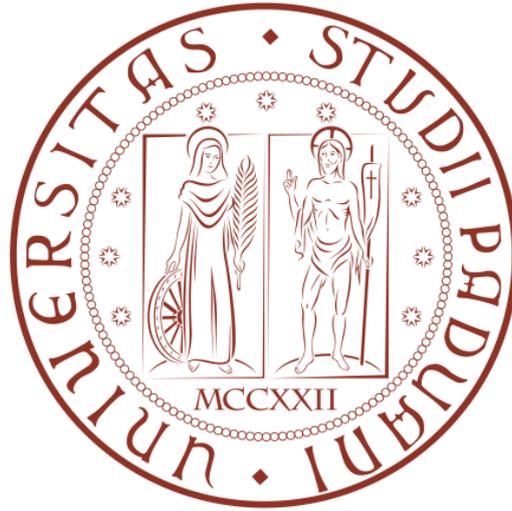


UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI FISICA ED ASTRONOMIA "Galileo
Galilei"



TESI DI LAUREA

DISCRIMINAZIONE DEL SEGNALE
DAL FONDO IN GERDA MEDIANTE L'
UTILIZZO DI UNA RETE NEURALE

Relatore: dott. ALBERTO GARFAGNINI

Laureando: LUCA MANNESCHI

ANNO ACCADEMICO 2013/2014

Indice

1	Introduzione	4
1.1	L' esperimento GERDA	4
1.2	Rivelatori al germanio	5
2	Catena d'acquisizione	7
2.1	GELATIO	8
3	Rete neurale e TMVA	9
3.1	La rete	9
3.2	Convergenza	10
4	Metodologia	13
4.1	Sorgente di Torio	13
4.2	Introduzione all' analisi	16
5	Analisi dei dati	17
5.1	Allenamento	17
5.2	Applicazione	20
6	Conclusioni	22

Capitolo 1

Introduzione

1.1 L' esperimento GERDA

L' apparato sperimentale GERDA [1] è situato ai laboratori Nazionali del Gran Sasso (LNGS) con l' intento di rivelare il doppio decadimento beta senza emissione di neutrini ($0\nu\beta\beta$) del ^{76}Ge . Il processo in questione è descritto dalla reazione :

$$(A, Z) \mapsto (A, Z + 2) + 2e^{-}$$

dove all' interno di un nucleo due neutroni decadono in due protoni e due elettroni. La scoperta di questo decadimento provverebbe la non conservazione del numero leptonic e avrebbe ripercussioni notevoli nella fisica particellare [2]. Il doppio decadimento beta con emissione di neutrini

$$(A, Z) \mapsto (A, Z + 2) + 2e^{-} + 2\nu_e$$

è un evento estremamente raro ed è osservato sperimentalmente con un tempo di dimezzamento di $10^{19} - 10^{20}$ anni a seconda del nucleo in questione. Per questo rivelare il decadimento è estremamente difficile.

In GERDA l' apparato sperimentale [1] è stato ideato e costruito per diminuire il più possibile il fondo, dovuto principalmente alla radioattività ambientale. Il cuore della strumentazione è costituito da diodi al germanio naturale, arricchiti con il 7.8% di ^{76}Ge , immersi in un criostato composto di rame e riempito di argon liquido (LAr) dal volume di 63m^3 , circondato a sua volta da un serbatoio d'acqua. La funzionalità generale di quasi ogni elemento citato sopra è quella di ridurre il fondo ambientale, nello specifico il LAr serve da liquido di raffreddamento, il criostato composto di rame protegge il suo interno dai raggi gamma, e il recipiente d' acqua completa il tutto facendo da schermo per i neutroni. L' apparato è sostenuto inoltre da una superstruttura d' acciaio ed è situato nei laboratori sotterranei LNGS, sovrastati da 1400 metri di roccia che funge da schermo per i raggi cosmici.

1.2 Rivelatori al germanio

I rivelatori al germanio [3] sono prima di tutto dei semiconduttori, e come tali devono la propria funzionalità alla connessione p-n. Un semiconduttore ha la proprietà di avere la differenza energetica tra banda di valenza e di conduzione molto bassa (per il germanio è per es. di 0.67eV), e grazie a questa caratteristica può essere drogato opportunamente per produrre degli elettroni (caso n) o delle lacune (caso p) portatori maggioritari di carica.

- *Semiconduttori n.* Il germanio è un elemento del quarto ordine della tavola periodica, e se viene drogato con un elemento del gruppo successivo, vi sarà un elettrone non impegnato nel legame. Il drogaggio, inoltre, altera i livelli aggiungendone uno proibito tra la banda di conduzione e la banda di valenza che può essere facilmente impegnato dall' elettrone libero per agitazione termica dovuta a radiazioni. Tale elettrone contribuirà alla conduzione.
- *Semiconduttori p.* Il semiconduttore p funziona in modo quasi antitetico rispetto al precedente. Si droga il Ge con un elemento del gruppo precedente, in modo tale che il sistema abbia bisogno di catturare un elettrone per stabilizzarsi, formando così una lacuna libera. Come prima il drogaggio crea un livello proibito, ma questa volta è la lacuna non impegnata nella banda di valenza a poter contribuire alla conduzione.

Il rivelatore in questione sfrutta la connessione tra un semiconduttore p ed uno n, in cui le cariche negative e le lacune andranno le une verso le altre lasciando dietro di sé ioni fissi positivi e negativi rispettivamente. Quest' ultimi formeranno un campo elettrico in prossimità della zona di giunzione che se polarizzato inversamente, può diventare abbastanza grande da raccogliere le cariche che si originano in seguito alla radiazione.

Di questo tipo sono i due rivelatori utilizzati in GERDA [4]:

- Rivelatori coassiali a simmetria cilindrica, con un campo elettrico come in figura 1.1b .
- rivelatori di tipo Broad Energy Germanium (BEGe), nei quali uno dei due elettrodi è quasi puntiforme, permettendo di generare un campo elettrico non uniforme. Quest' ultima caratteristica è molto importante e permette di sviluppare tecniche basate sulla forma dell' impulso per discriminare gli eventi di fondo dal segnale.

