

Sintesi relazione finale

Candidato: Bolognini Gaia Stella, 838719
Titolo prova finale: Classificatore con reti neurali quanto-classiche
Relatore: Prof. Andrea Giachero
Correlatori: Prof. Alberto Leporati, Prof. Claudio Ferretti
Data della seduta della prova finale: Sessione del 21/22 Luglio
Corso di laurea: Corso di Laurea Triennale in Fisica
Recapito telefonico: +393332944502

Il *Machine Learning*, ovvero l'insieme di meccanismi che permettono a una macchina di migliorare le proprie capacità e prestazioni nel tempo agendo su dati, ha permesso di ottenere importanti risultati e di risolvere notevoli problemi, non solo in ambiti specifici ma anche nella nostra vita quotidiana, ne è un caso il riconoscimento vocale degli smartphone o le pubblicità traccianti. Le *reti neurali* sono state sviluppate costruendo macchine con architettura ispirata al sistema nervoso umano per compiti di Machine Learning. All'inizio del XXI secolo, una maggior disponibilità di dati e maggior potenza di calcolo hanno permesso di evidenziare l'applicabilità di questi modelli e del *Deep Learning*; l'apprendimento avviene tramite modelli matematici e algoritmi di calcolo statistico in maniera automatica. Ci sono alcuni particolari ambiti, come nel caso del riconoscimento di immagini, dove l'intelligenza artificiale ha raggiunto o ha anche superato le capacità umane. *Teoricamente* le reti neurali sono capaci di simulare ogni algoritmo di apprendimento se possiedono un sufficiente numero di dati, soddisfacendo in alcuni casi la condizione di *Turing Completeness*. Può però sorgere un problema legato proprio al set di dati a cui si fa riferimento che, se di grandi dimensioni, può fortemente incidere sul tempo impiegato per l'apprendimento. E' in questo quadro che si inserisce la possibilità dell'utilizzo dei *computer quantistici*.

I computer quantistici fanno uso delle proprietà della meccanica quantistica, come *l'interferenza* tra funzioni d'onda, *l'entanglement* e la *sovrapposizione* degli stati, per eseguire calcoli e computazione. Come i computer classici codificano l'informazione e i dati in bit (0 o 1), i computer quantistici utilizzano i bit quantistici, *qubits*, che si possono trovare in sovrapposizione di stati, assumendo un *continuo* di valori. Grazie a questa proprietà, queste macchine potrebbero risolvere problemi in tempi minori rispetto ai computer classici: Google ha annunciato nell'ottobre del 2019 di aver raggiunto la *Quantum Supremacy*, ovvero aver risolto un problema con un notevole risparmio di tempo su un computer quantistico rispetto al più performante computer classico. In maniera analoga alle porte logiche che si trovano in computazione classica, diverse operazioni possono essere eseguite sui qubits: si parla in questo caso di *gates* quantistici.

Esistono diversi modelli di computer quantistici, per esempio *Quantum Turing Machines*, *Adiabatic quantum computing* e il modello a *Circuiti quantistici* e attualmente diverse aziende, tra cui Google, IBM, Intel, Rigetti e D-Wave, stanno sviluppando i loro componenti. La ricerca in questi settori si concentra sia sugli algoritmi che possono essere utilizzati su questo tipo di computer sia sulle sfide tecnologiche, entrando nella fisica della materia per sfruttarne le proprietà quantistiche. Diverse sono le tecnologie esplorate per capire quale possa essere la migliore per la costruzione dei componenti, per esempio utilizzando *superconduttori*, *ioni intrappolati* o ancora *qubit topologici*. Si stima che nel vicino futuro macchine note come NISQ (Noisy Intermediate Scale-Computers), ovvero soggette a rumore e limitate nel numero di qubits a disposizione, possano portare a vantaggi in diversi contesti.

La possibilità di utilizzare i computer quantistici al fine di ridurre in maniera significativa il tempo di computazione e le dimensioni del set di dati utili all'apprendimento della macchina vede la nascita del *Quantum Machine Learning*. La ricerca in questo ambito è ancora allo stato dell'arte ma sono diverse le idee sviluppate che potrebbero portare a benefici. Per esempio, alcune operazioni di ottimizzazione potrebbero essere eseguite più velocemente da processori

basati su *Quantum Annealing*, dove viene impiegata la fisica quantistica al fine di trovare stati a energia più bassa per un sistema. Nel breve termine gli algoritmi ibridi, ovvero funzionanti in parte su un computer classico e in parte su uno quantistico, sono i più promettenti. In questo contesto si inseriscono le reti neurali ibride, ovvero circuiti quantistici dipendenti da alcuni parametri da ottimizzare sui dati in ingresso per operazioni di classificazione. La forma più comune per questo tipo di modelli è basata su due componenti fondamentali: una funzione che abbia il compito di codificare i dati classici nel computer quantistico e una funzione variazionale con i parametri liberi da ottimizzare che abbia il compito di effettuare la classificazione vera e propria.

