



Classificazione dei bosoni elettrodeboli con una rete neurale al Large Hadron Collider

6 Novembre 2024

Candidato:
Jacopo Lancione

Relatore:
Prof. Emanuele Roberto Nocera



Sommario

- Introduzione
 - LHC
 - Machine Learning
- Dataset
 - Produzione dei bosoni
 - Preprocessing
- Reti Neurali
 - Architettura e principi
- Risultati
- Conclusioni



UNIVERSITÀ
DI TORINO

Introduzione

Large Hadron Collider - CMS



UNIVERSITÀ
DI TORINO

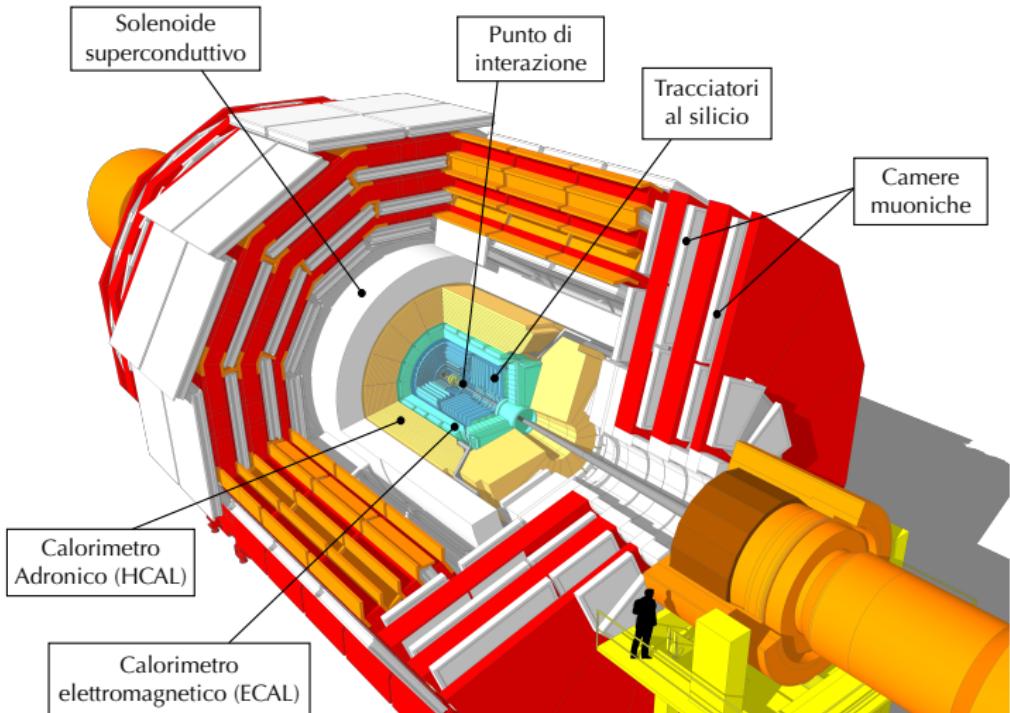
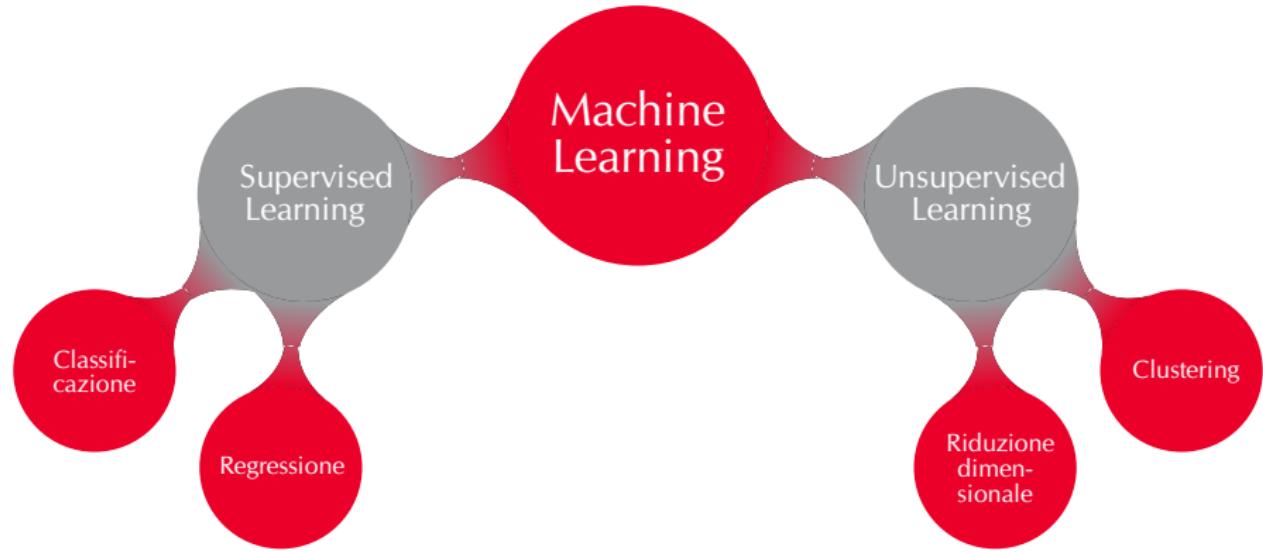
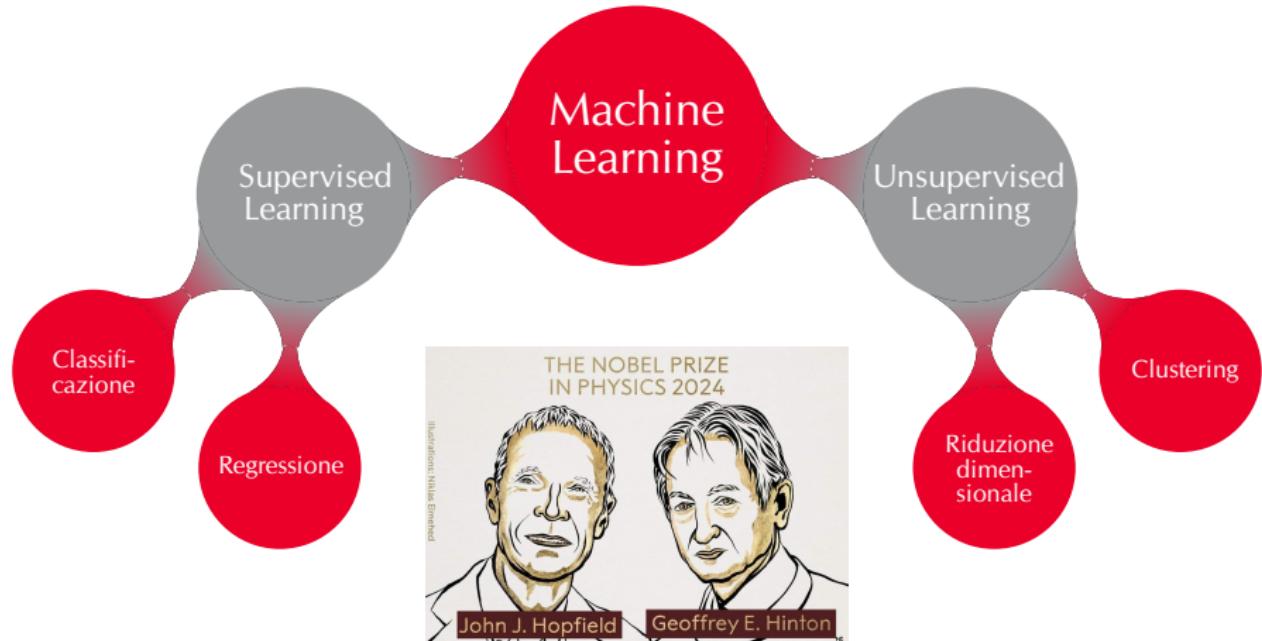


