

(Doctorat : D4)
RESUME DE THESE¹

Nom et Prénom du candidat : Quasdane Mohamed **Année de la 1^{ère} Inscription :** 2020/ 2021

Formation Doctorale : Recherche et Développement en Sciences & Ingénierie

Etablissement de domiciliation : ENSAM/Meknès

Centre d'Etudes Doctorales : Sciences et techniques et sciences médicales

Titre de la thèse	Regularization and Hyperparameter Optimization for Deep Neural Networks: Sparsification and Pruning
Discipline/ Spécialité	Science de l'ingénieur / Mathématiques Appliquées et Intelligence Artificielle
Nom et Prénom du Directeur de thèse	Tawfik MASROUR
Structure de Recherche/Etablissement d'Attache	Artificial Intelligence, Data Technologies and Modeling / ENSAM-UMI
Nom et Prénom du responsable de la Structure de Recherche	EL OSSMANI Mustapha
Nom du Co-encadrant de thèse	Hassan RAMCHOUN
Structure de Recherche/Etablissement d'Attache	Artificial Intelligence, Data Technologies and Modeling / ENSAM-UMI

Résumé : (150 mots)

L'efficacité des réseaux de neurones profonds repose souvent sur des architectures profondes et surparamétrées, ce qui engendre plusieurs défis majeurs : une tendance marquée au surapprentissage, des besoins importants en ressources de calcul et de mémoire, ainsi qu'une difficulté d'utilisation sur des appareils aux ressources limitées. Cette thèse aborde ces problématiques en explorant des méthodes de régularisation parcimonieuse et d'élagage, dans le but d'optimiser les architectures de réseaux de neurones et d'améliorer leurs performances. Ce travail est motivé par la nécessité de développer des techniques performantes expérimentalement et fondées théoriquement, afin de concevoir automatiquement des architectures optimisées.

Dans ce but, nous proposons tout d'abord une revue approfondie accompagnée d'une analyse mathématique des fondements de l'apprentissage profond, ainsi que des méthodes de régularisation et des stratégies d'élagage existantes. En se basant sur ces fondements, nous présentons deux contributions originales dans le domaine de la régularisation et l'optimisation des hyperparamètres. Premièrement, pour les perceptrons multicouches, nous introduisons une méthode d'élagage fondée sur des masques lissés apprenables, permettant d'approximer de manière continue des décisions binaires concernant l'importance des poids. La méthode intègre de nouveaux termes de régularisation spécifiquement conçus pour contrôler l'importance individuelle des poids et des neurones, permettant ainsi une réduction efficace de l'architecture tout en préservant les performances du réseau. Deuxièmement, pour les réseaux de neurones convolutifs, nous développons une nouvelle méthode de régularisation lisse et parcimonieuse, visant à imposer une parcimonie structurée au niveau des groupes. Cette approche combine les avantages de deux régularisations parcimonieuses, L_0 et $L_{1/2}$, tout en surmontant leur non différentiabilité et leur complexité combinatoire grâce à des approximations lisses. Une analyse approfondie ainsi que des évaluations empiriques menées sur des bases de données de référence, démontrent l'efficacité des méthodes proposées.

Cette thèse apporte des contributions importantes à l'évolution et à la compréhension des stratégies de régularisation et d'élagage, considérées comme des cadres essentiels pour l'optimisation des architectures de réseaux de neurones.

Mots clés : Apprentissage profond, surapprentissage, surparamétrisation, régularisation parcimonieuse, élagage, optimisation d'hyperparamètres, approximation lisse, masques lissés apprenables, régularisation L_0 , régularisation $L_{1/2}$.

¹ Le présent résumé sera publié conformément à l'article 31 des NSPCD- 2023.