

# UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA

Riassunto dell'elaborato della prova finale

Scuola di Scienze  
Dipartimento di Fisica "Giuseppe Occhialini"  
Corso di Laurea Magistrale in Fisica  
Curriculum di Fisica delle Particelle e Fisica Applicata



Feature Extraction Neural Networks for Quantum  
Kernel classifiers: low-energy background rejection in  
Xenon-Doped LArTPC detectors

**Relatore:**

Dott. Andrea Giachero

**Correlatore:**

Dott. Daniele Guffanti

**Relatore esterno:**

Dott. Michele Grossi

**Candidato:**

Roberto Moretti

**Matricola:**

825617

**Recapito telefonico:**

+39 331 248 99 59

Sessione di Laurea di Ottobre. Anno Accademico 2021/2022

# Abstract

In the last decade, the development of Machine Learning (ML) and Deep Learning (DL), and their application in several branches of physics contributed significantly to the advancements of the latter. One of the most innovative proposals in this field is Quantum Machine Learning (QML), an interplay between Machine Learning and Quantum Computing (QC), aiming to achieve an edge over traditional ML through the paradigm of quantum information. The main goal of the present thesis was to develop and characterise QML models for solving the task of supervised classification, a recurring problem in particle physics. The supervised classification consists of tagging elements of a vector space according to the value of their components, called *features*. An example of a classification algorithm is the Support Vector Machine (SVM), which is based on mapping the features into a vector space of greater dimension than the initial one, in which to define a separation hyperplane which optimizes the distance between elements of different classes and the hyperplane itself. A crucial aspect of SVM that affects the model's performance is the choice of the feature map. The latter occurs implicitly, by defining a positive, semidefinite scalar function of the features, called *kernel*. The SVM's kernel determines its expressivity, which is the capability to model the non-linear relations between variables that are useful for classification. Introducing Quantum Computing gives access to a new class of kernels, called Quantum Kernels, that are intractable with classic techniques. The SVMs that classify data using Quantum Kernels are called Quantum SVMs (QSVMs) and are examples of algorithms that are executable on modern NISQ devices (Noise Intermediate Scale Quantum devices).

The present thesis explores the potentials of QSVM within the framework of the Deep Underground Neutrino Experiment (DUNE) for the analysis of the neutrinoless double beta decay ( $0\nu\beta\beta$ ) of the  $^{136}\text{Xe}$  isotope. The observation of this hypothetical decay would be the first proof of lepton number violation and would confirm the Majorana theory on the equivalence between neutrinos and anti-neutrinos. DUNE is a Long-Baseline experiment designed for the measurement of neutrino flavour oscillation, the search for proton decay and the detection of supernovae and solar neutrinos. The DUNE Far Detector facility (FD) will be made of four modules, each one consisting of a 17 kton Liquid Argon Time Projection Chambers (LArTPCs), a class of detectors capable of reconstructing the track of ionizing particles in addition to their energy deposition. The  $0\nu\beta\beta$  analysis proposal consists of employing a xenon-doped FD module, enriched in the  $^{136}$  isotope, and searching for the decay at the expected energy peak, i.e. the Q-value of the reaction  $Q_{\beta\beta}^{^{136}\text{Xe}} = 2.458 \text{ MeV}$ . To increase the sensitivity of the experiment, it is necessary to implement background mitigation strategies. The latter is dominated by  $\beta$  emissions from the  $^{42}\text{Ar}$  decay chain and neutron capture.

This thesis investigates the possibility of carrying out background mitigation by developing ML and QML classifiers, able to exploit the track information produced by the TPC. This goal is challenging and complicated by the sub-optimal tracking efficiency of the DUNE LArTPCs at the few-MeV energy scale, offering the opportunity to test quantum algorithms where the classical ones underperform. A necessary step to implement QSVM on limited-size quantum devices was to reduce the track information by going through a feature extraction process. This was performed by developing two Deep Learning models: a Convolutional Neural Network (CNN) and a Transformer, or Attention Neural Network (ATN). The first one achieves the task by making use of convolutional filters, while the second one is based on the innovative mechanism of self-attention. Both models were used as feature extractors for QSVM and as independent classifiers. Once the goal of feature extraction was reached, the QSVMs were trained with Quantum Kernels, varying the number of qubits and the corresponding quantum circuits. The Quantum Kernels were first chosen arbitrarily and later optimized with a meta-heuristic genetic algorithm. The data used for the training of CNN, ATN and QSVM consisted of  $0\nu\beta\beta$  and  $\beta$  background tracks, reconstructed by an ideal LArTPC simulated using the Geant4 Monte Carlo simulation toolkit. To complete the study, the CNN was also tested for mitigating radiological background in one of the DUNE planned FD modules, called Horizontal Drift.