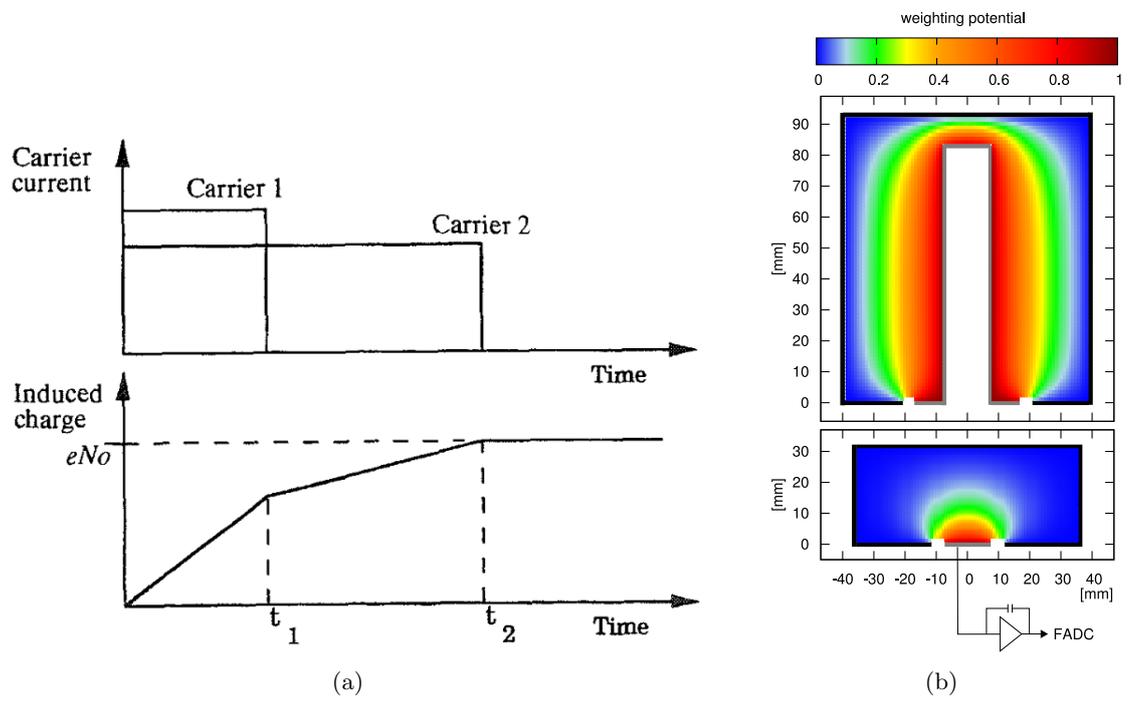


Figura 1.1: (a) esempio di come i portatori di carica generino un impulso [3]. (b) linee di campo per le due tipologie di rivelatori usati in GERDA: coassiali (top), rivelatori BEGe (bottom) [4].

Capitolo 2

Catena d'acquisizione

I dati di GERDA sono stati digitalizzati in due tracce differenti da un Flash-ADC a 14-bit che acquisisce alla frequenza massima di 100MHz. La prima traccia è più breve, lunga $4\mu s$, campionata a 100MHz, contiene il picco principale del segnale e permette di operare la discriminazione. La seconda è di $16\mu s$, campionata a 25MHz e permette invece di ricostruzione lo spettro. I dati sono poi organizzati di step in step in una catena di cosiddetti tier, in cui ogni tier indica un punto di avanzamento diverso dell'analisi. I dati raw, appena uscita dal FADC, vengono denominati tier0, mentre i dati convertiti in una struttura a tree di Root grazie a Gelatio sono detti tier1. I tier2 si ottengono poi dall'elaborazione dei tier1 mediante i moduli ¹ di Gelatio, e i tier3 sono il prodotto di un'ulteriore analisi dei dati. Per una rappresentazione grafica di come è composta questa catena gerarchica di file, si guardi la figura 2.1.

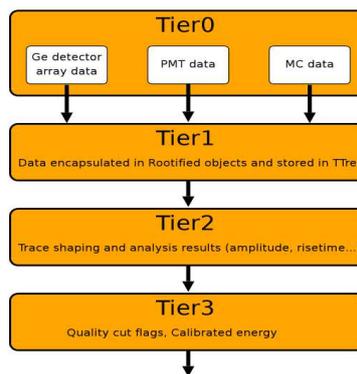


Figura 2.1: Struttura tier dei dati GERDA. (Preso da [5]).

¹

2.1 GELATIO

Gelatio [6] è un ambiente di analisi per processare i segnali d'onda digitalizzati a più canali. Esso è stato ideato e progettato appositamente per GERDA, e si basa sulle librerie di MGDO. Gelatio compie un'analisi dell'onda mediante un'interfaccia grafica composta da una serie di moduli, ossia programmi, con scopi diversi. Un esempio può essere il modulo *baseline*, che calcola appunto il valore della linea di base di un evento, o il modulo *derivative*, che calcola la derivata del segnale. Gelatio, inoltre, permette di passare da un tier all'altro, e per questo il suo utilizzo è fondamentale nell'analisi dei dati di GERDA.

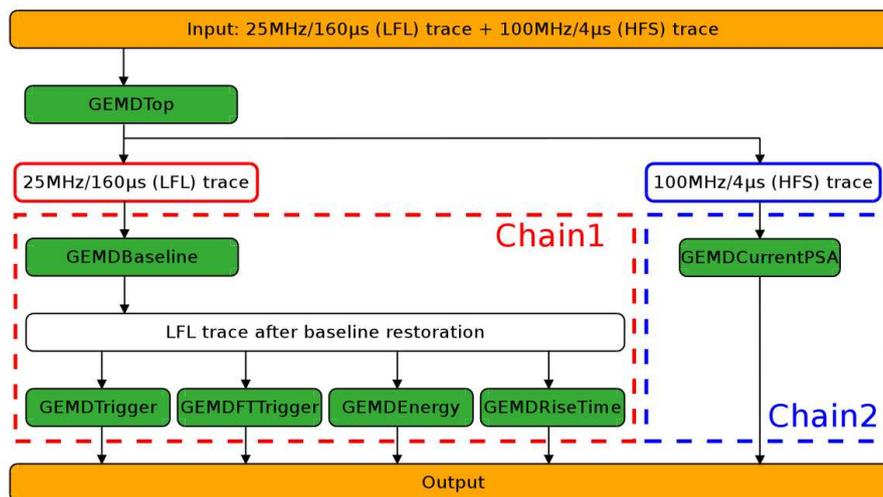


Figura 2.2: Flusso del segnale in GERDA (presa da [2].)

Dall'immagine appare chiaro che le due tracce, campionate una a 100MHz e l'altra a 25MHz, vengano processate attraverso due differenti catene dei moduli di Gelatio; quella campionata a frequenza maggiore serve per ricostruire l'intensità di corrente del segnale, mentre quella acquisita a frequenza minore può venire analizzata da moduli diversi, tra i quali vi è per esempio il modulo che calcola la risetime del segnale.

