







DALLE PARTICELLE ELEMENTARI ALL'INTELLIGENZA ARTIFICIALE E RITORNO

STEFANO FORTE
UNIVERSITÀ DI MILANO & INFN



UNIVERSITÀ DEGLI STUDI DI MILANO
DIPARTIMENTO DI FISICA

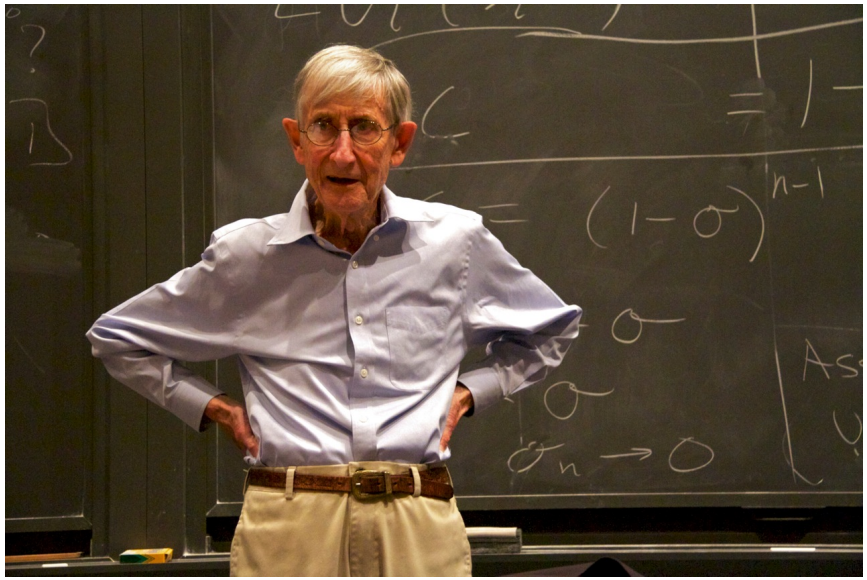


CISF 2019

MILANO, 6 MARZO 2019

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 740006

FISICA DELLE PARTICELLE E LHC



"There are historical reasons not to expect too much from the LHC. (...)

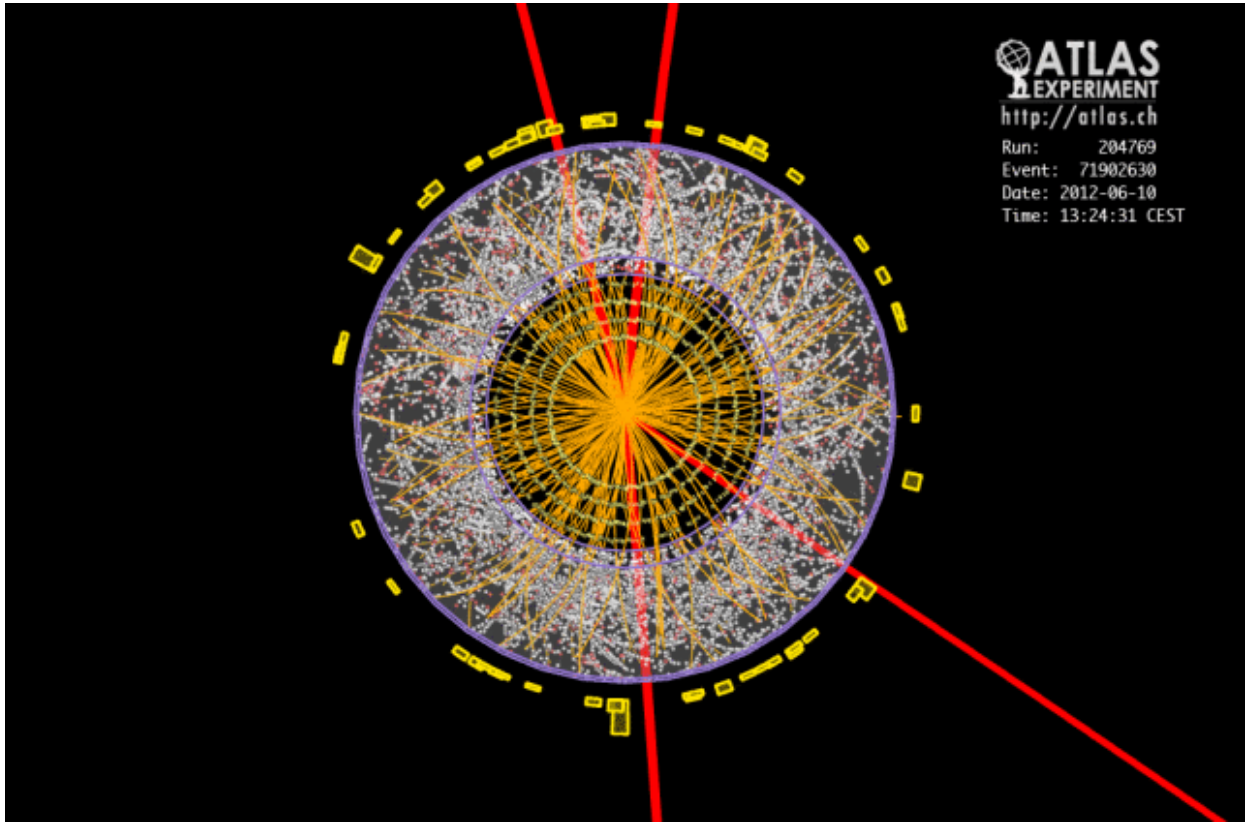
There have been sixteen important discoveries" (in HEP)
"between 1945 and 2008:
four discoveries on the energy frontier,
four on the rarity frontier,
eight on the accuracy frontier"

Freeman Dyson, 2008

QUAL È IL PROBLEMA?

"There are two reasons to be skeptical about the importance of the LHC: one technical and one historical".

Un bosone di Higgs prodotto in collisioni pp,
decade in due coppie $\mu^+\mu^-$

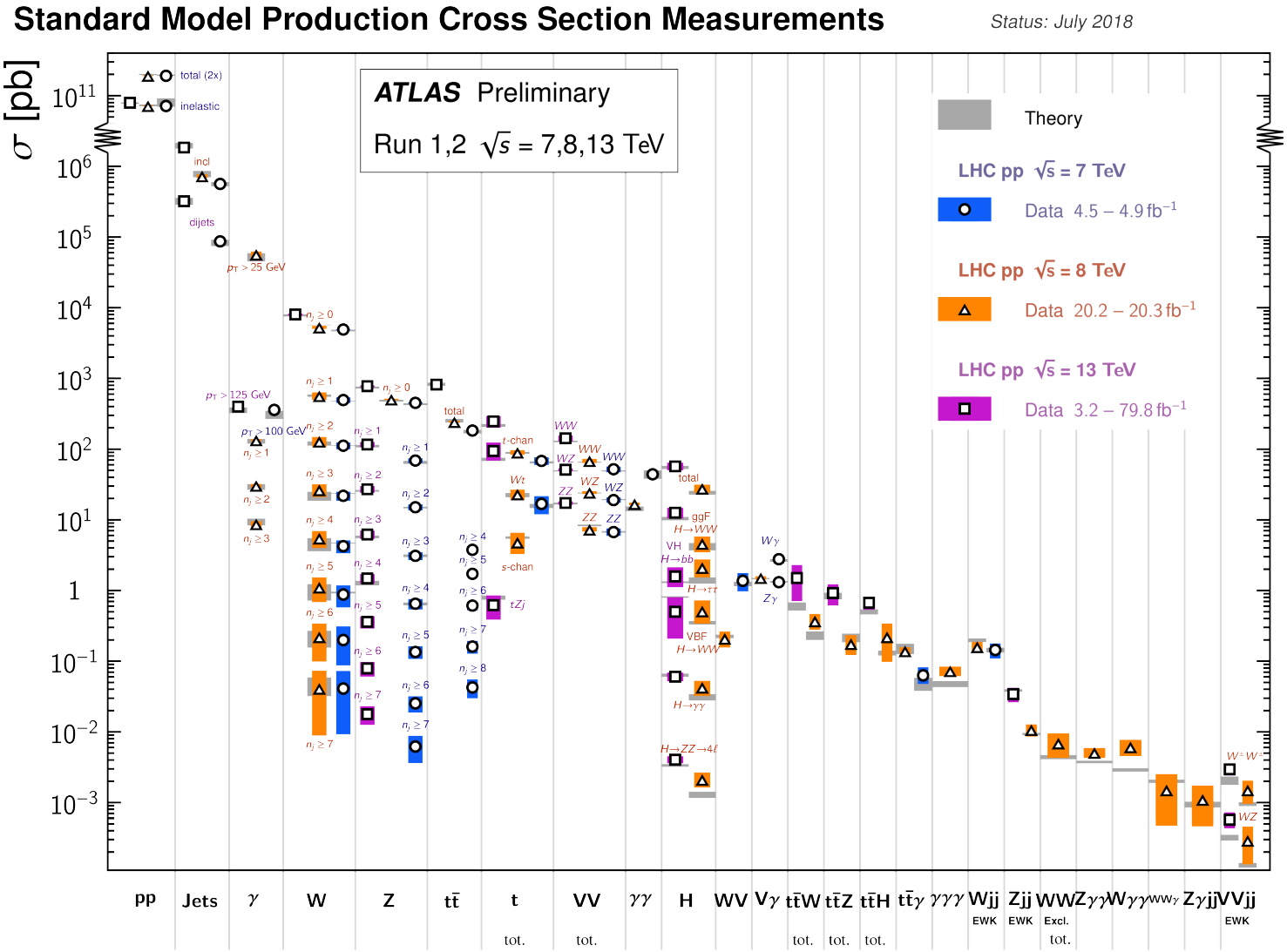


"The technical weakness of the LHC arises from the nature of the collisions that it studies. These are collisions of protons with protons, and they have the unfortunate habit of being messy"