Nella tesi si studiano le caratteristiche di una particolare implementazione di Quantum Machine Learning utilizzando delle reti neurali ibride per il riconoscimento di immagini, digit MNIST, in un contesto di *Supervised Machine Learning* dove è nota la classe di appartenenza dei dati in ingresso. All'interno delle *layer* nascoste di una rete neurale viene inserita una layer costruita con un circuito quantistico dipendente da alcuni parametri. I parametri sono il risultato delle layer precedenti e vengono utilizzati come angoli di rotazione per i gate interni al circuito quantistico. Possono essere implementate diverse operazioni sui qubits del circuito, per esempio l'entanglement o inserendo gate dedicati, o possono essere aggiunti qubits. La misura sul circuito viene utilizzata per continuare la computazione classica ottenendo un output dalla rete neurale che possa essere confrontata numericamente con il risultato atteso, a partire dalla *loss function*, definita in maniera analoga al caso classico. I parametri inseriti nel circuito vengono poi ottimizzati insieme ai parametri delle layer classiche con l'utilizzo dell'algoritmo di *backpropagation*, dove il metodo di ottimizzazione viene implementato per essere compatibile con il circuito quantistico, attraverso il meccanismo di *parameter shift rule*. L'obiettivo è minimizzare la loss function e creare un modello che sia in grado di fare previsioni corrette sui dati in ingresso. I programmi sono stati implementati con le librerie *Qiskit* e *Pytorch* fornite per il linguaggio Python, che permettono di effettuare computazione rispettivamente nell'ambito del *Quantum Computing* e del *Machine Learning*. I vari programmi sono stati testati su simulatori e su computer quantistici, utilizzando la piattaforma online *IBM Quantum Experience*, che mette a disposizione via cloud i computer quantistici costruiti da IBM.

Sono stati dunque effettuati diversi esperimenti a partire dal modello sopra descritto, modificando le caratteristiche del circuito quantistico e il numero di qubit. Sono stati inoltre aggiunti errori gaussiani con varie deviazioni standard nel set di dati di validazione, per testare il modello, e nel set di dati di apprendimento, per rafforzarlo. In particolare, è stata implementata una rotazione in tre dimensioni, gate U3, per un singolo qubit che ha permesso di ottenere un'accuratezza sui risultati del 99.9% eseguendo il programma sia sul simulatore che sul computer quantistico reale. Inoltre, validando il modello con dati rumorosi, si è ottenuta un'accuratezza superiore al 99% fino a deviazioni standard di 0.7. E' stato poi implementato un circuito con più qubit, ciascuno soggetto a una rotazione in una sola dimensione, che ha dato dei risultati pari al 99% di accuratezza per set di dati non rumorosi. Nel caso di dati di validazione rumorosi, i risultati erano ancora accettabili se il set di dati di apprendimento era a sua volta rumoroso (deviazione standard pari a 0.2). In seguito sono stati testati un circuito a più qubit entangled, soggetti inoltre a una rotazione dipendente dai parametri di apprendimento, e un circuito ispirato ad un algoritmo ibrido di ottimizzazione, il QAOA che richiede a sua volta l'utilizzo dell'entanglement. In questo caso si è ottenuta un'accuratezza rispettivamente del 99% e del 98% sui devices quantistici.

La tesi è suddivisa in tre capitoli:

1. Nel primo capitolo viene riportato un approfondimento riguardo alla teoria sottostante al Machine Learning e ai Computer Quantistici, con particolare attenzione alle caratteristiche utilizzate per implementare il lavoro di tesi.
2. Nel secondo capitolo la prima parte è dedicata ad una visione di insieme riguardo al Quantum Machine Learning. Vengono poi evidenziate le caratteristiche tecniche dei programmi implementati nel progetto di tesi e vengono descritti i vari esperimenti che sono stati effettuati.
3. Il terzo capitolo è dedicato ad un'analisi critica sui risultati ottenuti, alle possibili future implementazioni del modello e alle conclusioni.