Capitolo 3

Rete neurale e TMVA

TMVA (Toolkit for MultiVariate Analysis) [7] è un programma di analisi multivariata presente in ROOT [8] che verrà utilizzato in questo lavoro per classificare il segnale dal rumore. Una discriminazione basata su TMVA presenta due fasi diverse: una di training, in cui il classificatore viene allenato a distinguere ciò che si è definito segnale da ciò che è stato definito fondo, ed una classificativa, in cui si applica il classificatore a dei nuovi dati.

3.1 La rete

Adesso, facendo riferimento alla figura 3.1, si descriveranno le parti del multilayer perceptron di TMVA.

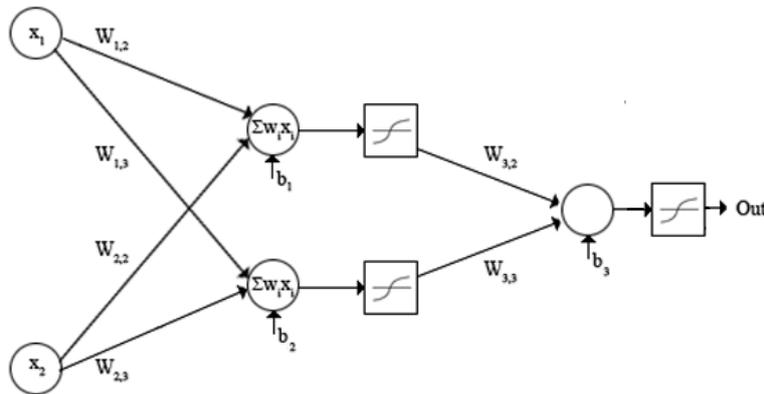


Figura 3.1: Immagine schematica di una rete neurale

- Gli x_i vengono detti nodi, e ad ognuno di essi è associata una variabile di stato n_j .
- w_{ij} sono i diversi pesi delle connessioni tra i nodi.

- ϑ_j è la soglia di attivazione del nodo.
- $f(\sum w_{ij}n_j - \vartheta_j)$ è la funzione di trasferimento¹, in genere una funzione di tipo Fermi (una sigmoide), una tangente iperbolica o anche una funzione lineare.

Il metodo più utilizzato per aggiustare i pesi in modo da ottimizzare la classificazione della rete è detto *back propagation*. Tale algoritmo è supervisionato e si fonda nel minimizzare l'errore tra l'output atteso y_a e il valore fornito dalla rete. Ci sono n eventi di training $x_a = (x_1, \dots, x_n)_a$. Si consideri per semplicità solo un hidden layer, che è uno strato di elaborazione interna della rete compreso tra il primo strato di input e il risultato. La funzione di trasferimento scritta utilizzando la *tanh* è

$$y_{ann} = \sum_j y_j w_{j1}^{(2)} = \sum \tanh(\sum x_i w_{ij}^{(1)}) w_{j1}$$

Il primo membro si riferisce al segnale di output dell'hidden layer verso il neurone finale, mentre nell'ultimo membro si è esplicitata la funzione di trasferimento riguardante i nodi del primo strato. Per questo nei pesi si sono utilizzati gli apici 1 e 2. Si definisce allora una funzione E errore, come

$$E = \sum_a E_a(x_a|w) = \sum_a \frac{1}{2}(y_{ann} - y_a)^2$$

Dove w è il vettore dei pesi della rete. Si cerca allora di aggiornare w sommandogli un vettore la cui direzione è quella per cui E decresce più rapidamente:

$$w^{\rho+1} = w^\rho - \eta \nabla E$$

La formula esplicita si trova poi sostituendo la funzione di trasferimento.

3.2 Convergenza

L'algoritmo esposto precedentemente descrive un metodo per l'aggiornamento dei pesi della rete neurale basato nel sommare a w , il vettore dei pesi della rete, un vettore volto a minimizzare l'errore tra l'output della rete e il responso del teacher ² Si vuole dare qui di seguito risposta alla domanda:

¹L'importante è che essa tenda ad un valore prefissato quando il suo argomento è positivo. In figura è descritta solo da una sommatoria, ma è non cambia a meno di una rinormalizzazione dei w_{ij}

²come si vedrà successivamente la risposta dell'insegnante nel nostro caso è costituita dallo spettro noto della sorgente del torio.

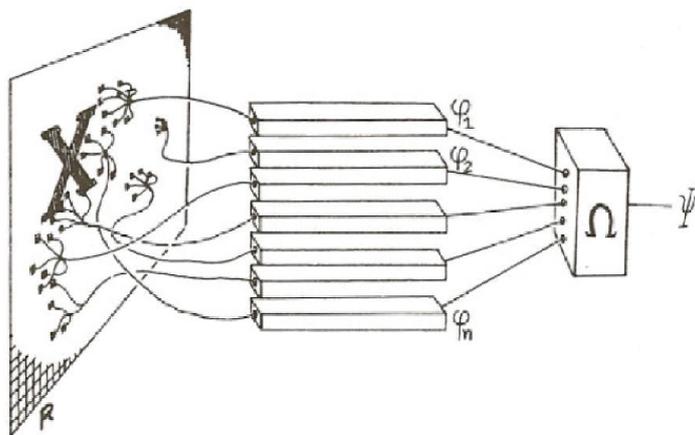


Figura 3.2: Immagine del Perceptron [9]

cosa ci garantisce a priori che la back propagation converga trovando in un numero finito di passaggi i pesi per svolgere la discriminazione? Si desidera rispondere a tale interrogativo non nell'ambito di una rete multilayer, ma in quello più ristretto del perceptrone, la più semplice macchina d'apprendimento composta da un singolo strato parallelo di connessioni disposto come in figura.

Nota: in questa parte si farà riferimento ad un formalismo e a dei concetti per cui si rimanda a [9]

Si immagini una macchina che deve imparare a discriminare due insiemi di figure F^+ ed F^- secondo uno specifico predicato: il primo insieme è quello a cui corrisponde una risposta affermativa, mentre il secondo è quello a cui ne corrisponde una negativa. Una prima macchina che può svolgere il lavoro in questione è quella che salva semplicemente tutte le figure in input già divise negli insiemi F^+ ed F^- per poi discriminare un nuovo elemento solo grazie all'esatto riconoscimento. Ossia, data una nuova figura, il programma va a vedere se ce ne è una uguale salvata in memoria e risponde al predicato. Una macchina del genere, però, possiede poco o niente del concetto di apprendimento.