Freeman Dyson, 2008

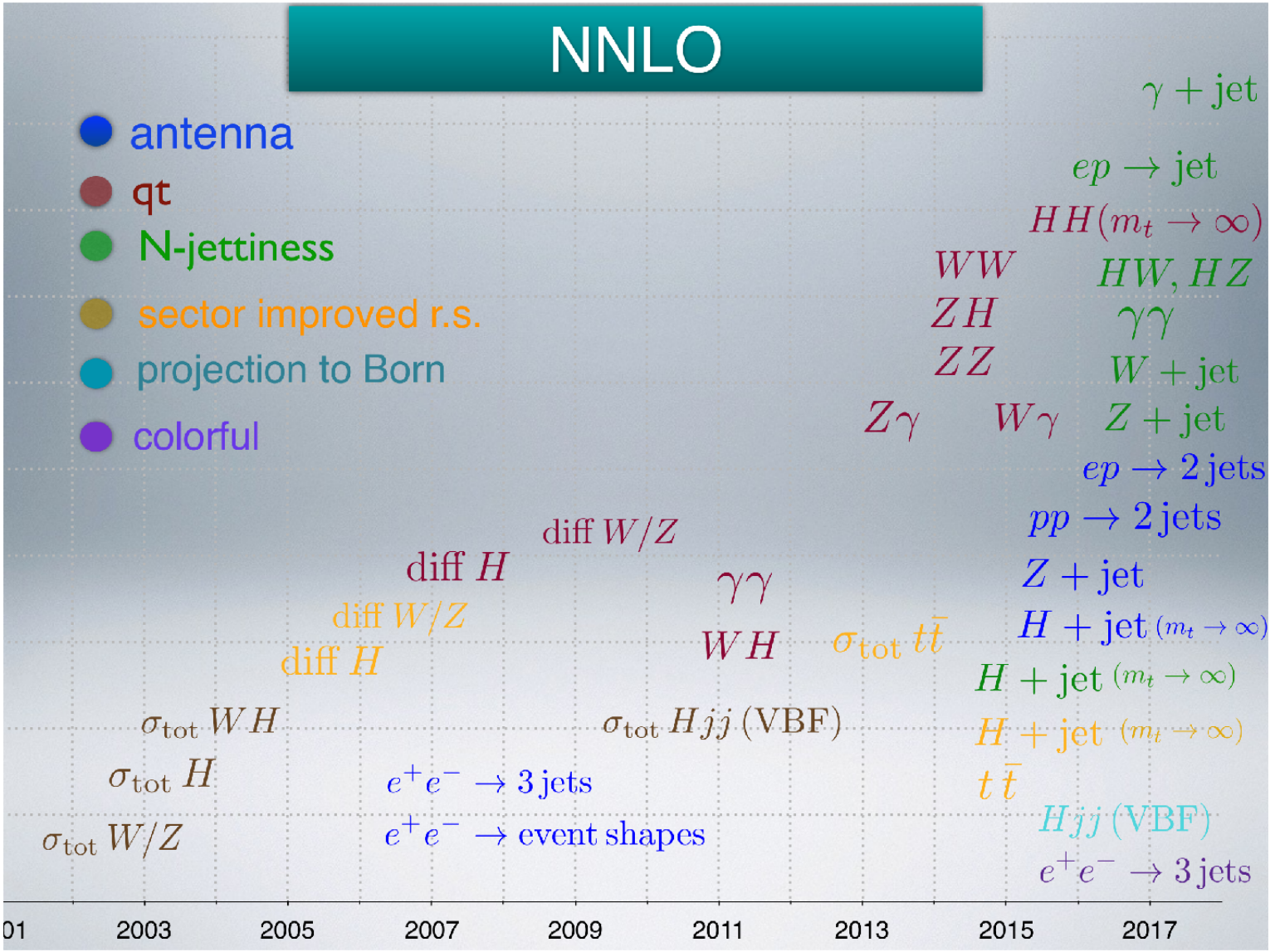
COME SONO ANDATE LE COSE VERAMENTE?

MISURA DI PRECISIONE DI PROCESSI DEL MODELLO STANDARD A LHC



DOVE HA SBAGLIATO DYSON?

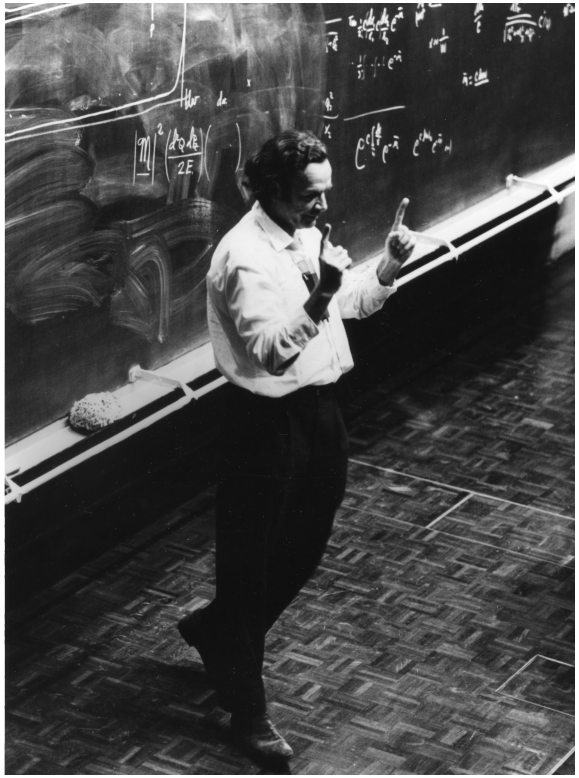
CALCOLI AL TERZO ORDINE PERTURBATIVO NELL'INTERAZIONE FORTE



(G. Heinrich, 2017)

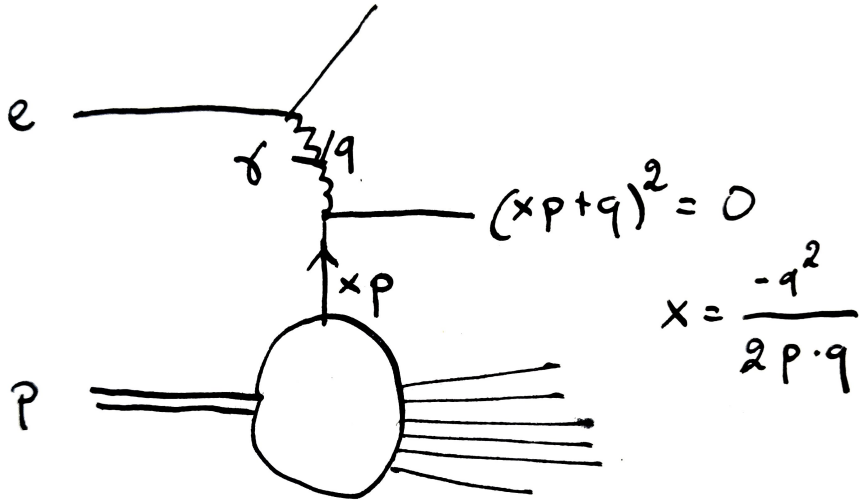
CALCOLI PERTURBATIVI?
IL "MIRACOLO" DELLA FATTORIZZAZIONE

I "PARTONI" DI FEYNMAN



Feynman spiega il modello
a partoni al CERN nel 1970

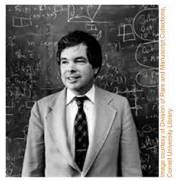
RADIOGRAFIA DEL PROTONE
CON UN FOTONE DI PICCOLA LUNGHEZZA D'ONDA λ:

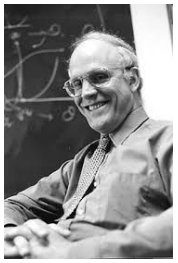


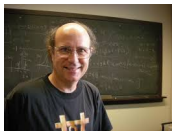
- SUPPONIAMO CHE IL FOTONE COLPISCA UN "PARTONE" LIBERO E SENZA MASSA (QUARK, GLUONE) CHE PORTA UNA FRAZIONE x DELL'ENERGIA DEL PROTONE $p_{\text{quark}} = xP_{\text{proton}}$
- IL VALORE DI x È MISURABILE: FISSATO DALLA CINEMATICA
- LA SEZIONE D'URTO È PROPORZIONALE ALLA PROBABILITÀ $q_i(x)$ DI ESTRARRE UN PARTONE DI TIPO i CON FRAZIONE DI ENERGIA x NEL PROTONE

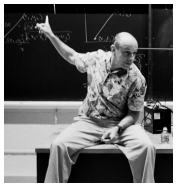
IL "MIRACOLO" DELLA FATTORIZZAZIONE

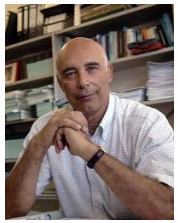
LA QCD PERTURBATIVA

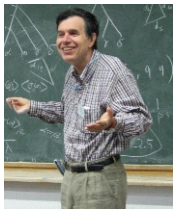

Wilson


Gross


Wilczek


Politzer

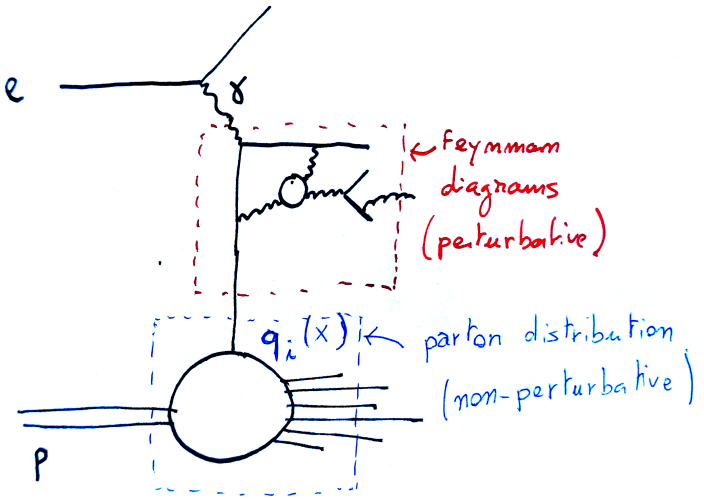

Altarelli


Parisi


Collins


Sterman

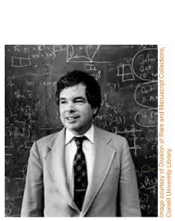
RADIOGRAFIA DEL PROTONE
CON UN FOTONE
DI PICCOLA LUNGHEZZA D'ONDA λ:



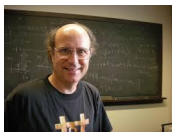
- IL MODELLO A PARTONI È IL PRIMO ORDINE DI UNA SERIE PERTURBATIVA
- A OGNI ORDINE PERTURBATIVO LA SEZIONE D'URTO FATTORIZZA:
  $\sigma_{\text{misurata}} = \hat{\sigma}_{\text{perturbativa}} \otimes q_i$
- SI PUÒ DIMOSTRARE RIGOROSAMENTE CON LO SVILUPPO DI WILSON

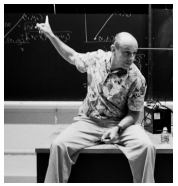
LA QCD PERTURBATIVA

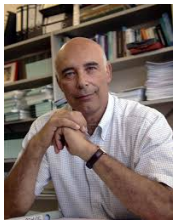
URTO PROTONE-PROTONE CON PICCOLA SCALA DI DISTANZA λ:

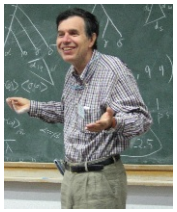

Wilson


Gross


Wilczek

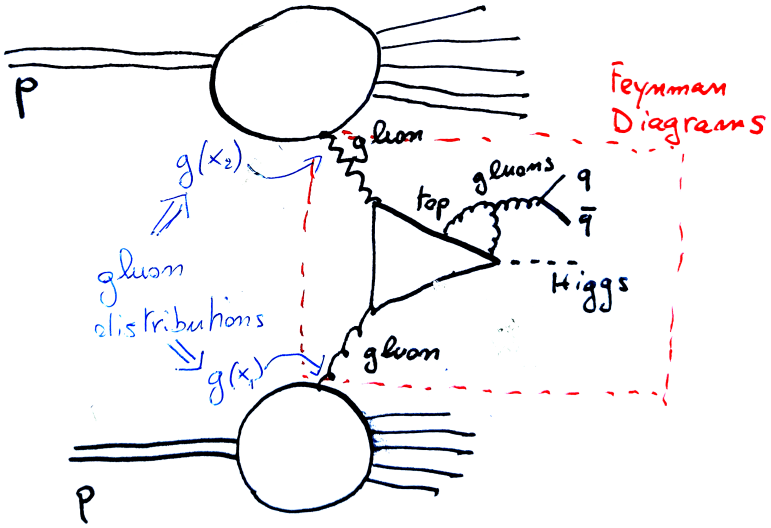

Politzer


Altarelli


Parisi


Collins


Sterman

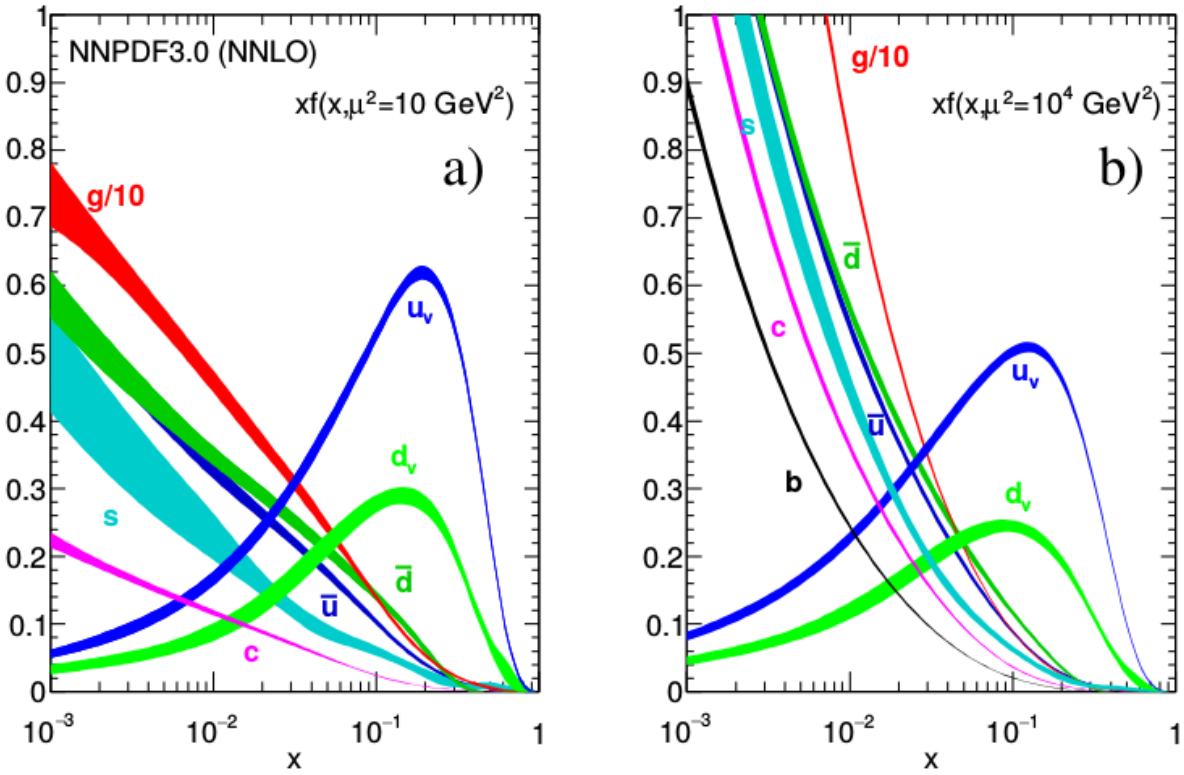


$$\lambda \sim \frac{1}{M_{\text{Higgs}}}$$

produzione di Higgs al quarto ordine perturbativo

- IL MODELLO A PARTONI È IL PRIMO ORDINE DI UNA SERIE PERTURBATIVA
- A OGNI ORDINE PERTURBATIVO LA SEZIONE D'URTO FATTORIZZA:
  $\sigma_{\text{misurata}} = \hat{\sigma}_{\text{perturbativa}} \otimes q_i \otimes q_j$
- SI PUÒ DIMOSTRARE RIGOROSAMENTE CON LO SVILUPPO DI WILSON

RITRATTO DEL PROTONE
VISTO DA UN BOSONE DI HIGGS



(PDG 2018)

LE "DISTRIBUZIONI PARTONICHE"

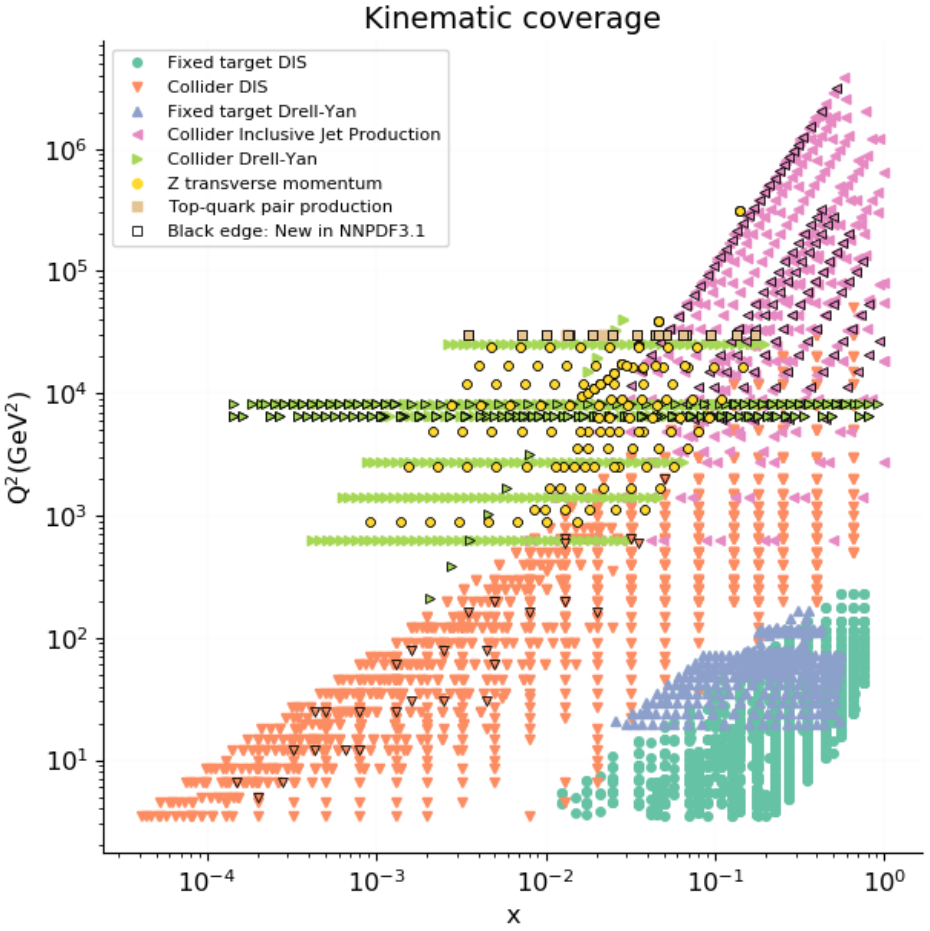
- DISTRIBUZIONI DI ENERGIA
  PER OGNI SPECIE DI QUARK, ANTIQUARK, E IL GLUONE
- $\int_0^1 f_{\text{up}}(x)\,dx = N_{\text{up}}, \ldots$
- "VALENZA" $N_{\text{up}} - N_{\text{antiup}} = 2$; $N_{\text{down}} - N_{\text{down}} = 2$
- DIPENDONO DALLA RISOLUZIONE $\mu \sim \frac{1}{\lambda}$ (SCALA DI ENERGIA $\sim \frac{1}{\text{distanza}}$);
  dipendenza dalla risoluazione calcolabile

DOVE HA SBAGLIATO DYSON?

