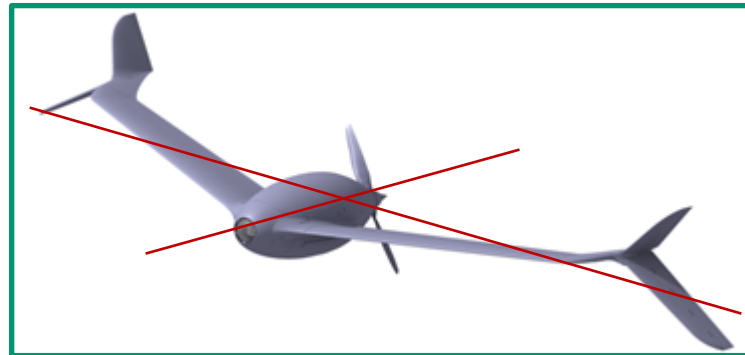
- Le traitement de données renvoie à une série de processus permettant d'extraire de l'information ou de produire du savoir à partir de données brutes.
- Ce traitement peut porter sur la fusion de données, l'extraction d'information ou la transformation de représentation :
  - La fusion de données consiste à combiner différentes sources de données afin de les compiler en une information plus sûre et plus précise dans toutes les conditions d'acquisition (données GPS et centrale inertielle par exemple)
  - L'exploration de données (fouille de données, prospection de données, data mining), a pour objet d'extraire de la connaissance, à partir d'une grande quantité de données, et de caractériser des phénomènes complexes pour mieux les comprendre.

- 
- L'exploration de données permet de résoudre des problèmes divers dans différents domaines, tels que :
    - la police pour élucider des affaires criminelles,
    - le marketing pour identifier des similarités de comportement et de consommation parmi les prospects et réduire ainsi le coût d'acquisition de nouveaux clients,
    - les centres d'appel pour adapter les réponses apportées et améliorer la qualité du service,
    - les banques et assurances pour repérer les clients profitables sans facteur de risque (scoring),
    - la recherche médicale pour découvrir le rôle des gènes dans le génome humain,
    - la reconnaissance d'images,
    - la maintenance pour détecter et localiser des défaillances, etc.

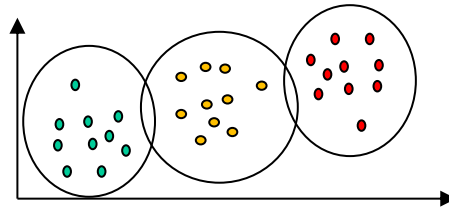
- 
- L'exploration est menée dans un but descriptif ou prédictif.
  
  - Outre la formulation de la problématique et le choix des données appropriées pour la résoudre, les traitements suivants peuvent être mis en œuvre :
    - préparation par suppression des doublons et des valeurs aberrantes,
    - discrétisation ou synchronisation sur des instants ou situations identiques ou similaires (phase de vol, état énergétique...),
    - complétude par ajout des données non renseignées (moyenne, plus proches voisins, interpolation...),
    - exploration au moyen de méthodes et algorithmes fondés sur les statistiques, l'informatique ou l'intelligence artificielle.



- L'analyse de la variance permet d'étudier le comportement d'une variable quantitative en fonction d'une ou de plusieurs variables explicatives.
- L'analyse factorielle discriminante cherche à identifier des groupes d'individus en représentant les données selon des combinaisons linéaires des variables prédictives (risque d'infarctus selon l'alimentation, le tabac, les antécédents familiaux...).
  - Le premier axe factoriel maximise la variance entre les groupes et la part de variance apportée par chacun des axes permet d'en apprécier l'importance.