E' conveniente considerare l'insieme dei coefficienti $\alpha(\psi)$ come un vettore nello spazio $|\Phi|$ -dimensionale e denotarlo con \mathbf{A} , e considerare allo stesso modo anche $\varphi(X)$ e chiamarlo $\Phi(X)$.

Si enuncia allora il seguente teorema, che dimostra come se due classi di figure sono linearmente separabili, allora il perceptrone riuscirà a discriminarle con un numero finito di passaggi.

Th. Il seguente algoritmo cambierà il valore di A un numero finito di volte.

Inizio: Si scelga un qualsiasi valore di A

Test: Si scelga una X da F^+ o F^-

Se $X \in F^+$ e $A * \Phi > 0$ torna a **Test**

Se $X \in F^-$ e $A * \Phi < 0$ torna a **Test**

Se $X \in F^+$ e $A * \Phi < 0$ vai a **Aggiungi**

Se $X \in F^-$ e $A * \Phi > 0$ vai a **Sottrai**

Aggiungi: $A = A + \Phi(X)$ Vai a **Test**

Sottrai: $A = A - \Phi(X)$ Vai a **Test**

Dim: si rimanda a [9].

Il teorema afferma quindi che questo processo di allenamento, nel caso in cui una soluzione esista, preveda un numero finito di passaggi. Si deve prestare attenzione tuttavia al fatto che l' algoritmo descritto nella dimostrazione può diventare laborioso nel caso in cui il vettore Φ sia troppo grande, in questo caso l' equilibrio potrebbe venire superato più e più volte, o troppo piccolo, in questo caso l' equilibrio sarebbe difficilmente raggiungibile. Per questo si preferisce utilizzare Φ normalizzato al posto di Φ . Inoltre, nel caso limite in cui la soluzione del predicato sia unica o costituita da un insieme troppo piccolo di valori il programma potrebbe anche non trovarla. Questi casi possono essere evitati con ipotesi più forti al teorema 2³

Hill-Climbing. Con *Hill-Climbing* si intende un semplice algoritmo per riuscire a ricavare la posizione di un massimo di una funzione $G : R^n \mapsto R$. Si inizializza in modo casuale un punto di partenza A e gli si aggiunge un vettore Φ per cui $A_{i+1} = A_i + \Phi$. Il vettore Φ può essere trovato in più modi, ma ci basti pensare ad esso come il primo vettore per cui $G(A + \Phi) \geq G(A)$. Ecco che il successo o meno di questo algoritmo è equivalente al successo del teorema di convergenza, mostrando quindi tutti i limiti di quest' ultimo per casi nemmeno troppo esotici. Perché il processo di *Hill-Climbing* non funzioni opportunamente basti pensare alla funzione in figura 3.3, $f(x, y) = \sin(x)\sin(y)/(xy)$.

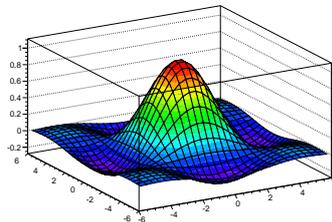


Figura 3.3: $f(x,y)=\sin(x)\sin(y)/(xy)$

³In pratica si richiede che, data una famiglia F di vettori normalizzati, esistano sempre un $\delta > 0$ e un A^* tali che $A\Phi > \delta \forall \Phi \in F$.

Capitolo 4

Metodologia

L'obiettivo dell'esperimento GERDA è di misurare il decadimento doppio beta ($0\nu\beta\beta$). Siccome gli unici prodotti del decadimento considerato sono due elettroni, l'energia del decadimento sarà ben definita, pari a $Q_{\beta\beta}$, che è la differenza energetica tra il nucleo padre ed il nucleo figlio (a riposo). Per il germanio $Q_{\beta\beta} = 2039\text{keV}$. Il decadimento cercato rilascerà l'energia in una zona specifica del rivelatore, e gli eventi con questa forte localizzazione sono chiamati single site event (SSE). Oltre ai SSE esistono poi eventi che, al contrario, depositano energia in punti diversi del rivelatore e sono denominati multiple site event (MSE).

Come è già stato accennato la difficoltà dell'analisi è dovuta al valore estremamente elevato di $T_{1/2}$, la vita media del decadimento ($0\nu\beta\beta$), che in GERDA suggerisce un valore di aspettazione del numero di eventi cercati pari a circa quattro in un intero anno. Si comprende allora come il rumore ambientale possa diventare così rilevante nei dati acquisiti, e come ci sia bisogno di un metodo di analisi che possa classificare il segnale dal fondo e isolare l'evento ricercato. Le tecniche che svolgono un'analisi con questo intento sono chiamate Pulse Shape Analysis. Eventi diversi corrispondono a segnali elettrici di forma diversa. Come si vede in figura 5.1, per un SSE l'evento è caratterizzato da una sigmoide la cui derivata è costituita da un solo picco, mentre per un MSE la sigmoide presenta un punto di flesso in più, e quindi una derivata a due picchi adiacenti.

Ci si aspetta che l'energia rivelata a $Q_{\beta\beta}$ possa essere quindi dovuta o al decadimento ($0\nu\beta\beta$) o a multiple site events di fondo.

4.1 Sorgente di Torio

Visto che si desiderano usare tecniche di analisi multivariata per compiere la discriminazione tra il doppio decadimento beta senza emissione di neutrini ed eventi di fondo, in particolare tra SSE e MSE, è importante avere dei dati con cui allenare la rete. Quest'ultima difatti deve compiere l'allenamento su dati che sono già noti, e per farlo si sono utilizzati degli eventi acquisiti dall'esposizione dei rivelatori di GERDA ai raggi γ provenienti da una sorgente di torio nota (^{228}Th). Ci si aspetta che le radiazioni γ che incidono sul rivelatore producano tutte e tre le interazioni possibili tra radiazione elettromagnetica e materia: effetto Compton, effetto fotoelettrico, e produzione di coppie (figura 4.2). La produzione di coppie è un processo in cui l'energia di un fotone viene convertita in un elettrone ed un