COME SI FA IL RITRATTO?

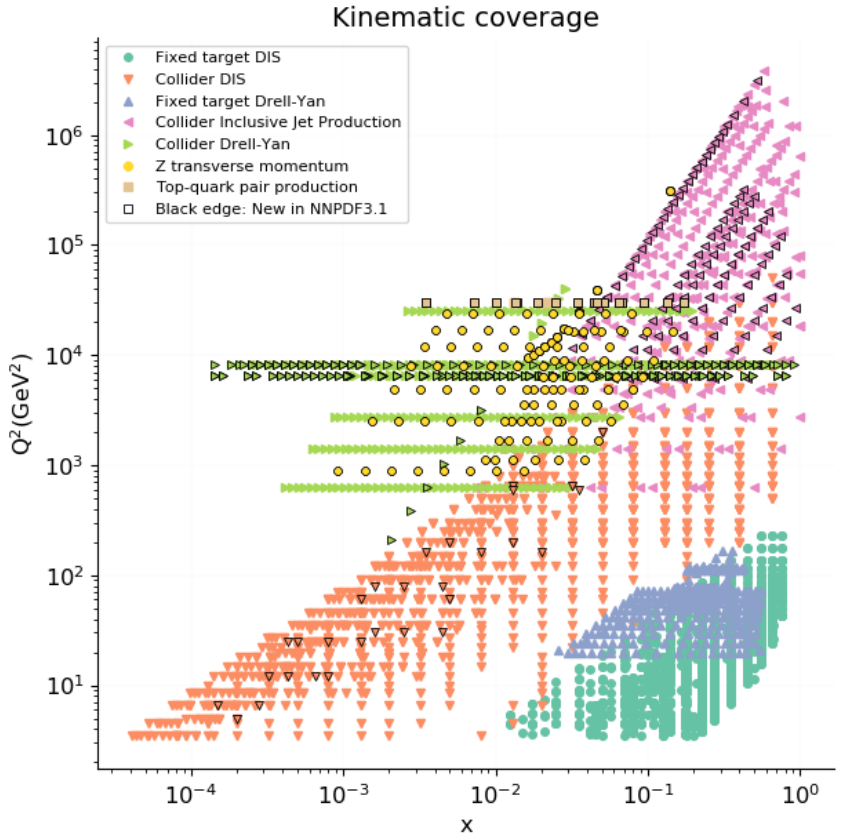
- CALCOLO DI UN PROCESSO FISICO $\Rightarrow$ DIPENDE DALLA PDF
- CONFRONTO CON I DATI
- INVERSIONE

UN TIPICO INSIEME DI DATI (NNPDF3.1)



DOVE HA SBAGLIATO DYSON?

UN TIPICO INSIEME DI DATI(NNPDF3.1)



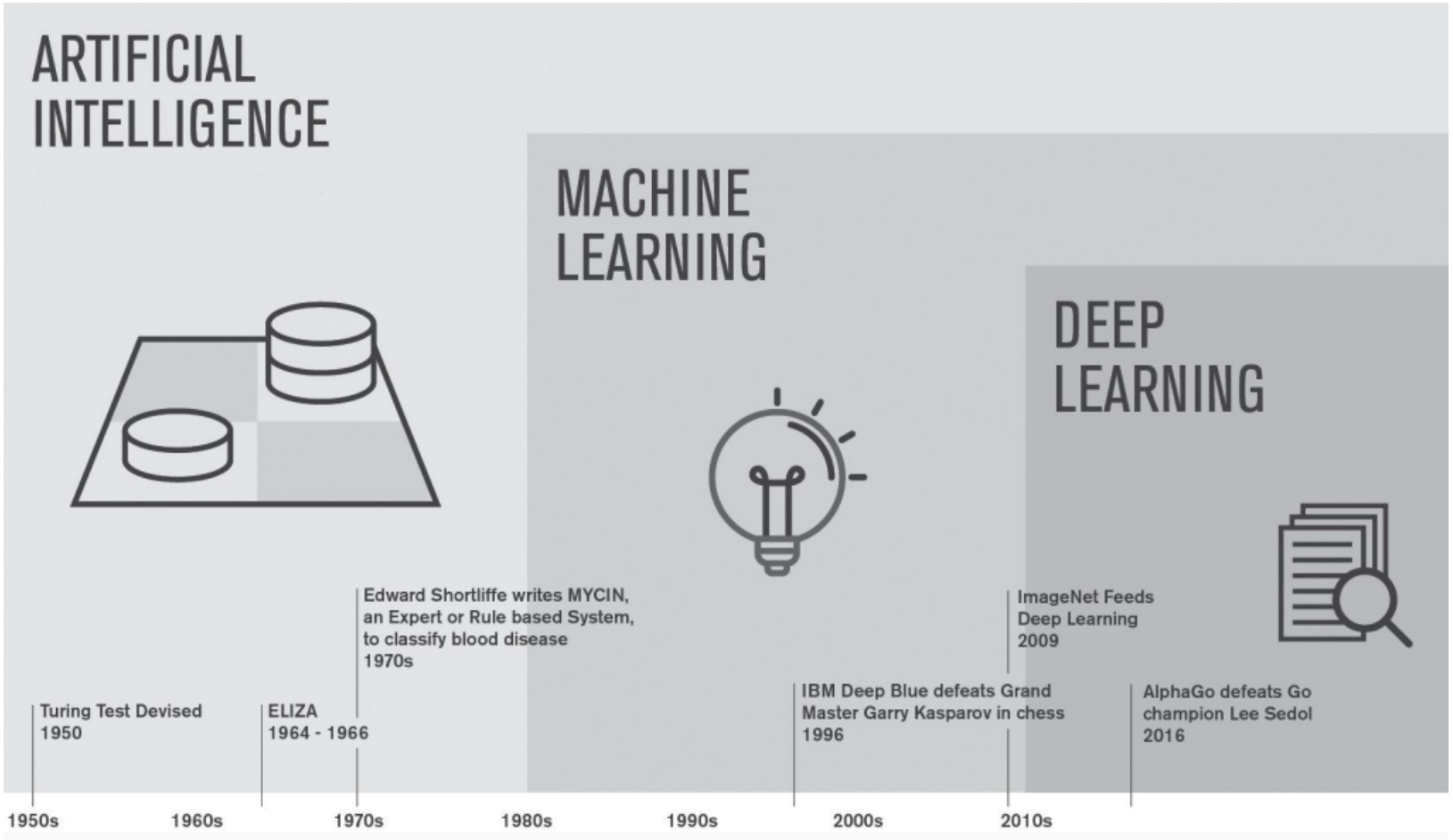
COME È POSSIBILE??

- OGNI PDF È UNA FUNZIONE una distribuzione, in realtà
  $\Rightarrow$ SI DEVE DETERMINARE UNA
  PROBABILITÀ IN UNO SPAZIO DI FUNZIONI

- SI DEVE DETERMINARE UN
  FUNZIONALE (CONTINUO) DA UN INSIEME DISCRETO DI DATI SPERIMENTALI

DOVE HA SBAGLIATO DYSON?

HA FATTO I CONTI SENZA L'INTELLIGENZA ARTIFICIALE!