Image adapted from <https://cds.cern.ch/record/2665537>

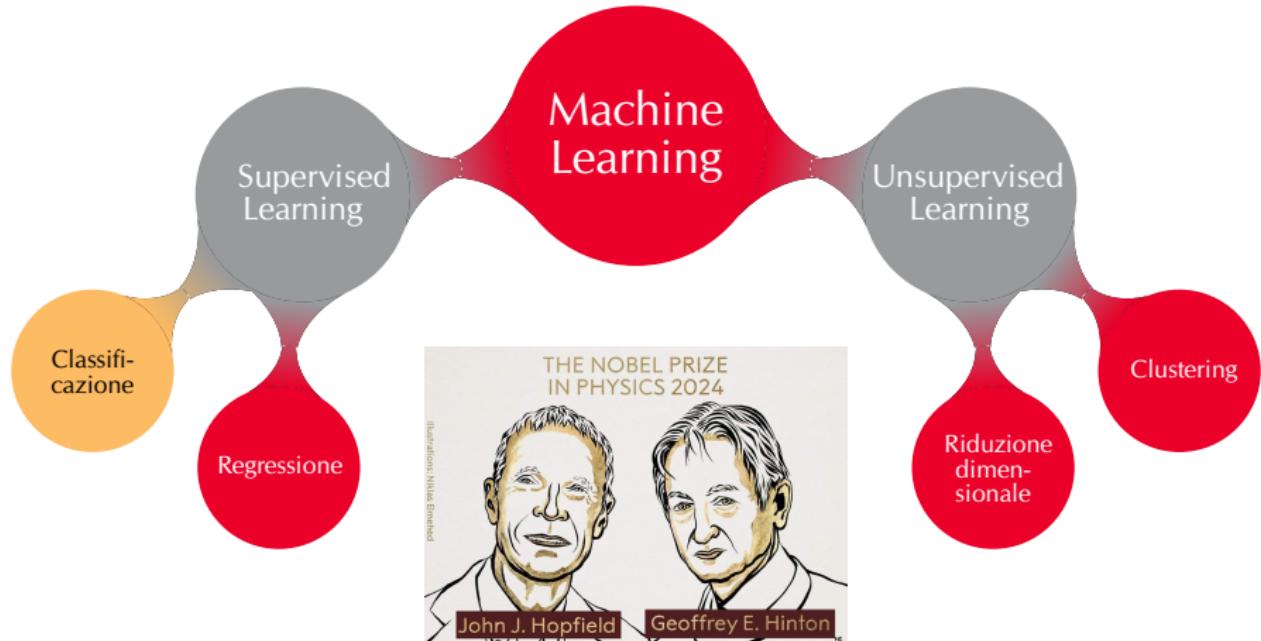
Machine Learning: uno sguardo d'insieme



Machine Learning: uno sguardo d'insieme



Machine Learning: uno sguardo d'insieme



Il Progetto di tesi

Classificare

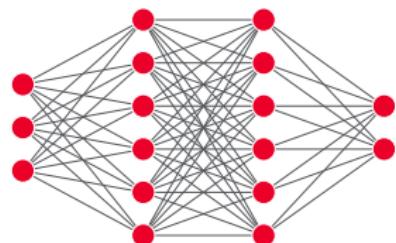
Bosoni elettrodeboli

con una

Rete neurale



Adapted from: https://it.wikipedia.org/wiki/Modello_std





UNIVERSITÀ
DI TORINO

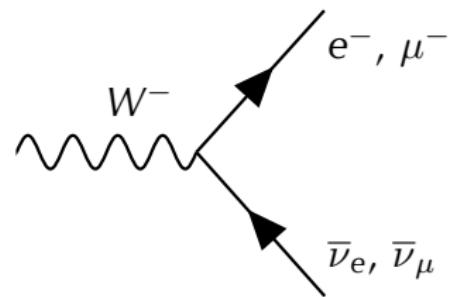
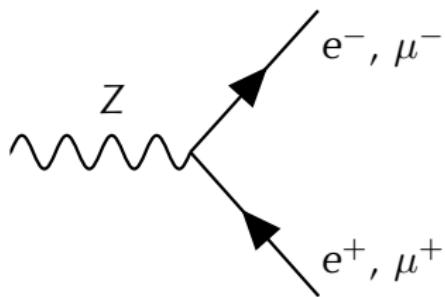
Il Dataset

