

# THÈSE DE DOCTORAT DE

L'UNIVERSITÉ DE NANTES

ÉCOLE DOCTORALE N° 601  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Informatique*

Par

**Mathilde MONVOISIN**

**Modèles graphiques probabilistes appliqués aux procédés de fabrication**

Thèse présentée et soutenue à Nantes, le 16/12/2022

Unité de recherche : Laboratoire des Sciences du Numérique de Nantes

## Rapporteurs avant soutenance :

Philippe WEBER Professeur à l'Université de Lorraine  
Karim TABIA Maître de conférence HDR à l'Université d'Artois

## Composition du Jury :

*Attention, en cas d'absence d'un des membres du Jury le jour de la soutenance, la composition du jury doit être revue pour s'assurer qu'elle est conforme et devra être répercutée sur la couverture de thèse*

Président :	Prénom NOM	Fonction et établissement d'exercice (à préciser après la soutenance)
	Philippe WEBER	Professeur à l'Université de Lorraine
	Karim TABIA	Maître de conférence HDR à l'Université d'Artois
	Véronique DELCROIX	Maître de conférence à l'Université Polytechnique Hauts-de-France
	Bruno CASTANIER	Professeur à l'Université d'Angers
Dir. de thèse :	Philippe LERAY	Professeur à Nantes Université
Co-enc. de thèse :	Mathieu RITOU	Maître de conférence HDR à Nantes Université

## Invité :

Adolfo SUAREZ ROOS Expert R&T IRT Jules Verne

---

**Titre :** Modèles graphiques probabilistes appliqués aux procédés de fabrication

**Mot clés :** Apprentissage, réseaux bayésiens, co-training, diagnostic, industrie 4.0

**Résumé :** La fabrication intelligente est un domaine de recherche prometteur pour l'amélioration de la productivité et de la compétitivité dans l'industrie, par l'exploitation des données numériques obtenues lors de procédés de fabrication, tel que l'usinage à grande vitesse. Les réseaux bayésiens ont fait leurs preuves en matière de classification et de diagnostic, et ils ont notamment l'intérêt d'être grandement interprétables. Cette thèse présente une architecture générique de réseaux bayésiens pour le diagnostic à partir de capteurs, incluant un mécanisme de sélection de variables basé sur l'information mutuelle. Le co-training est un champ émergent des algorithmes d'apprentissage à partir de don-

nées, et l'exploration de cette famille d'algorithmes est jusqu'à présent essentiellement limitée à un apprentissage supervisé ou semi-supervisé. Ce manuscrit propose plusieurs stratégies de co-training non-supervisées utilisables par tout modèle probabiliste, et détaille leur utilisation sur plusieurs jeux de données. L'ensemble des contributions théoriques est mis à profit dans un cas d'usage sur l'usinage à grande vitesse, dans lequel deux réseaux bayésiens avec la structure générique proposée permettent d'exploiter les données de capteurs d'une électrobroche en conditions réelles d'utilisation, et dont les paramètres sont appris grâce aux stratégies de co-training non-supervisées.

---

**Title:** Probabilistic graphical models applied to manufacturing processes

**Keywords:** Learning, Bayesian networks, co-training, diagnosis, industry 4.0

**Abstract:** Smart manufacturing is a promising area of research for improving productivity and competitiveness in industry, by exploiting digital data obtained during manufacturing processes, such as high-speed machining. Bayesian networks have proven their worth in classification and diagnosis, and they have the particular advantage of being highly interpretable. This thesis presents a generic Bayesian network architecture for sensor-based diagnosis, including a variable selection mechanism based on mutual information. Co-training is an emerging field of data-driven learning algorithms, and the exploration of this

family of algorithms is so far mostly limited to supervised or semi-supervised learning. This manuscript proposes several unsupervised co-training strategies that can be used by any probabilistic model, and details their use on several datasets. All the theoretical contributions are put to use in a use case on high speed machining, in which two Bayesian networks with the proposed generic structure are used to exploit sensor data of an electrospindle in real conditions of use, and whose parameters are learned thanks to the unsupervised co-training strategies.