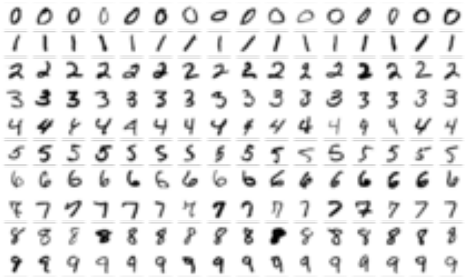


INTELLIGENZA ARTIFICIALE: PARADIGMI

"KNOWLEDGE BASED" AI

- APPRENDIMENTO E IMPLEMENTAZIONE DI UN INSIEME DI REGOLE
- BENE PER GLI SCACCHI,
  MALE PER LA VITA VERA



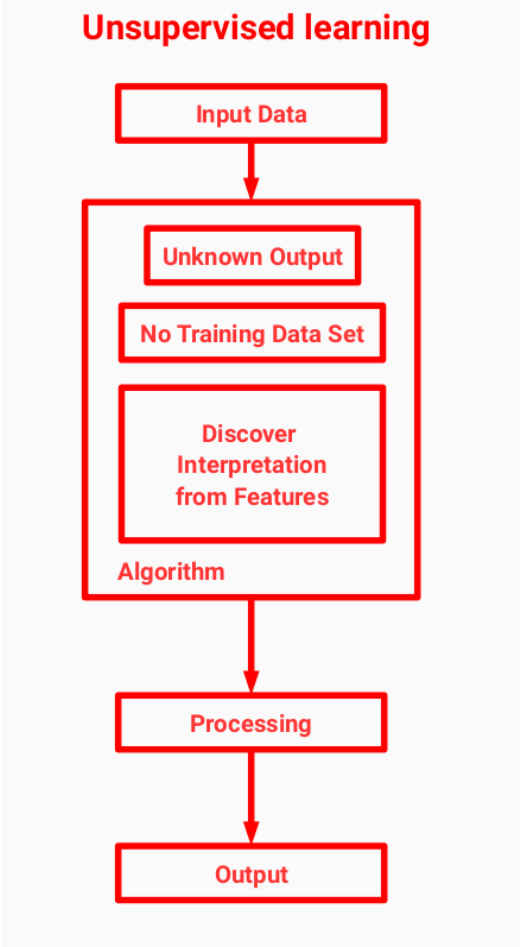


MACHINE LEARNING

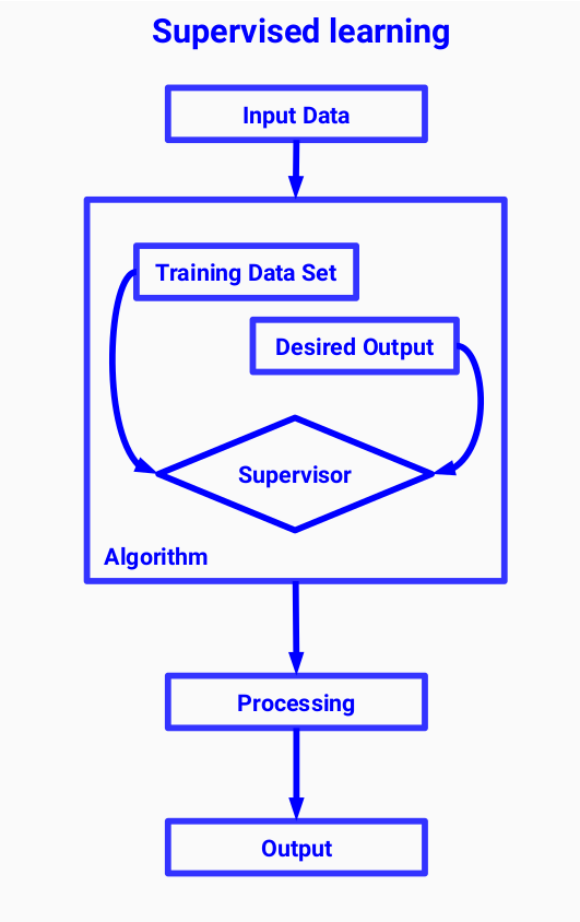
- RAPPRESENTAZIONE "INTUITIVA"
- IL SISTEMA AI ACQUISISCE LA PROPRIA CONOSCENZA



INTELLIGENZA ARTIFICIALE: ALGORITMI



ESTRAE E OTTIMIZZA
LE CARATTERISTICHE DEI DATI



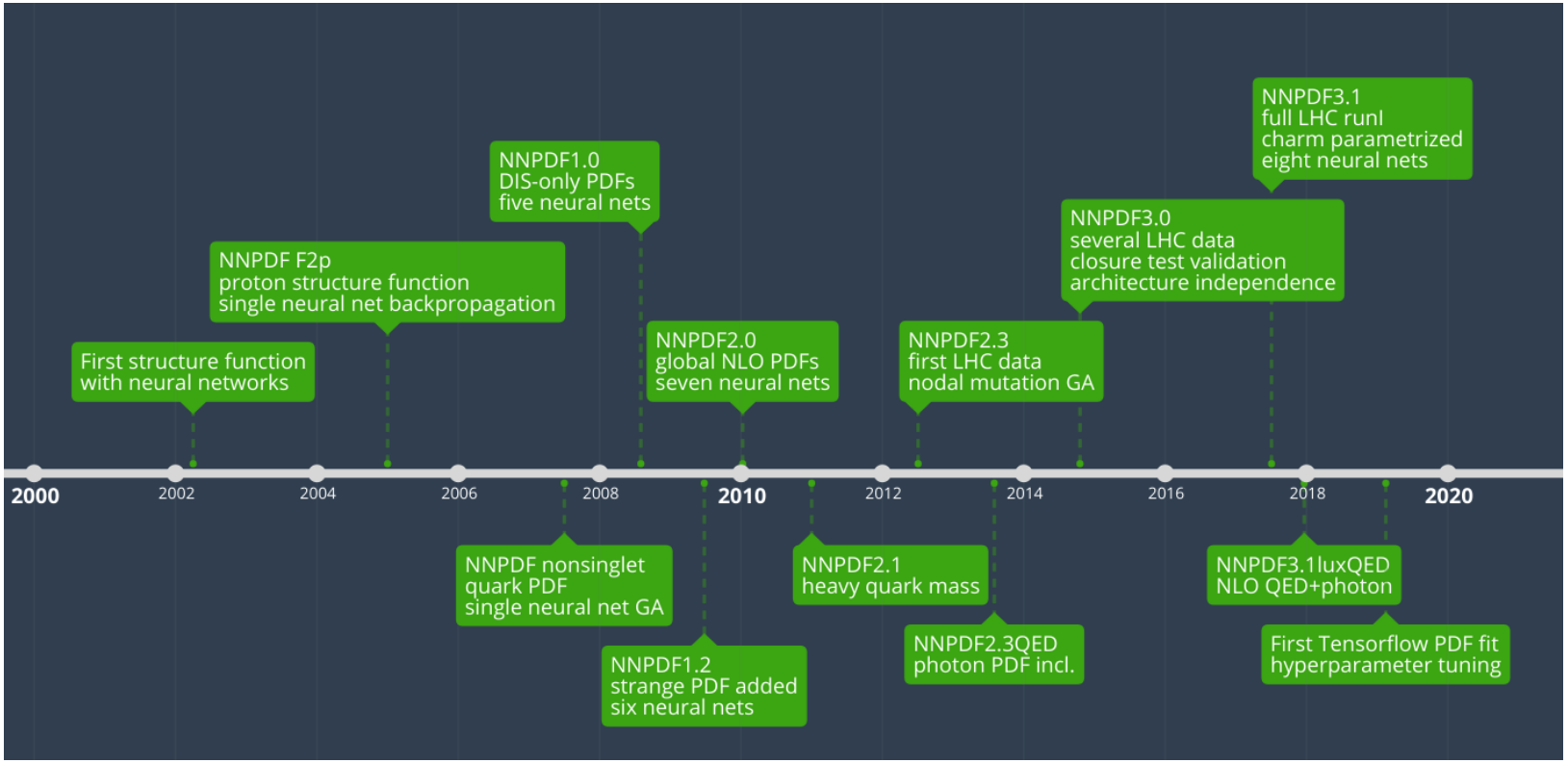
IMPARA DAI DATI OTTIMIZZANDO
UNA PROPRIETÀ



IMPARA DAI DATI LA STRATEGIA
DI APPRENDIMENTO

LA STRUTTURA DEL PROTONE COME PROBLEMA DI AI: NNPDF

(s.f., Latorre: 2002)



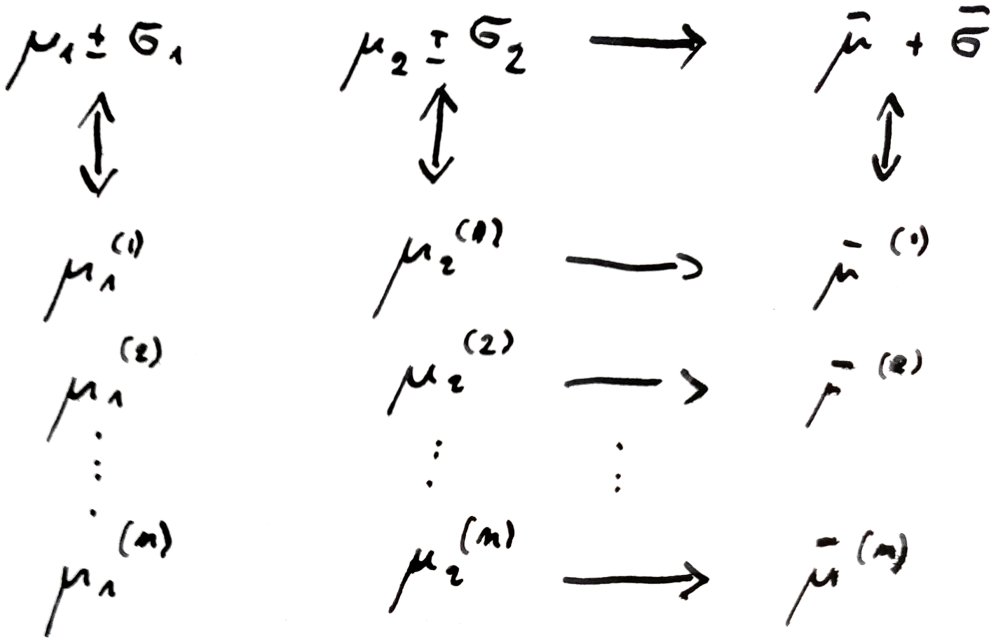
MONTECARLO + RETI NEURALI

COMBINAZIONE DI DATI CON IL METODO MONTECARLO

DUE DATI: $\mu_1 \pm \sigma_1$; $\mu_2 \pm \sigma_2$

COMBINAZIONE DI MASSIMA VEROSIMIGLIANZA: $\bar{\mu} \pm \bar{\sigma}$; $\bar{\mu} = \dfrac{\frac{\mu_1}{\sigma_1^2} + \frac{\mu_2}{\sigma_2^2}}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$; $\bar{\sigma}^2 = \dfrac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2}}$