$$pp \rightarrow Z \rightarrow \ell \bar{\ell}$$

$$pp \rightarrow W \rightarrow \ell \bar{\nu}, \nu \bar{\ell}$$

$(91.188 \pm 0.002) \text{ GeV}/c^2$

$(80.369 \pm 0.013) \text{ GeV}/c^2$



La struttura del dataset

Energia nel centro di massa
7 TeV

- Run
- Event
- Parametri dei leptoni (Q , p_t , η , ϕ)
- Parametri del detector

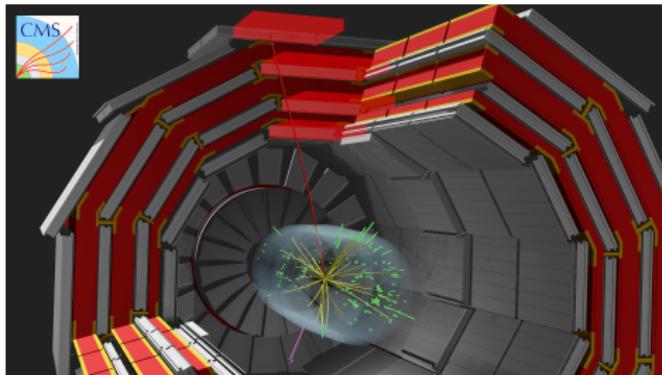


Image from: <https://cds.cern.ch/record/2909335?ln=en>

Datasets derived from the Run2011A:
<https://opendata.cern.ch/record/545>

Preprocessing

Trattamento degli outliers → Correlazioni → Normalizzazione

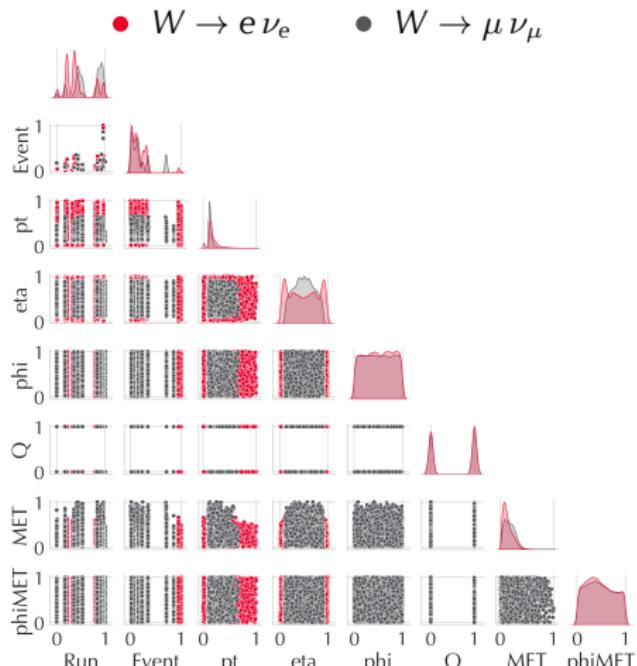
- $Z \rightarrow \mu^+ \mu^-$



- $Z \rightarrow e^+ e^-$

- $W \rightarrow e \nu_e$

- $W \rightarrow \mu \nu_\mu$



Dataset

Bisogna evitare di insegnare alla rete il rumore statistico dei dati

- Test: per verificare che la rete generalizzi
- Train: per allenare la rete
- Validation: per valutare l'allenamento della rete



Suddivisione del dataset

T
e
s
t

Bisogna evitare di insegnare alla rete il rumore statistico dei dati

- Test: per verificare che la rete generalizzi
- Train: per allenare la rete
- Validation: per valutare l'allenamento della rete



Suddivisione del dataset



Bisogna evitare di insegnare alla rete il rumore statistico dei dati

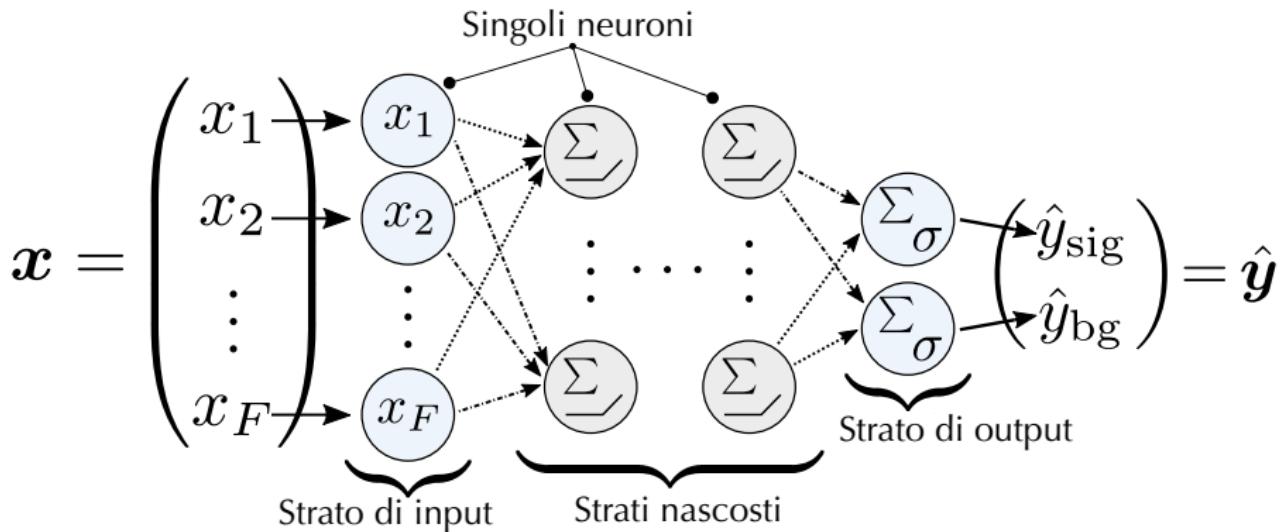
- Test: per verificare che la rete generalizzi
- Train: per allenare la rete
- Validation: per valutare l'allenamento della rete