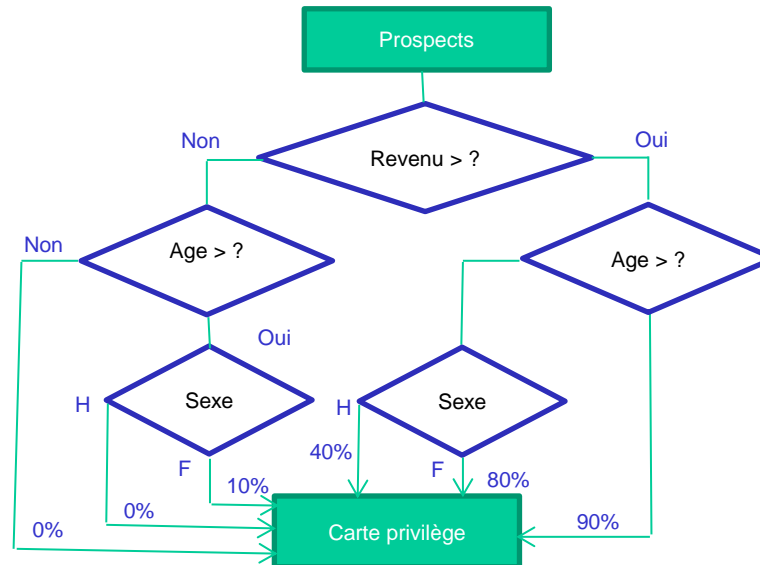


- ❑ L'analyse en composantes principales (ACP) transforme des variables corrélées en nouvelles variables indépendantes, appelées composantes principales.
- ❑ En nombre réduit, celles-ci expliquent presque aussi bien les données que l'ensemble des variables.
  - L'ACP détermine, par exemple, les deux axes d'inertie maximale qui expliquent au mieux la dispersion des pixels dans une image.

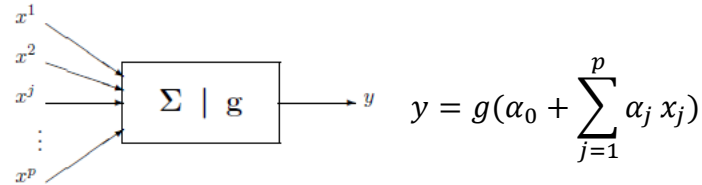


- ❑ La discrimination consiste à discerner et distinguer les éléments les uns des autres afin de pouvoir les séparer ou les regrouper.
  - La segmentation ou partitionnement (data clustering) cherche à découvrir des groupes et des structures similaires au sein des données.
  - La classification consiste à ranger des données nouvelles dans les différents segments ou clusters identifiés.
- ❑ Menée à partir de variables qualitatives ou quantitatives, la discrimination peut permettre d'établir un pronostic (ex : prévoir le comportement d'un individu).
- ❑ Menée à partir de variables quantitatives, la régression consiste à établir un modèle de données afin d'en prédire les valeurs futures.
- ❑ Les méthodes les plus utilisées issues de l'intelligence artificielle sont les arbres de décision, les réseaux de neurones et les réseaux bayésiens.

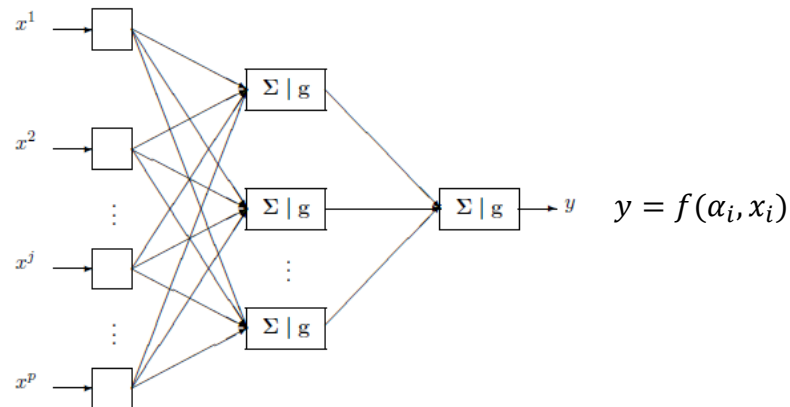




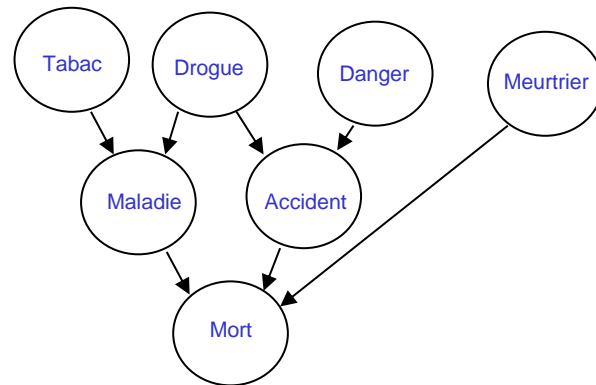
- ❑ L'arbre de décision représente un processus de décision.
- ❑ Les feuilles de l'arbre correspondent aux décisions possibles qui sont atteintes en fonction des choix pris à chacun des nœuds.
- ❑ L'arbre de décision permet de répartir une population d'individus en groupes homogènes selon un ensemble de variables discriminantes.
- ❑ Il présente l'avantage d'être lisible, rapide d'exécution et bien adapté à l'apprentissage automatique.