RAPPRESENTAZIONE MONTECARLO

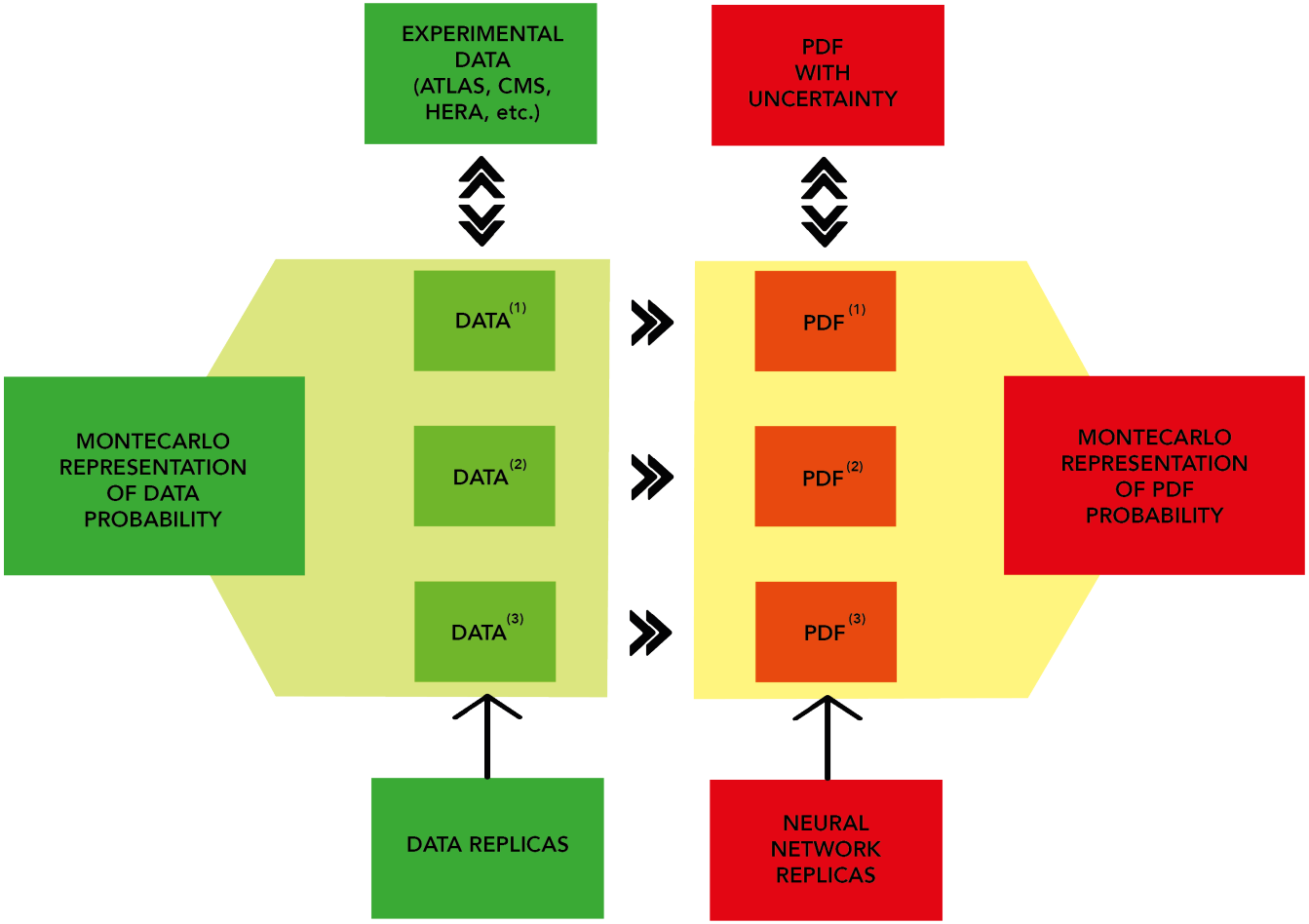


$\mu^{(i)} \Leftrightarrow$ REPLICHE $\Leftrightarrow$ DISTRIBUZIONE DI PROBABILITÀ

LA STRUTTURA DEL PROTONE COME PROBLEMA DI AI

IL MONTECARLO FUNZIONALE

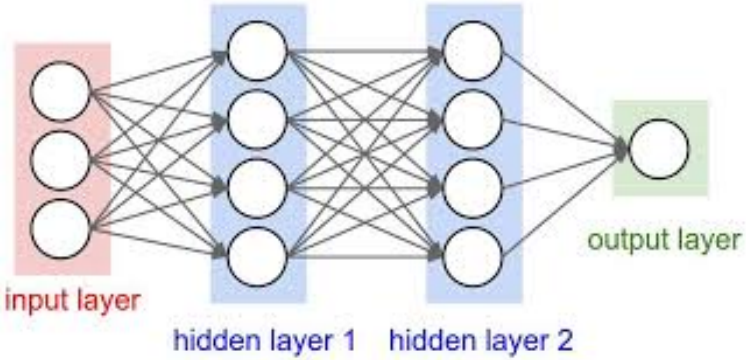
REPLICHE DI FUNZIONI $\Leftrightarrow$ PROBABILITÀ IN UNO SPAZIO DI FUNZIONI



PDF OUTPUT: $f_i^{(a)}(x,\mu)$; i =up, antiup, down, antidown, strange, antistrange, charm, gluone; $j = 1, 2, \ldots N_{\rm rep}$

LA RETE NEURALE COME AGENTE AI UNIVERSALE

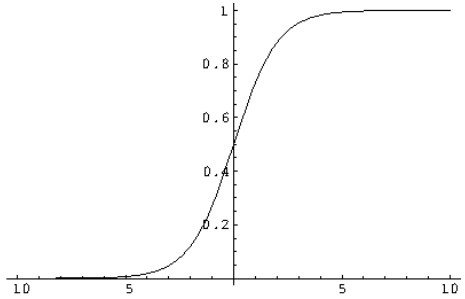
ARCHITETTURA



PARAMETERI

- PESI ω_{ij}
- SOGLIE θ_i

FUNZIONE DI ATTIVAZIONE



$$F^{(i)}_{\text{out}}(\vec{x}_{\text{in}}) = F\left(\sum_j \omega_{ij} x^j_{\text{in}} - \theta_i\right)$$

IL PIÙ SEMPLICE ESEMPIO

1-2-1

$$f(x) = \frac{1}{1+e^{\theta^{(3)}_1 - \frac{\omega^{(2)}_{11}}{1+e^{\theta^{(2)}_1 - x\omega^{(1)}_{11}}} - \frac{\omega^{(2)}_{12}}{1+e^{\theta^{(2)}_2 - x\omega^{(1)}_{21}}}}}$$

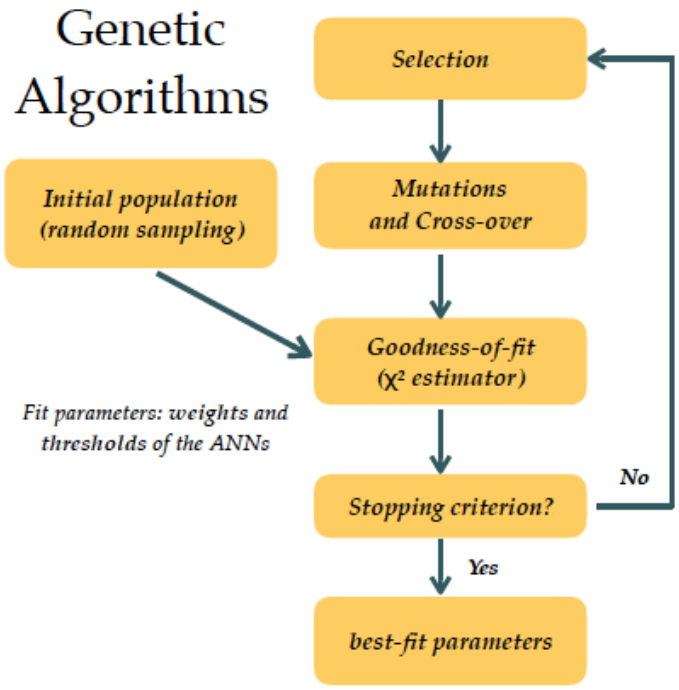
NNPDF: $2-5-3-1$ NN PER OGNI PDF: $37 \times 8 = 296$ PARAMETETRI

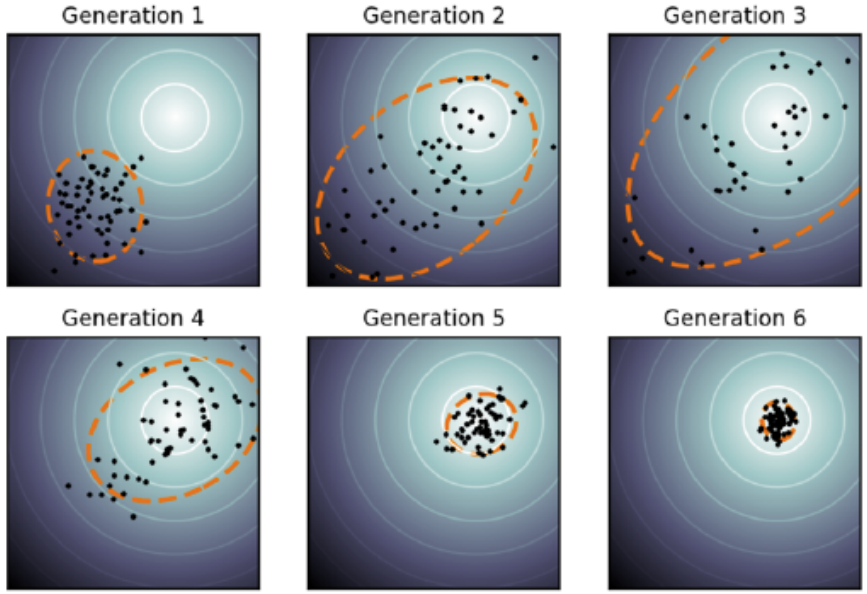
SUPERVISED LEARNING

L'ALGORITMO GENETICO

IDEA:

- MUTAZIONE CASUALE DEI PARAMETRI DELLA RETE
- SELEZIONE DEL PIÙ ADATTO

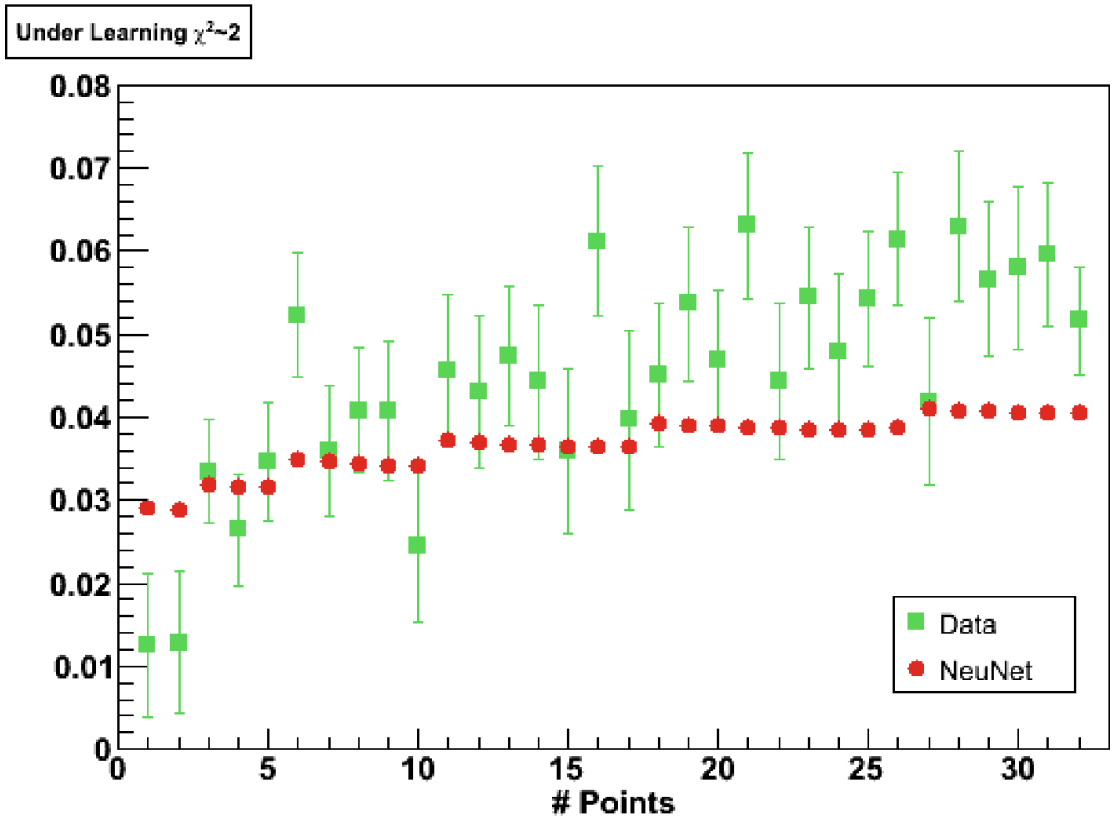




RETI NEURALI: APPRENDIMENTO DI UNA LEGGE

- LA COMPLESSITÀ AUMENTA DURANTE L'APPRENDIMENTO
- FINO A IMPARARE IL RUMORE
- QUANDO FERMARSI?

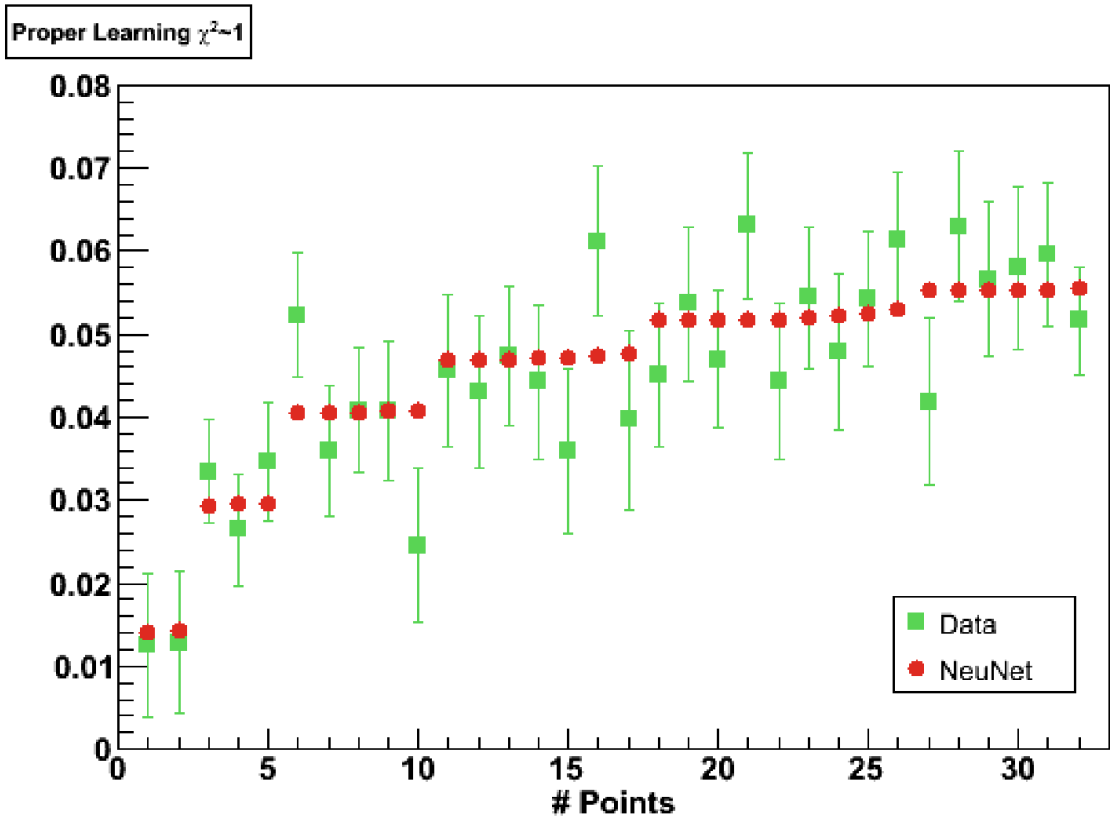
SOTTO-APPRENDIMENTO ("UNDERLEARNING")



RETI NEURALI: APPRENDIMENTO DI UNA LEGGE

- LA COMPLESSITÀ AUMENTA DURANTE L'APPRENDIMENTO
- FINO A IMPARARE IL RUMORE
- QUANDO FERMARSI?

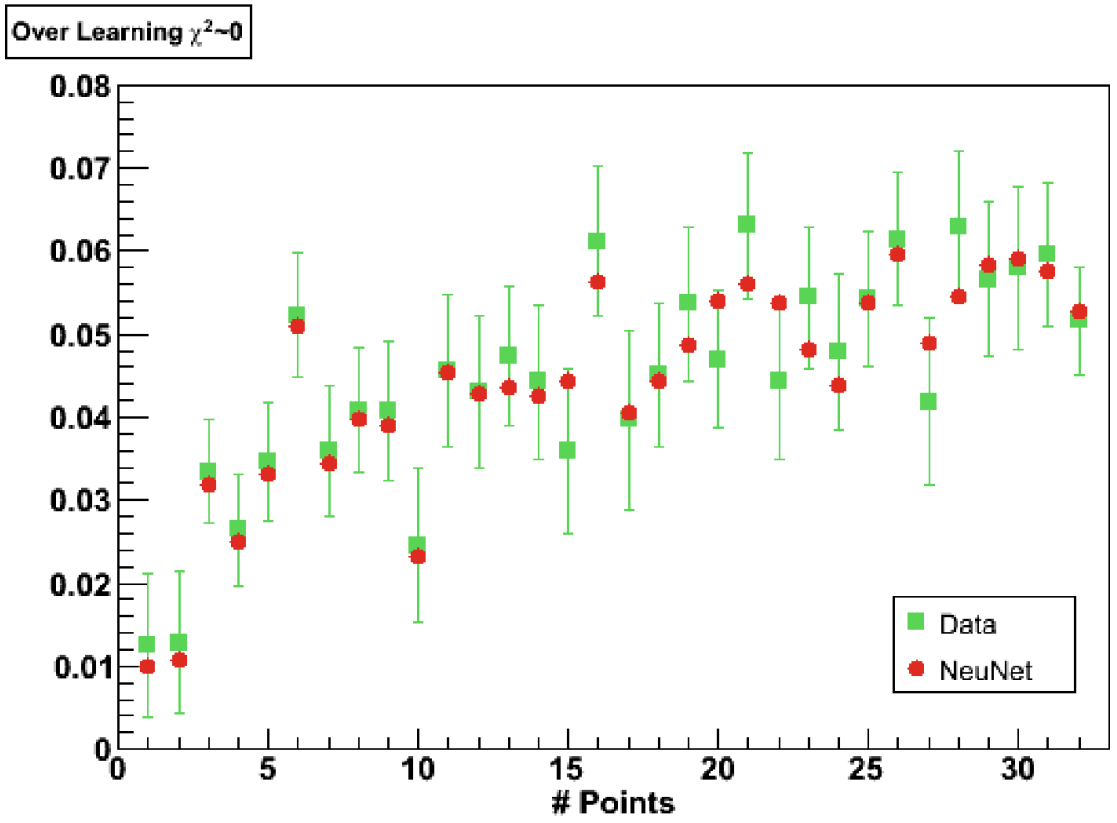
APPRENDIMENTO ("PROPER LEARNING")



RETI NEURALI: APPRENDIMENTO DI UNA LEGGE

- LA COMPLESSITÀ AUMENTA DURANTE L'APPRENDIMENTO
- FINO A IMPARARE IL RUMORE
- QUANDO FERMARSI?