UNIVERSITÀ
DI TORINO

Reti Neurali

L'architettura



Adapted from: <https://github.com/ejmastnak/fmf-seminar/blob/main/media/vector/fcn-architecture-simple.pdf>

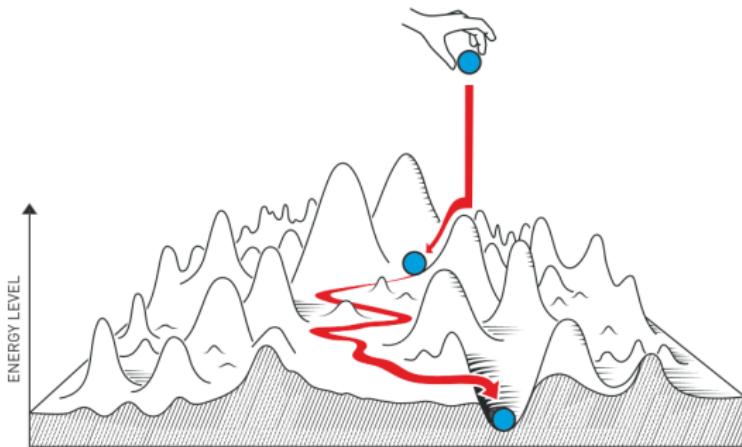
Loss function

Algoritmi di ottimizzazione



$$P(\mathcal{D}|\mathbf{w}) = \prod_{i=1}^n q(x_i|\mathbf{w})^{p(x_i)} (1 - q(x_i|\mathbf{w}))^{1-p(x_i)}$$

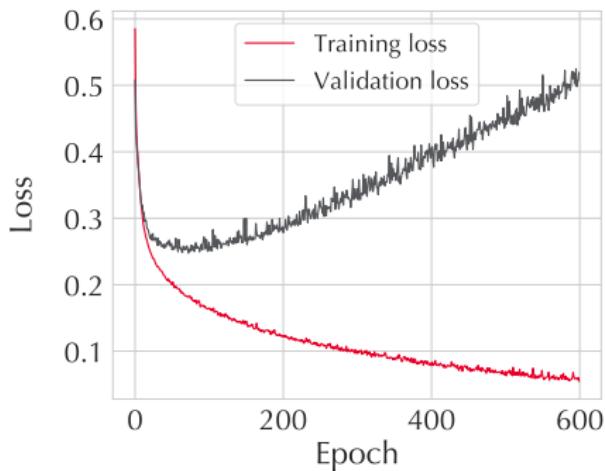
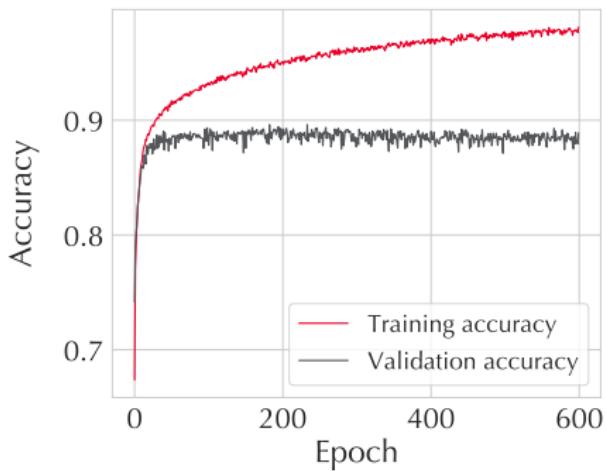
$$\rightsquigarrow H(p, q) = - \sum_{i=1}^n p(x) \log q(x)$$



©Johan Jarnestad/The Royal Swedish Academy of Sciences

- Gradient Descent
- RMS
- Adam
- Nadam
- ...