- ❑ Les réseaux de neurones sont inspirés du fonctionnement des neurones biologiques.
- ❑ Un neurone artificiel est doté d'une fonction de transfert qui transforme des entrées en sortie selon des règles précises, telles que la sommation des entrées puis la comparaison à un seuil.
- ❑ Chaque entrée est modulée par un coefficient synaptique et différents neurones peuvent être associés en réseaux selon des topologies diverses.
- ❑ Les réseaux de neurones permettent d'opérer très rapidement des classifications qui peuvent être améliorées par des règles d'apprentissage portant sur la valeur des seuils et des coefficients synaptiques.



- ❑ Les réseaux de neurones ont été améliorés avec l'invention du perceptron multicouche capable de traiter des phénomènes non linéaires.
- ❑ Un réseau reste cependant une boîte noire qui donne une réponse à une configuration d'entrées sans fournir une justification facile à interpréter.
- ❑ Cette réponse exploite toute la connaissance acquise jusqu'alors par apprentissage mais n'est pas toujours satisfaisante dans les cas de configurations inconnues (problème de la voiture autonome).

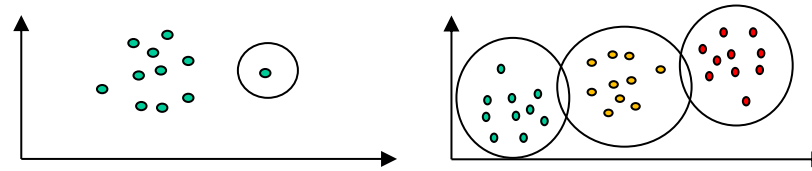


- ❑ Un réseau bayésien est un modèle de représentation graphique des dépendances entre des variables aléatoires de type booléen (vrai, faux), discret, voire continu.
- ❑ Les variables sont représentées par des cercles et les liens de dépendance par des flèches ou arcs orientés.
- ❑ Le sens des flèches représente un lien de parenté (parent → enfant) généralement causal (cause → effet).
- ❑ A partir d'une observation, le réseau peut calculer la probabilité de chacun des événements du graphe (risque de mort d'un fumeur en présence d'un meurtrier, probabilité d'assassinat en présence d'un mort...).

- Fondé sur l'inférence bayésienne, l'apprentissage automatique permet d'estimer les paramètres du réseau, voire la structure même du graphe à partir d'observations successives.
  - La fréquence d'apparition des événements sous certaines conditions permet d'estimer les probabilités conditionnelles et leur connaissance a priori.
  - L'analyse des corrélations entre les variables permet d'ajouter ou de supprimer des arêtes et détecter divers types de structure (graphe non causal).
- Les réseaux bayésiens sont notamment utilisés pour le diagnostic médical ou industriel.

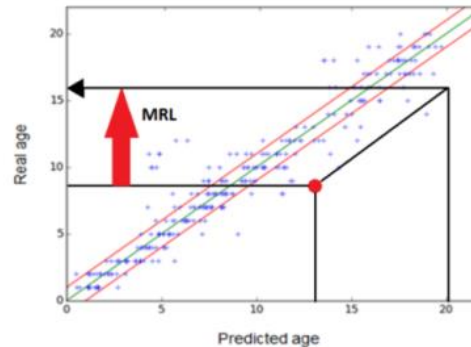
- L'apprentissage automatique (machine learning) donne aux ordinateurs la capacité de résoudre des tâches pour lesquelles ils ne sont pas explicitement programmés :
  - la mémorisation concerne l'assimilation de nombreux exemples,
  - la généralisation est la capacité à traiter des exemples distincts de ceux qui ont été appris.
- Un modèle d'apprentissage est élaboré en conception à partir des données disponibles puis est enrichi lors de la vie opérationnelle.
- L'apprentissage est supervisé s'il s'effectue sous le contrôle d'un expert (qui connaît la réponse attendue dans chacune des situations) ou non supervisé s'il est réalisé de manière autonome.
- L'apprentissage est sujet à des biais qui dépendent de la pertinence des modèles sous-jacents et de la qualité des données utilisées (un détecteur de panneaux de signalisation entraîné dans un pays peut avoir de mauvais résultats ailleurs).