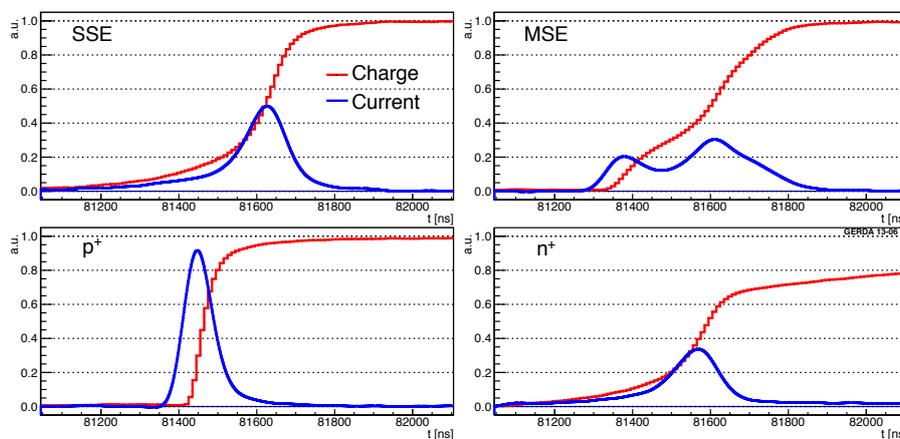


Figura 4.1: Forma del segnale (e sua derivata) per diverse tipologie di eventi nei rivelatori al germanio.

positrone, e perchè questo processo avvenga è necessario che l' energia del fotone sia almeno uguale all' energia dell' elettrone e del positrone a riposo, pari a 1022keV (511keV per particella). La sua sezione d' urto, inoltre, aumenta all' aumentare dell' energia, ed è quindi più probabile osservare la produzione di coppie ad energie elevate. Il positrone successivamente interagirà con gli elettroni del materiale più vicino e si annichilerà nuovamente per produrre una nuova coppia di fotoni in accordo con la conservazione del quadrimomento. A seconda che vengano rivelati entrambi i fotoni, solo uno, o nessuno, si possono presentare tre casi differenti:

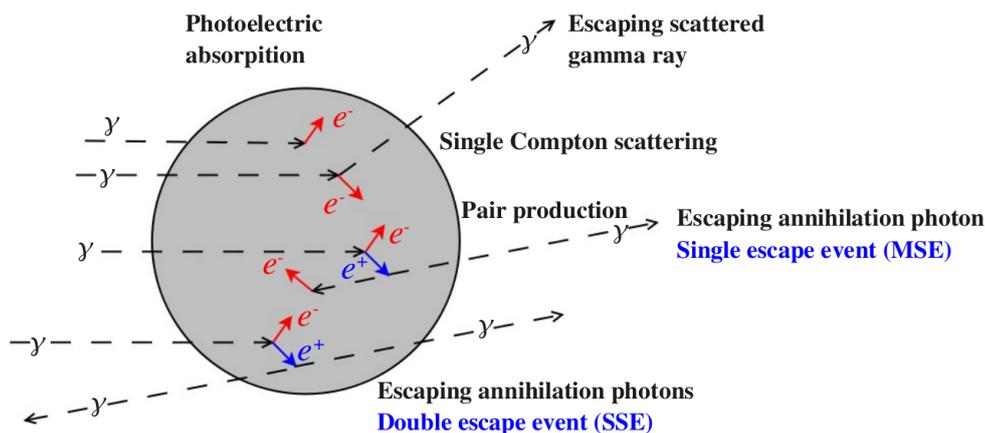


Figura 4.2: Interazioni del germanio. Sono mostrati in figura l' effetto fotoelettrico, l' effetto Compton, e la produzione di coppie; per quest' ultima si raffigura la differenza tra i SSE e i MSE.

- Full energy peak (FEP). Tutta l' energia del fotone iniziale viene rilasciata nel rivelatore. Questo evento può essere conseguente sia alla pair production di cui si è appena parlato, sia per esempio all' efetto fotoelettrico o ad una combinazione di questi.
- Single escape peak (SEP), uno dei fotoni prodotti dall' annichilazione del positrone non viene rivelato. Questo evento presenta cosi' un'energia di 511keV inferiore al FEP.
- Double escape peak (DEP), entrambi i fotoni prodotti dall' annichilazione del positrone non vengono rivelati. L' evento ha un'energia di 511 keV inferiore rispetto al SEP.

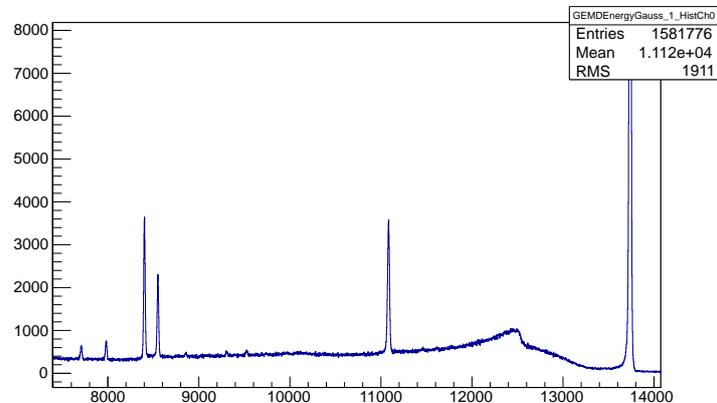


Figura 4.3: Spettro di energia non calibrato per una sorgente di ^{228}Th in un rivelatore al germanio. Partendo da destra sono visibili il FEP, il SEP, e dopo il picco del Bi e quello del DEP

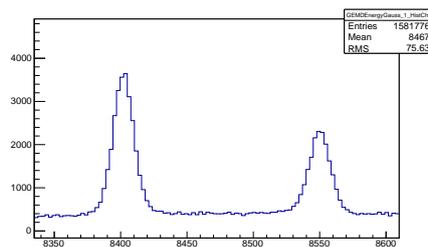


Figura 4.4: Picco DEP (sulla sinistra), picco del bismuto (sulla destra).

La figura 4.4 mostra la parte che ci interessa (almeno per l' allenamento) dello spettro del torio. Per il training della rete neurale si sono scelti infatti i due picchi di Figura 4.4: il picco a sinistra con l' energia di 1592keV è il DEP, mentre quello a destra con l' energia di 1620keV è dovuto al bismuto (^{212}Bi) ed è composto quasi interamente da MSE.

4.2 Introduzione all' analisi

Il segnale che si è scelto per l' analisi e l' allenamento è stato acquisito dal rivelatore esposto alla sorgente di torio.

I dati uscenti dal FADC sono dei tier1, quindi il primo passo consiste nel convertire questi rowdata nel formato tier2, leggibile da ROOT. Per farlo, come si è accennato precedentemente, si è utilizzato GELATIO. Successivamente si sono letti gli eventi provenienti dal tier1 osservando la forma della funzione d' onda (si riveda figura 4.1).