Overfitting





UNIVERSITÀ
DI TORINO

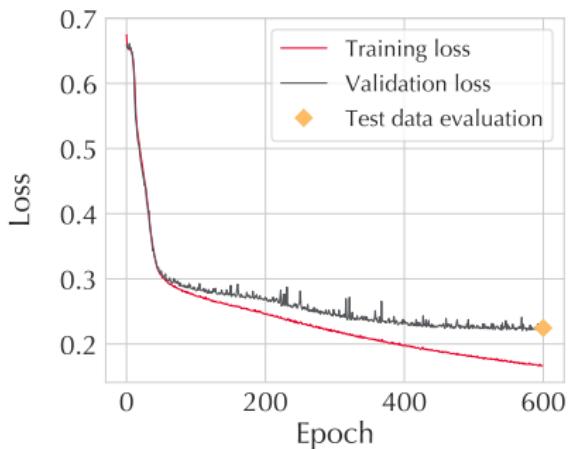
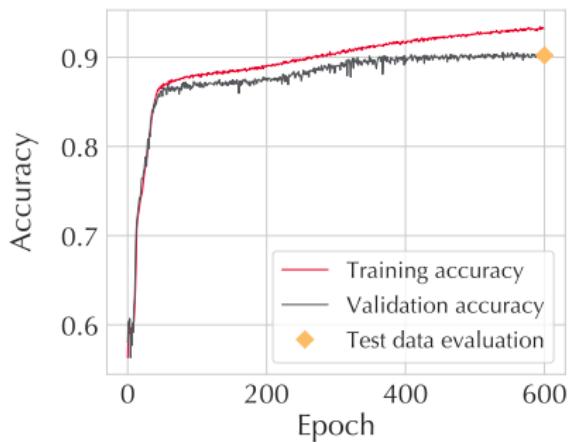
Risultati

Risultati - Dataset Z

architettura: (64,64)

algoritmo: Adam

attivazione: sigmoide



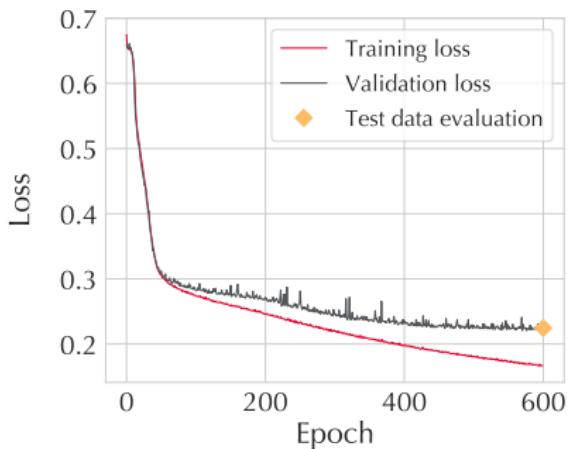
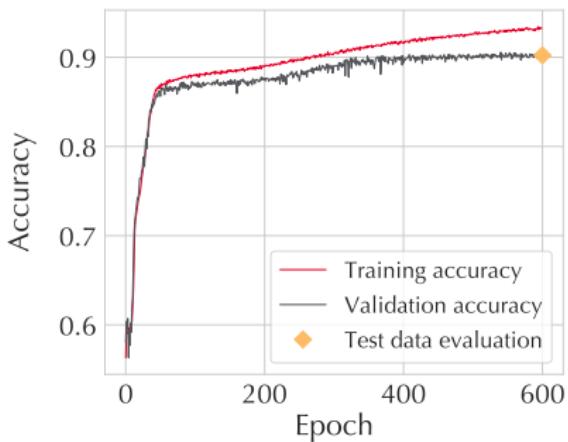
[https://github.com/jlancione/thesis_notebooks]

