الجمهورية الشعبية الديمقراطية الجزائرية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي و البحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
المدرسة العليا للإعلام الآلي 08٠ ماي 5491٠ بسيدي بلعباس
École Supérieure en Informatique
-08 Mai 1945- Sidi Bel Abbès

THESIS

To obtain the diploma of **Master**
Field: **Computer Science**
Specialty: **Systèmes d'Information et Web (SIW)**

Theme

Optimizing Federated Learning Advancements in Model Aggregation and Client Selection

---

Presented by:
Amira MEKKI

Submission Date: **Sept, 2024**
In front of the jury composed of:

Mr. BENDAOUD Fayssal	President
Mr. OUCHANI Samir	Supervisor
Mr. BENSLIMANE Sidi Mohammed	Co-Supervisor
Mr. MALKI Abdelhamid	Examiner

Academic Year : 2023/2024

Abstract

In recent years, Federated Learning (FL) has emerged as a powerful approach to addressing critical challenges in various domains, such as industry, healthcare, and communications. Due to its decentralized nature, FL offers solutions to concerns like data privacy and communication efficiency, which are paramount in these fields. However, traditional FL methods often face challenges such as non-IID data distributions and the varied computational capabilities of participating devices, which can hinder model performance and convergence. To address these limitations, this master's thesis delves into the integration of Reinforcement Learning (RL) into FL, aiming to optimize client selection processes. By prioritizing clients with high-quality data and strong computational resources, RL-based client selection strategies can significantly enhance the overall performance and efficiency of FL systems. This thesis provides a comprehensive background on the relevant ML and FL techniques, followed by an in-depth analysis of state-of-the-art advancements in model aggregation and client selection. The research contributions are discussed, and future directions in FL research are outlined.

Keywords— Federated Learning, Machine Learning, Reinforcement Learning, Model Aggregation, Client Selection.

ملخص

في السنوات الأخيرة، أصبح التعلم الفيدرالي نهجًا قويًا لمواجهة التحديات الحرجة في مجالات مختلفة مثل الصناعة والرعاية الصحية والاتصالات. وبفضل طبيعته اللامركزية، يقدم التعلم الفيدرالي حلولًا لمخاوف مثل خصوصية البيانات وكفاءة الاتصالات، وهي أمور ذات أهمية كبيرة في هذه المجالات. ومع ذلك، غالبًا ما تواجه الطرق التقليدية للتعلم الفيدرالي تحديات تشمل التوزيعات غير المتماثلة للبيانات وقدرات الحوسبة المتفاوتة للأجهزة المشاركة، مما قد يؤثر سلبًا على أداء النموذج ومعدل التقارب. للتغلب على هذه القيود، يستكشف هذا البحث التكامل بين التعلم المعزز والتعلم الفيدرالي بهدف تحسين عمليات اختيار العملاء. من خلال إعطاء الأولوية للعملاء الذين يمتلكون بيانات عالية الجودة وموارد حوسبة قوية، يمكن لاستراتيجيات اختيار العملاء القائمة على التعلم المعزز أن تحسن بشكل كبير من الأداء العام وكفاءة أنظمة التعلم الفيدرالي. يقدم هذا البحث إطارًا شاملاً حول تقنيات التعلم الآلي والتعلم الفيدرالي، متبوعًا بتحليل عميق لأحدث التطورات في مجال تجميع النماذج واختيار العملاء. كما يتم مناقشة إسهامات هذا البحث وتقديم توجهات البحث المستقبلية في مجال التعلم الفيدرالي.

التعلم الفيدرالي، التعلم الآلي، التعلم المعزز، تجميع النماذج، اختيار العملاء

Résumé

Ces dernières années, l'apprentissage fédéré s'est imposé comme une approche puissante pour relever des défis cruciaux dans divers domaines, tels que l'industrie, la santé, et les communications. Grâce à sa nature décentralisée, le FL offre des solutions à des préoccupations telles que la confidentialité des données et l'efficacité des communications, qui sont primordiales dans ces domaines. Cependant, les méthodes traditionnelles de FL sont souvent confrontées à des défis, notamment les distributions de données non-IID et les capacités informatiques variées des dispositifs participants, ce qui peut nuire à la performance du modèle et à la convergence. Pour surmonter ces limitations, ce mémoire de master explore l'intégration de l'apprentissage par renforcement dans le FL, dans le but d'optimiser les processus de sélection des clients. En priorisant les clients disposant de données de haute qualité et de ressources informatiques solides, les stratégies de sélection des clients basées sur le RL peuvent améliorer de manière significative la performance globale et l'efficacité des systèmes de FL. Ce mémoire fournit un cadre complet sur les techniques d'apprentissage automatique et de FL, suivi d'une analyse approfondie des avancées les plus récentes en matière d'agrégation de modèles et de sélection des clients. Les contributions de cette recherche sont discutées, et les futures orientations de la recherche en FL sont également présentées.

Mots-clés— Apprentissage fédéré, Apprentissage automatique, Apprentissage par renforcement, Agrégation de modèles, Sélection des clients.