Il passo successivo è stato quello di ricercare delle variabili per la discriminazione.¹

- *risetime*. La prima scelta è ricaduta sulla risetime, che è il tempo di durata del segnale dal trigger fino alla stabilizzazione dell' onda. La risetime può difatti discriminare i SSE dai MSE per il fatto che i multiple site event hanno una funzione $q(t)$ più dilatata dovuta alla presenza di due eventi successivi. Per calcolarla si sono usati prima i moduli di GELATIO, e poi si è costruita una macro personalmente che, acquisiti i dati dal tier1, pulisce l' evento con un algoritmo di smooth, calcola la baseline, il massimo del segnale, e tramite essi riesce a ricavare la risetime.
- A/E . Questa variabile è il rapporto tra l' ampiezza del picco della derivata dell' evento, ossia il picco dell' intensità di corrente in funzione del tempo, e l' energia. Nei MSE si prende solo l' ampiezza del primo picco, che in questo modo sarà più bassa rispetto a quella dei SSE. Visto inoltre che l' energia, corrispondente all' integrale del segnale, rimane circa la stessa, A/E presenterà valori diversi a seconda se si stia trattando il picco del bismuto o quello dovuto al double escape peak.

Ricavate le variabili, si è utilizzata la rete neurale schematizzata in figura 4.5.

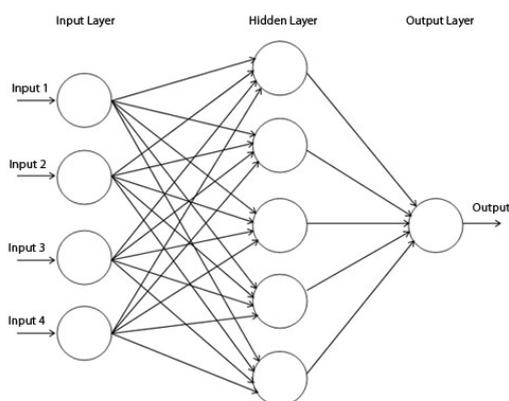


Figura 4.5: Multilayer perceptron

Il multilayer perceptron (MLP) è un perceptron a più layer con il schema di nodi come in figura 4.5. Si è utilizzata questa rete neurale per l' analisi dati per discriminare i SSE dai MSE.

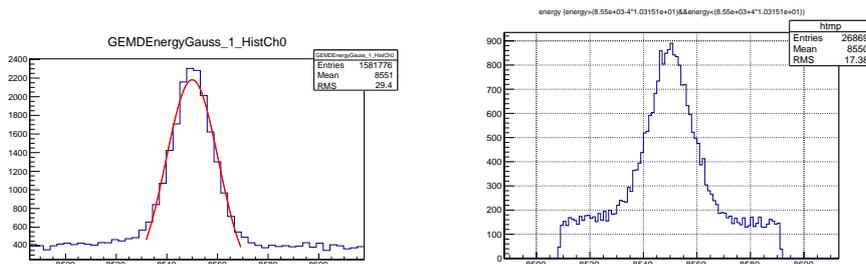
¹Si faccia riferimento alla figura 4.1 fino alla fine del paragrafo

Capitolo 5

Analisi dei dati

5.1 Allenamento

Come si è scritto sopra, si sono presi dallo spettro del torio il picco relativo al bismuto come esempio di background, ed il picco relativo al DEP come esempio di segnale.



(a) Esempio di fit sul picco del bismuto

(b) Picco del bismuto, presenza del fondo

Figura 5.1: Picco del bismuto fittato con una gaussiana

Come si vede dalla 5.1 , anche selezionando i due picchi sarà sempre presente una parte di fondo significativa. Fittati poi i due picchi separatamente con una gaussiana, si è scelta una larghezza opportuna, misurata in sigma, per effettuare il taglio e ricavarne i dati da analizzare. La larghezza della curva di Gauss non dovrà essere nè troppo grande, in questo caso si considererebbe una percentuale di fondo troppo elevata, nè troppo piccola, perchè si trascurerebbero molti dati e si avrebbe una statistica più debole. Il taglio si è effettuato allora a due sigma, con una percentuale del fondo calcolata del 30%.

Prima di applicare il multilayer perceptron ai dati, si sono prima osservate graficamente le differenze dei valori delle variabili risetime e A/E tra gli eventi DEP e quelli provenienti dal picco del bismuto. La figura 5.2 raffigura in blu i valori delle variabili del bismuto, e in rosso i valori provenienti dal DEP .

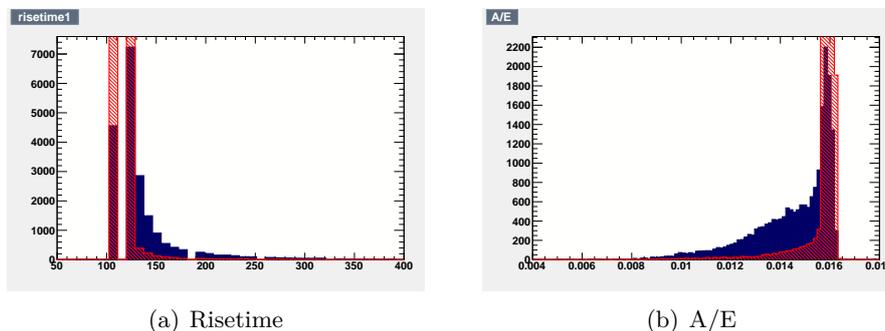


Figura 5.2: Variabili risetime ed A/E a confronto

Siccome è possibile notare a colpo d'occhio la diversità dei due set di dati, si presuppone che la rete neurale riesca a compiere la discriminazione proposta. Ecco allora i grafici 5.3 , ricavati con TMVA.