- Le diagnostic est le raisonnement menant à l'identification et la caractérisation d'un problème (pannes et niveaux de dégradation) et éventuellement à la compréhension de sa cause à partir d'observations (symptômes).
- Le suivi de l'état fonctionnement était couvert jusqu'alors par :
  - la surveillance de données observables,
  - le dépassement de seuil par certaines données,
  - la comparaison de données habituellement corrélées (monitoring),
  - le suivi régulier des données par des spécialistes du domaine.
- Les techniques automatisées de fouille de données permettent :
  - d'identifier des changements ou évolutions lentes dans des séries d'observations chronologiques (signaux faibles),
  - de comparer des signatures caractéristiques d'un état de dégradation.

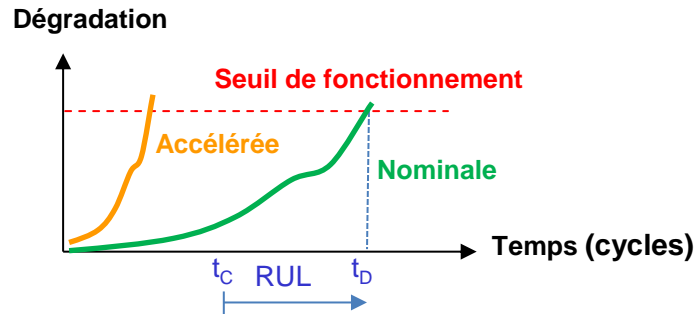


- ❑ Les données sont comparées entre des fenêtres temporelles successives (pouvant se chevaucher) ou entre des zones particulières d'intérêt (phases ou modes de fonctionnement).
- ❑ Une signature caractéristique des observations est élaborée par une méthode de fouille de données (indicateur statistique, calcul symbolique, réseau de neurones, etc.).
- ❑ Une distance entre les diverses observations est alors établie à partir des signatures afin de pouvoir identifier des comportements singuliers ou caractériser des équipements plus ou moins âgés ou dégradés (clustering).





- ❑ Le pronostic est le processus d'évaluation de la dégradation future d'un matériel et de sa durée de vie restante avant défaillance.
- ❑ Le pronostic peut résulter d'un modèle physique (model base), des observations (data driven), du retour d'expérience (REX) ou d'une approche hybride.
- ❑ Les modèles construits à partir de données observées peuvent se fonder sur des techniques automatisées de fouille de données faisant l'objet d'un apprentissage de signatures caractéristiques de dégradation.



- ❑ Les modèles peuvent également se fonder sur des lois de dégradation, telles que des processus de Lévy, ajustées à partir des données observées.
- ❑ Les lois de dégradation peuvent être rendues non stationnaires par changement de variable, et se coupler à des modèles d'accélération (Arrhenius...) pour tenir compte des conditions d'utilisation et d'environnement.
- ❑ Le choix d'un seuil de dégradation acceptable permet alors d'obtenir un modèle de fiabilité en ligne.
- ❑ Les modèles peuvent être élaborés en opération mais aussi en fin de conception au cours des essais accélérés de durabilité ou de fiabilité.

- ❑ Les techniques de fouille de données et d'apprentissage issues de l'intelligence artificielle permettent d'améliorer significativement le diagnostic en sachant détecter des signaux faibles ou des signatures caractéristiques de l'état de dégradation.
- ❑ Son apport est plus limité en ce qui concerne le pronostic qui requiert un apprentissage de tous les comportements opérationnels, dans toutes les conditions d'utilisation et d'environnement.
- ❑ La maintenance prédictive ne peut donc se suffire de l'intelligence artificielle que pour des produits peu sophistiqués fonctionnant dans des conditions stables ou des systèmes complexes dotés d'un long retour d'expérience opérationnel dans un marché de masse.



