

Quantum Computing for Hybrid Machine Learning: Current Limitations and Future Perspectives

Sabri Mdaini

October 2, 2025

1 Introduction

Quantum computing (QC) offre un potentiel prometteur pour le machine learning (ML), notamment à travers les architectures hybrides combinant des parties classiques et des circuits quantiques paramétrables (VQC, Quantvolutional Filter, etc.). Cependant, l'utilisation actuelle des QC pour des applications ML complètes reste limitée. Ce rapport décrit l'état actuel de l'usage des QC pour le ML hybride, les stratégies pratiques, et les perspectives à court terme.

2 Validation de l'idée avec un prototype

Lorsqu'un modèle hybride est entraîné sur un GPU/CPU avec la partie quantique simulée, il est possible de réaliser un Proof of Concept (POC) en testant l'inférence sur un vrai QC. Cette étape permet de :

- Valider le concept : montrer que la partie quantique améliore l'extraction de features.
- Tester le modèle dans un environnement quantique réel, y compris l'effet du bruit et des erreurs de mesure.
- Préparer le workflow pour une future intégration dans un environnement de production lorsque les QC seront plus performants.

3 Limitations actuelles des QC pour le ML

Les limitations actuelles des QC empêchent leur utilisation directe en production pour le ML :

1. **Nombre limité de qubits** : Les QCs actuels disposent de quelques dizaines à quelques centaines de qubits. Les images ou vecteurs d'entrée ML sont souvent beaucoup plus grands, rendant l'utilisation directe impraticable.
2. **Bruit et décohérence** : Les erreurs de portes et de mesure introduisent de l'incertitude dans les résultats, affectant la fiabilité de l'inférence.

3. **Temps d'exécution long** : L'envoi des circuits au cloud, leur exécution sur le QC et le retour des résultats peuvent prendre plusieurs minutes par batch, ce qui est incompatible avec un usage en production.
4. **Coût** : L'accès aux QC via le cloud peut être limité et potentiellement payant pour un usage intensif.

4 Stratégie pratique pour le court terme

Pour exploiter les QC de manière réaliste aujourd'hui, la stratégie suivante est recommandée :

1. **Entraînement sur GPU/CPU** : La partie quantique du modèle est simulée, permettant un entraînement rapide et fiable.
2. **Test / inference sur vrai QC** : Une fois le modèle entraîné, envoyer uniquement la partie quantique sur le QC pour vérifier le comportement dans un environnement réel.
3. **Préparation pour la production future** : Concevoir l'architecture pour que la partie quantique puisse être intégrée directement lorsque les QC auront plus de qubits, moins de bruit et des temps d'exécution rapides.

5 Conclusion et recommandations

- Actuellement, l'usage des QC pour le ML est principalement limité aux POC et aux démonstrations.
- La production directe avec des QC réels n'est pas encore réaliste à cause des limitations techniques.
- Une approche hybride réaliste consiste à entraîner sur GPU/CPU et tester la partie quantique sur un QC pour validation.
- La stratégie consiste à préparer le modèle dès maintenant pour qu'il soit compatible avec les QC plus performants du futur.