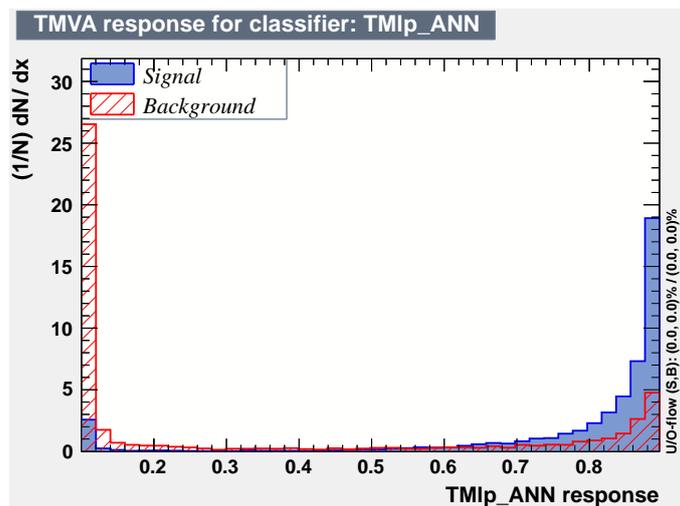
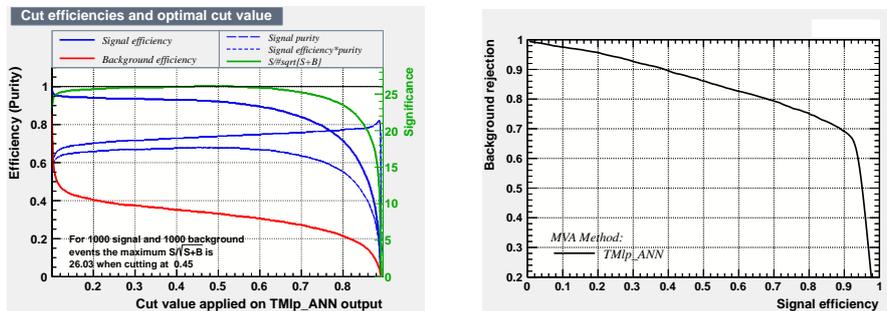


Figura 5.3: Risposta del multilayer perceptron di root. La figura mostra la distribuzione del segnale e del fondo ricavate dall' output del classificatore.

Come si vede da 5.3 l' overlap delle distribuzioni del segnale e del background è circa un quarto dell' area totale dell' istogramma; inoltre il segnale è molto distaccato dalla zona in cui è presente la maggior parte del fondo. Questi due aspetti ci suggeriscono che la classificazione sia andata a buon fine. Il valore della significanza al variare della posizione del taglio e la reiezione del fondo al variare dell' efficienza del segnale, rappresentati in 5.1 (a) e (b), ci confermano inoltre la bontà della discriminazione. La significanza presenta difatti valori superiori a venti sigma per più dell' 80% dei tagli possibili, mentre la curva in 5.1 presenta valori stabili e maggiori di 0.7 per la maggior parte del grafico. La figura 5.1 (a) ci fornisce inoltre il valore ottimale su cui tagliare i dati.

Successivamente si è rifatta più volte la discriminazione con un taglio della



(a) Efficienza e significanza al variare del taglio del campione

(b) Background rejection vs signal efficiency

Figura 5.4: Efficienza e significanza al variare del taglio.

gaussiana più ampio; il risultato è stato un innalzamento sistematico della zona centrale dell' istogramma per entrambe le figure rappresentate. Ciò suggerisce in modo quantomeno intuitivo¹ che il taglio dei picchi del DEP e del bismuto giochino un ruolo non trascurabile nella bontà della discriminazione.

Si desidera ora riportare il grafico 5.5, raffigurante le coordinate parallele.

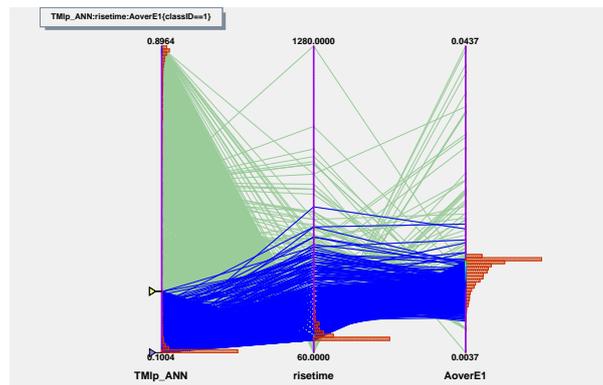


Figura 5.5: Coordinate parallele

Questo grafico rappresenta un modo per mostrare la corrispondenza di ogni variabile, e fornisce un metodo visivo per controllare le ricorsività del comportamento dei vari dati. Se si considera lo spazio tridimensionale composto dalle variabili in questione, ogni spezzata del grafico rappresenterebbe un punto. In questo modo tale grafico ci fornisce un' idea di come il discriminatore ha lavorato, potendo guardare il percorso compiuto da ogni singolo dato. Sulle linee di demarcazione delle diverse variabili vengono poi mostrati il set di valori più influenti di ciascuna variabile tramite dei piccoli istogrammi. Gli eventi con i colori più sgargianti sono i più importanti.

¹Si tenga conto che la rete neurale è una sorta di black box

Il multilayer perceptron usato nell' analisi precedente è composto da due hidden layer e compie duecento cicli. Cambiare il numero di hidden layer riducendolo ad uno (visto che si sono utilizzate due variabili è l' unica possibilità) dovrebbe comportare dei cambiamenti nella classificazione. Nel nostro caso non si riporta l' istogramma relativo al perceptrone ad un layer perchè le variazioni dal 5.3 sono trascurabili. L' unico modo per avere un output della rete neurale molto diverso con questi dati sarebbe quello di diminuire il numero di cicli fermando la rete prima che essa potesse convergere.

5.2 Applicazione

Dopo aver allenato la rete neurale e aver salvato in un file i pesi con cui essa si era stabilizzata, si è voluto applicarla per discriminare i MSE dai SSE nello spettro del torio 4.3. La figura dello spettro così ottenuta è la seguente, 5.6, in cui in blu si sono contrassegnati gli eventi del segnale, mentre in rosso il fondo.

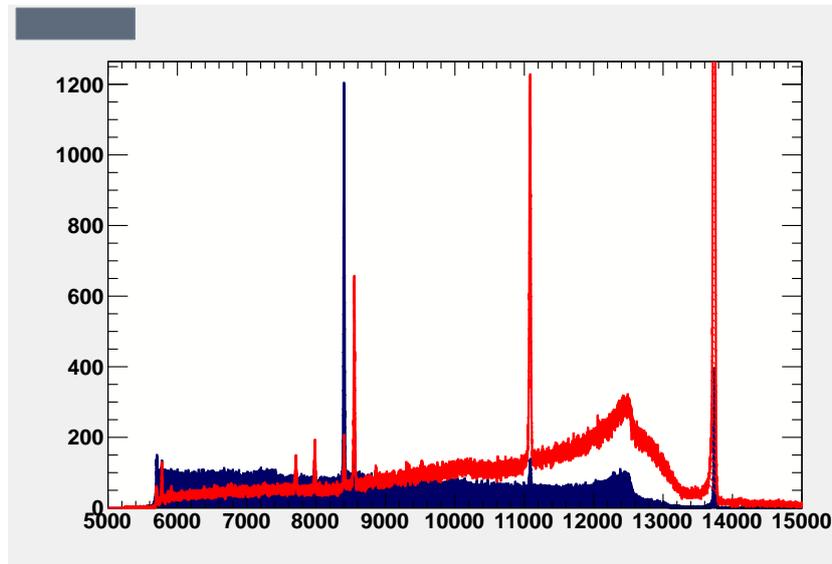


Figura 5.6: Spettro del torio, SSE e MSE

Si riportano il grafico 5.7, relativo alla spalla Compton, ed il 5.8, raffigurante il FEP.