Risultati - Dataset W

architettura: (64,64)

algoritmo: Adam

attivazione: sigmoide



Il modello generalizza al test set

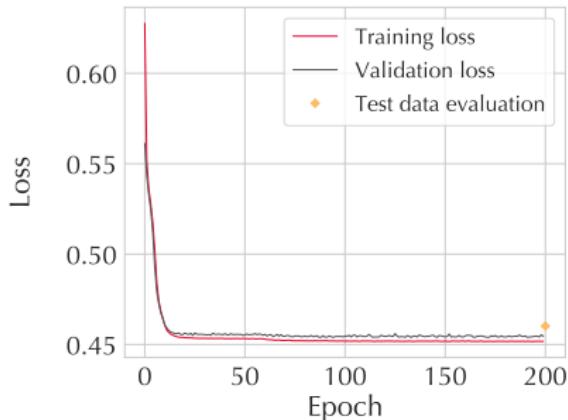
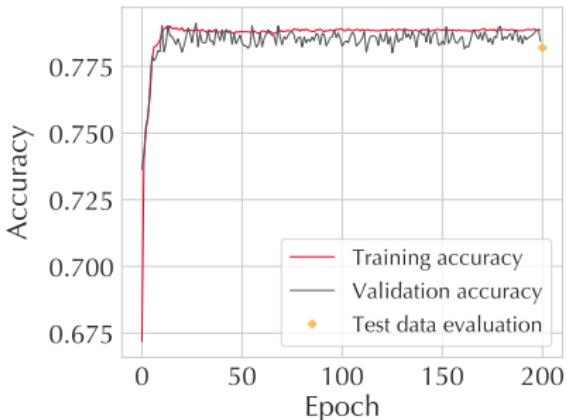
[https://github.com/jlancione/thesis_notebooks]

Risultati - Dataset W

architettura: (3,3)

algoritmo: Adam

attivazione: relu



Il modello generalizza al test set

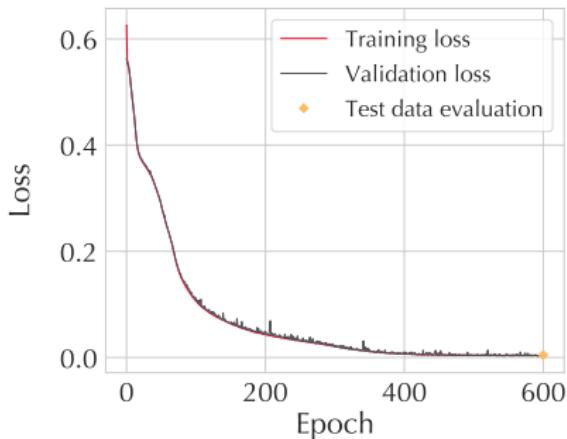
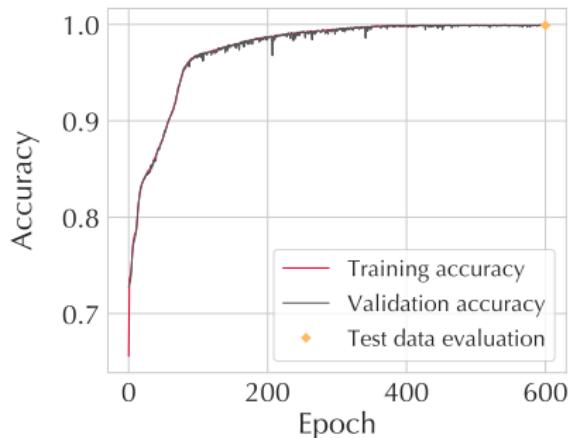
[https://github.com/jlancione/thesis_notebooks]

Risultati - Dataset W

architettura: (24,12,24)

algoritmo: Adam

attivazione: sigmoide



Il modello generalizza al test set

[https://github.com/jlancione/thesis_notebooks]



UNIVERSITÀ
DI TORINO

Conclusioni



Conclusioni

- Le reti neurali si prestano molto bene all'identificazione di particelle
- L'allenamento è efficace perché **generalizza** al test set
- Ho identificato una classe di **modelli equivalenti**

Ulteriori sviluppi:

- Combinare i dataset rimuovendo le labels per allenare una rete a distinguere i bosoni tra loro
- Utilizzare una di queste reti come modello generativo per simulare decadimenti



Conclusioni

- Le reti neurali si prestano molto bene all'identificazione di particelle
- L'allenamento è efficace perché **generalizza** al test set
- Ho identificato una classe di **modelli equivalenti**

Ulteriori sviluppi:

- Combinare i dataset rimuovendo le labels per allenare una rete a distinguere i bosoni tra loro
- Utilizzare una di queste reti come modello generativo per simulare decadimenti

Grazie