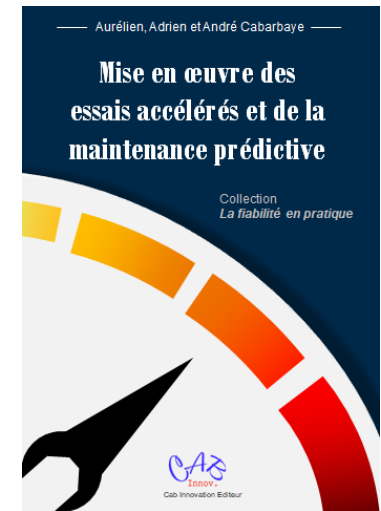
 Institut pour la Maîtrise des Risques  
 Sûreté de Fonctionnement - Management - Cindyniques

INSTITUT | PRODUITS | LES ENTRETIENS DU RISQUE | LE CONGRÈS LAMBDA MU | ACADÉMIE IMdR | CONTACT

**Prochain(s) événement(s) - Manifestation IMdR**

**3 juin 2022**

L'IMdR invite un auteur : **André CABARBAYE** « Mise en œuvre des essais accélérés et de la maintenance prédictive »



## RÉSUMÉ :

La maintenance prédictive, ou prévisionnelle, constitue le Graal des opérateurs de systèmes car elle améliore la disponibilité et la sécurité, réduit les coûts d'exploitation et diminue la consommation de ressources et la production de déchets. Mais sa quête est difficile et ne peut se fonder sur l'intelligence artificielle que pour des produits simples ou dotés d'un long retour d'expérience dans un marché de masse. Elle implique l'emploi d'un modèle prédictif capable de décrire les comportements dans diverses conditions de dégradation, d'environnement et d'utilisation (température, vibration, humidité, sollicitation...) afin d'élaborer un pronostic à partir de l'état courant. Aussi, est-il paru opportun de rassembler dans un même ouvrage la maintenance prédictive et les essais accélérés qui utilisent un tel modèle pour démontrer la capacité des produits à réaliser leur mission. Ce livre vulgarise leurs fondements théoriques, dont ceux de l'intelligence artificielle, des modèles d'accélération et des processus de dégradation, stationnaires ou non. Il propose un guide d'application pratique fondé sur l'exploitation de l'information disponible, plus que sur l'ajout de divers capteurs, et complète différents ouvrages des mêmes auteurs, dont « La fiabilité aujourd'hui », une encyclopédie de la Sûreté de Fonctionnement et de l'optimisation des systèmes.

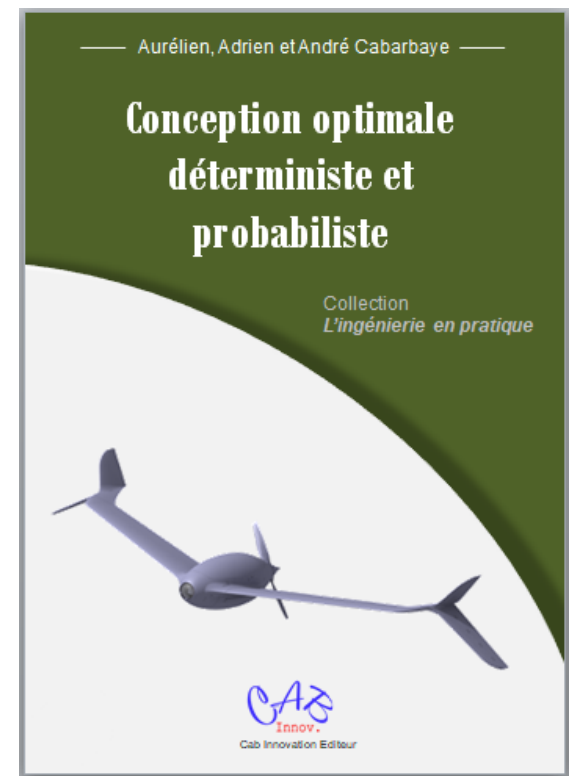
Cab Innovation éditeur



ISBN : 979-10-97287-12-2



ISBN : 979-10-97287-13-9



A paraître

*Boutique en ligne : <https://bit.ly/edition-cabinnovation>*