This work provides interesting results from QML and DL points of view as well as from a physics perspective, assessing DUNE's potential for the search of the  $0\nu\beta\beta$  decay of the  $^{136}\text{Xe}$  isotope. QSVM classifiers have been implemented on IBM quantum hardware and have proven to be competitive, achieving classification accuracies equivalent to the ones of the best classic SVM of common use, based on a Gaussian Kernel: 66% and 64% with features extracted from the CNN and ATN, respectively. Further studies were made, varying parameters affecting the trace reconstruction, showing that ATN outperforms the CNN in case of high track resolution, while on the contrary, CNN prevails at low resolutions. The topological classification accuracy of  $0\nu\beta\beta$  events and the main radiological background in the DUNE FD Horizontal Drift detector settles at 63% with the use of the CNN. This result adds a benchmark for DUNE's potential in the search for the  $0\nu\beta\beta$  decay and on the topological background mitigation for other low-energy channels proposed for DUNE, such as solar neutrinos and WIMPs.

The thesis activities have been carried out during an International Mobility period with the Erasmus Traineeship programme, at the *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas* (CIEMAT) of Madrid (ES), in collaboration with the CIEMAT Neutrino Physics working group and the CERN Quantum Technology Initiative group (CERN QTI).

# Sommario

Nell'ultimo decennio, lo sviluppo ed il crescente utilizzo di algoritmi di Machine Learning (ML) e Deep Learning (DL) in numerose aree della fisica hanno dato forti contributi al loro avanzamento, tanto da costituirne oggi un elemento fondamentale. Tra le proposte più recenti in questo ambito si annovera il Quantum Machine Learning (QML), un insieme di strategie basate sull'utilizzo congiunto di Machine Learning e Quantum Computing (QC), con lo scopo di trarre vantaggio dal paradigma dell'informazione quantistica. La presente tesi pone come obiettivo principale lo sviluppo e la caratterizzazione di modelli di QML con applicazioni in fisica delle particelle, in particolare la classificazione supervisionata. Quest'ultima consiste nel categorizzare elementi di uno spazio vettoriale in base al valore delle loro componenti, dette *features*. Un esempio di algoritmo di classificazione è la Support Vector Machine (SVM), la quale si basa sul mappare le features di partenza in uno spazio vettoriale di dimensione maggiore di quello iniziale, in cui definire un iperpiano di separazione in grado di ottimizzare distanza che intercorre tra elementi di classi differenti e l'iperpiano stesso. Un aspetto cruciale della SVM, che ne determina la sua efficacia, è la scelta della funzione di mappatura, o *feature map*. Quest'ultima avviene in modo implicito, tramite la definizione di una funzione scalare semidefinita positiva delle features, detta *kernel*. Il kernel della SVM determina la sua espressività, ovvero la capacità di modellare relazioni non lineari tra le variabili ai fini della classificazione. In questo contesto, l'utilizzo di un processore quantistico permette di accedere in modo naturale ad una nuova classe di kernel, detti Quantum Kernels, intrattabili con tecniche classiche. Le SVM che classificano dati mediante Quantum Kernel sono denominate Quantum SVM (QSVM), e sono esempi di algoritmi implementabili su dispositivi quantistici attualmente esistenti, denominati NISQ devices (Noise Intermediate Scale Quantum devices).

Nel presente lavoro di tesi vengono esplorate le potenzialità della QSVM nell'ambito del Deep Underground Neutrino Experiment (DUNE), nell'analisi del decadimento doppio beta senza emissione di neutrini ( $0\nu\beta\beta$ ) dell'isotopo  $^{136}\text{Xe}$ . L'osservazione di questo ipotetico processo proverebbe per la prima volta la violazione del numero leptonico e confermerebbe la teoria di Majorana sull'equivalenza di neutrino ed anti-neutrino. DUNE è un esperimento di tipo Long-Baseline progettato per lo studio dell'oscillazione di sapore del neutrino, la ricerca del decadimento del protone e la rivelazione di neutrini di origine solare e da supernova. Il Far Detector (FD) di DUNE sarà costituito da quattro moduli Liquid Argon Time Projection Chambers (LArTPCs), ciascuno contenente 17 kton di argon liquido, in grado di ricostruire la traccia di particelle ionizzanti e misurarne l'energia. La proposta di analisi  $0\nu\beta\beta$  consiste nel drogare uno dei moduli FD con xeno arricchito nell'isotopo  $^{136}$  e ricercare il decadimento al picco atteso, pari al suo Q-valore  $Q_{\beta\beta}^{^{136}\text{Xe}} = 2.458 \text{ MeV}$ .