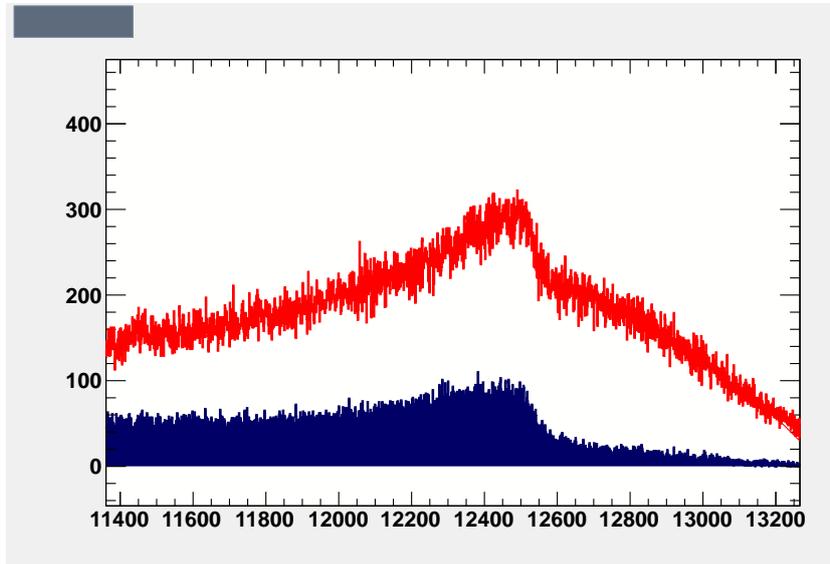


Figura 5.7: Spettro del torio, spalla Compton. Nel grafico viene raffigurata la spalla Compton per lo spettro del torio dopo la discriminazione compiuta dalla rete neurale.

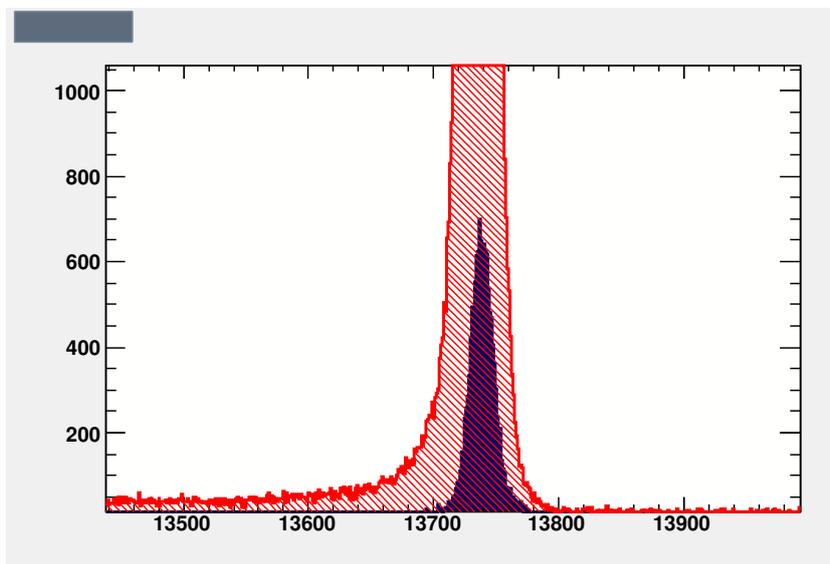


Figura 5.8: Spettro del torio. FEP dopo la discriminazione compiuta dalla rete neurale.

Capitolo 6

Conclusioni

In questo lavoro l'obiettivo è stato l'utilizzo di una rete neurale per compiere una discriminazione tra i multiple site event (MSE) e i single site event (SSE) nello spettro del torio (^{228}Th). Per ricavare due set di dati con un'alta percentuale o di MSE o di SSE si è effettuato un taglio attorno al picco del double escape peak (SSE) e del picco corrispondente al decadimento del bismuto (MSE). Successivamente si sono utilizzate le variabili rilettime ed il rapporto tra l'ampiezza dell'intensità di corrente e l'energia per allenare la rete neurale e compiere la discriminazione. Come si è visto, i risultati provenienti dal classificatore mostrano che la rete neurale riesce a discriminare i dati accettando l'80% degli eventi single site contro il 20% degli eventi double site. Si ritiene quindi che la classificazione sia avvenuta con successo. Si vuole focalizzare l'attenzione sul fatto che la bontà dei risultati dipenda in modo non trascurabile dal taglio effettuato per i due picchi considerati nello spettro del torio. Aumentando la larghezza dei picchi gli eventi del bismuto e i DEP sono sempre più simili tra loro e la rete neurale riesce sempre di meno a compiere la classificazione. Si conclude infine che, visti i risultati, il multilayer perceptron ha un'ottima efficienza nel discriminare i SSE dai MSE.

Bibliografia

- [1] K. H. Ackermann et al. (GERDA Collaboration), *Eur. Phys. J C* 73 (2013) 2330
- [2] W. Rode Johann et al., *J. Phys G.: Nucl. Part. Phys.* 39 634 (2012) 124008
- [3] G. F. Knoll, “Radiation Detection and Measurement, III edizione (2000) John Wiley and Sons
- [4] M. Agostini et al, (GERDA Collaboration) *Eur. Phys. J. C* 73 (2013) 2583.
- [5] M. Agostini et al., Offline data processing and analysis for GERDA experiment, *Journal of Physics: Conference Series* 368 (2012) 012047, arXiv/1111.3582
- [6] M. Agostini et al., GELATIO: a general framework for modular digital analysis of high-purity Ge detector , *JINST* 6 P080 (2011), arXiv/1106.1780
- [7] A. Hecker et al. TMVA 4 Toolkit for Multivariate Data Analysis with ROOT, Novembre 2009
- [8] R. Brum and F. Rodenakers, “ROOT-An Object Oriented Data Analysis Framework”, proceeding AIMENP '96, *Nucl. Inst. Math* A389 (1997) 81
- [9] M. L. Minsky, Seymour A. Papert, *Perceptrons, An Introduction to Computational Geometry*, The MIT Press 1988