

Prescription non supervisée des paramètres de fonctionnement d'un processus industriel avec un Auto-Encodeur Variationnel

L'Ollivier, Brendan, blollivier@fieldbox.ai, FieldBox.ai, Bordeaux (Orateur)
Tabti, Sonia, stabti@fieldbox.ai, FieldBox.ai, Bordeaux

Thème 4 : Analyse des données

Résumé : En industrie, la prescription des paramètres de fonctionnement d'un processus de production présente plusieurs défis : grande dimension, non-linéarité des corrélations entre les variables et diversité des données. Les réseaux de neurones génératifs sont de bons candidats pour les relever, en particulier avec les Auto-Encodeurs Variationnels (VAE). Dans cette étude on propose d'exploiter la capacité des VAE à représenter des distributions complexes en faible dimension. Ceci permet d'appliquer une méthode de recherche dans un graphe autrement non tractable, qui isole une zone de fonctionnement optimale au sein des données historiques de production.

Mots clés : Industrie, Prescription, Auto-Encodeur Variationnel, Réduction de dimensions

1. Introduction

L'intelligence artificielle joue un rôle prépondérant dans l'industrie 4.0 en tant qu'outil d'aide à la décision dans l'usine. Un état de l'art sur l'optimisation de processus industriels par le machine learning [1] décrit l'enjeu de développer des algorithmes performants sur des données de grande dimension et adaptés aux non-linéarités des corrélations entre paramètres d'entrée et de sortie, tout en restant assez génériques pour être utilisés dans des contextes industriels variés. Ces facteurs influencent positivement l'adoption du machine learning dans l'industrie. Cette étude traite un problème d'optimisation qui consiste à prescrire des paramètres de production et propose une solution générique capable, par construction, de s'affranchir d'un des principaux facteurs de complexité et de spécificité: la dimensionnalité de l'espace des paramètres. La méthode est testée sur un jeu de données réelles issu d'un processus d'**extrusion de film plastique**, sélectionné pour son grand nombre de variables (468) et d'échantillons historiques (225 000).

2. Méthodologie

L'algorithme de prescription prend en entrée un historique des paramètres de production du processus étudié, aussi bien actionnables que non-actionnables, et les valeurs associées de la variable cible à optimiser. L'objectif est d'isoler, au sein de l'espace des points de fonctionnement historiques, un sous-ensemble continu (au sens de la densité locale) où les valeurs moyennes locales de la variable cible respectent un seuil fixé. L'algorithme se divise en deux étapes:

- Entraînement d'un Auto-Encodeur Variationnel (VAE) [2] : il s'agit d'un réseau de neurones bayésien performant pour l'apprentissage non supervisé de distributions complexes. Il calcule une représentation à faible dimension, appelée espace latent, des données de production, ouvrant ainsi la voie à des techniques d'optimisation autrement non tractables.

- Maillage orthogonal de l'espace latent : cette étape permet le moyennage local de la variable cible ainsi que l'évaluation de la densité locale à partir des points du jeu de données historiques. Les mailles adjacentes sont alors liées les unes aux autres avec un graphe *radius-neighbor*, tout en ignorant celles ne respectant pas des seuils d'optimalité sur la variable cible et de densité locale. Les zones connectées de l'espace latent sont identifiées avec l'algorithme *Breadth-First Search* [3]. On sélectionne finalement la zone prescrite comme étant celle maximisant un critère de population (si l'on privilégie la représentativité historique), ou de volume (si l'on privilégie l'étendue de la zone en question).

3. Originalité / perspectives

Cette approche s'affranchit d'un paradigme de l'optimisation industrielle qui est la modélisation supervisée du lien entre les paramètres de production et la variable cible, limitant ainsi les erreurs de prédiction tout en maintenant l'indépendance des étapes de modélisation du processus et de prescription. L'espace latent peut alors être utilisé indifféremment avec différents objectifs d'optimisation (variables cibles, seuils, information a priori). De plus, la prescription d'une zone de fonctionnement au lieu d'un unique point, offre une garantie statistique de sa qualité et de sa stabilité.

En outre, cette méthode offre des perspectives d'application à la surveillance grâce à des intervalles de tolérance calculés à partir des frontières de la zone prescrite. L'espace latent peut aussi être exploité pour des tâches de d'apprentissage sur des variétés, notamment l'interpolation d'un chemin, depuis la zone de fonctionnement actuelle jusqu'à la zone prescrite, qui assure la viabilité des états intermédiaires du processus.

Références

[1] Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., & Wrobel, S. (2019). A review of machine learning for the optimization of production processes. *The International Journal of Advanced Manufacturing Technology*, 104(5-8), 1889-1902.

[2] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

[3] Bundy, A., & Wallen, L. (1984). Breadth-first search. In *Catalogue of artificial intelligence tools* (pp. 13-13). Springer,

Berlin, Heidelberg.