Per aumentare la sensibilità dell'esperimento è necessario attuare strategie di riduzione del fondo, dominato dall'emissione  $\beta$  proveniente dalla catena del  $^{42}\text{Ar}$  e da cattura neutronica. La presente tesi esamina la possibilità di effettuare tale mitigazione costruendo classificatori ML e QML basati sull'informazione prodotta dalla TPC. Questo obiettivo risulta ambizioso, complicato dalla limitata efficienza di tracciamento di elettroni da pochi MeV nelle LArTPC di DUNE, e costituisce un banco di prova per lo sviluppo di algoritmi quantistici dove la controparte classica risulta poco performante.

Un passaggio necessario per implementare la QSVM su dispositivi quantistici di dimensione limitata è la riduzione del contenuto informativo prodotto dalla TPC, tramite un processo di estrazione delle features. In questo elaborato sono stati utilizzati due modelli di Deep Learning: la Convolutional Neural Network (CNN) ed il Transformer, o Attention Neural Network (ATN). Il primo approccio è caratterizzato dall'uso di filtri convoluzionali, mentre il secondo si basa sull'innovativo meccanismo della *self-attention*. Entrambi i modelli sono stati impiegati sia come ausilio alla QSVM che come classificatori indipendenti. Una volta raggiunto l'obiettivo dell'estrazione di features, la QSVM è stata addestrata con diverse tipologie di Quantum Kernel, al variare del numero di qubit ed in funzione di diversi circuiti quantistici. I Quantum Kernel, dapprima selezionati con criteri arbitrari, sono poi stati ottimizzati tramite un algoritmo meta-euristico di tipo genetico. I dati utilizzati per addestrare i modelli CNN, ATN e QSVM corrispondono a tracce ricostruite da una LArTPC ideale, a partire da eventi  $0\nu\beta\beta$  e fondo  $\beta$  simulati mediante l'uso del *toolkit* di simulazione Monte Carlo Geant4. A completamento dello studio, la CNN è stata adoperata per la classificazione di eventi  $0\nu\beta\beta$  e fondo radiologico da una simulazione del FD Horizontal Drift, uno dei moduli pianificati per DUNE.

Il presente lavoro di tesi ha prodotto risultati di interesse nell'ambito del Quantum Machine Learning e nel Deep Learning, ed ha contribuito ad asserire le potenzialità dell'esperimento DUNE nella ricerca del  $0\nu\beta\beta$  dell'isotopo  $^{136}\text{Xe}$ . I classificatori quantistici QSVM sono stati implementati su hardware IBM e si sono dimostrati competitivi, ottenendo un'accuratezza di classificazione equivalente a quella del miglior SVM classico di uso comune, basato su kernel gaussiano: 66% e 64% con feature estratte rispettivamente da CNN e ATN. Ulteriori studi, condotti variando parametri che incidono sulla ricostruzione delle tracce, dimostrano che le performance della ATN in condizioni ad alta risoluzione superano quelle della CNN, mentre nel caso contrario si verifica un'inversione di tendenza. L'accuratezza nella classificazione topologica di eventi  $0\nu\beta\beta$  e dei principali fondi radiologici nel detector DUNE FD Horizontal Drift si assesta al 63% con l'uso della CNN. Tale risultato costituisce un parametro di riferimento sulle potenzialità di DUNE nella ricerca del decadimento  $0\nu\beta\beta$  e sulle capacità di mitigazione del fondo in altri canali di basse energie proposti per DUNE, come neutrini solari e la ricerca di WIMPs.

Le attività di tesi sono state svolte durante un periodo di Mobilità Internazionale tramite il programma Erasmus Traineeship, presso il *Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas* (CIEMAT) di Madrid (ES), in collaborazione con il Neutrino Physics working group di CIEMAT ed il gruppo di CERN Quantum Technology Initiative (CERN QTI).