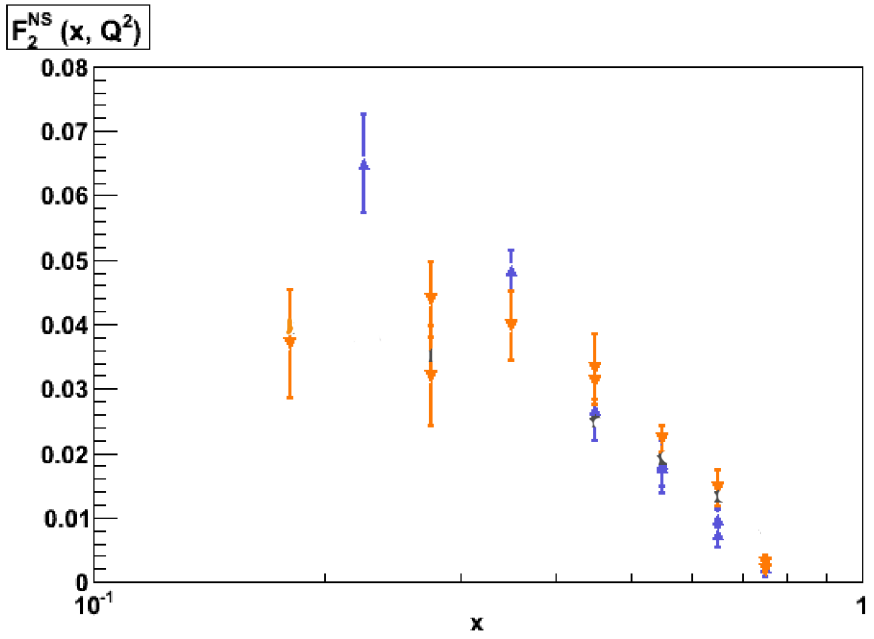
SOVRA-APPRENDIMENTO ("OVERLEARNING")



CONVALIDA INCROCIATA!

ALGORITMO GENETICO:
AD OGNI GENERAZIONE, IL COSTO (χ^2) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

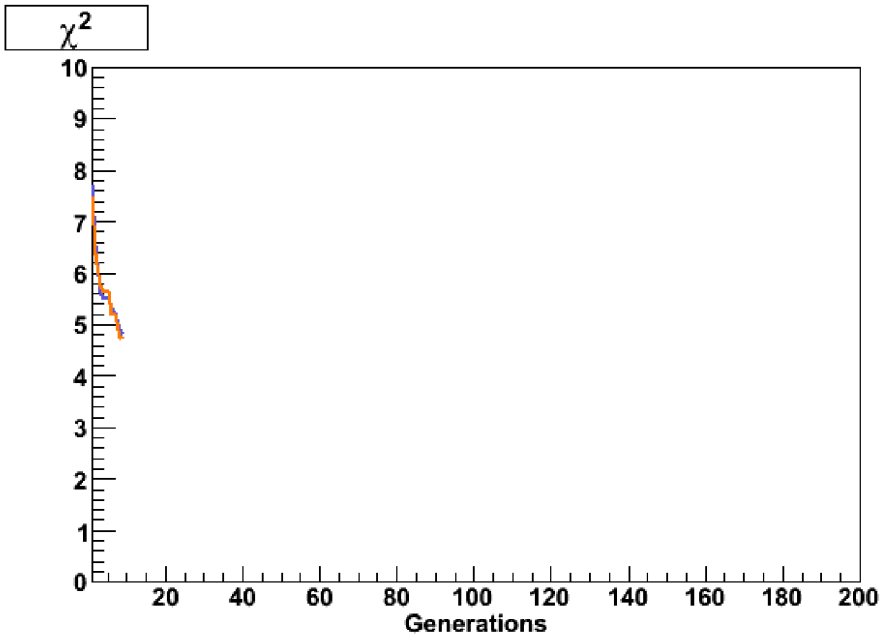


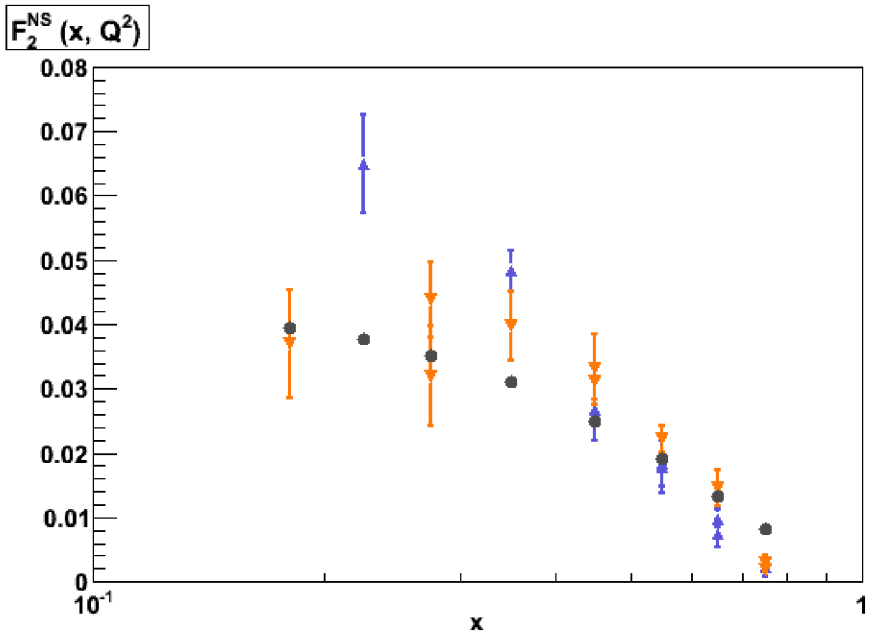
CONVALIDA INCROCIATA!

ALGORITMO GENETICO:
AD OGNI GENERAZIONE, IL COSTO (χ^2) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

GO!



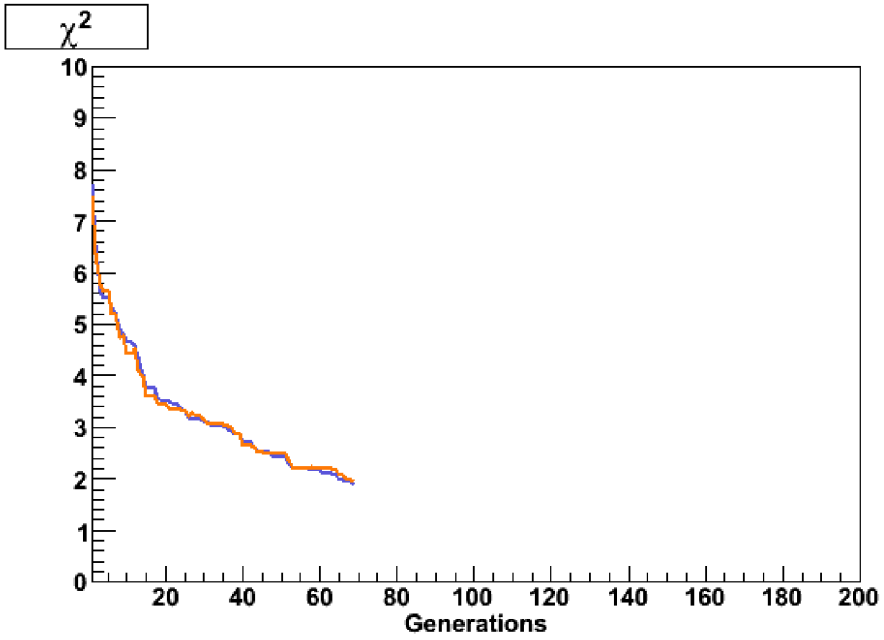


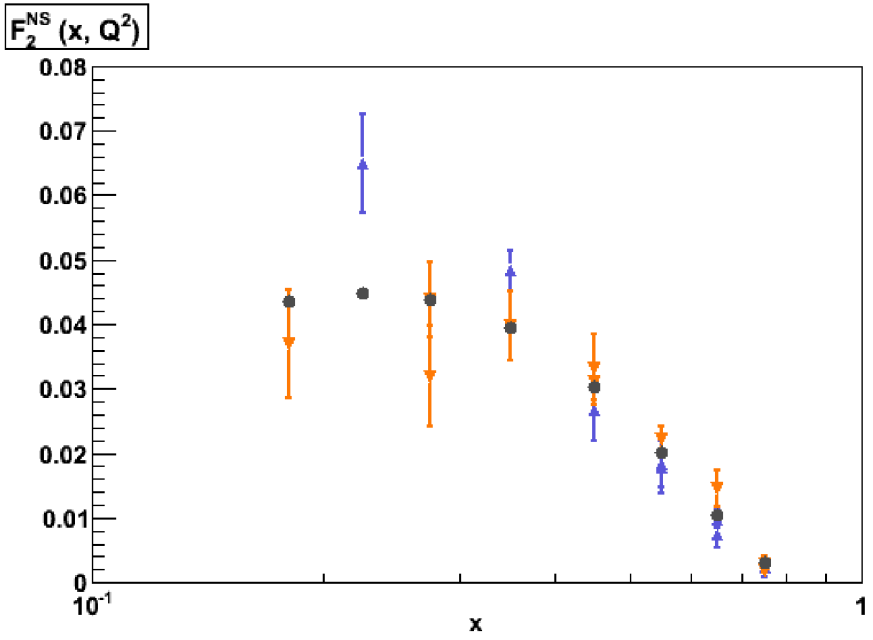
CONVALIDA INCROCIATA!

ALGORITMO GENETICO:
AD OGNI GENERAZIONE, IL COSTO (χ^2) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

STOP!



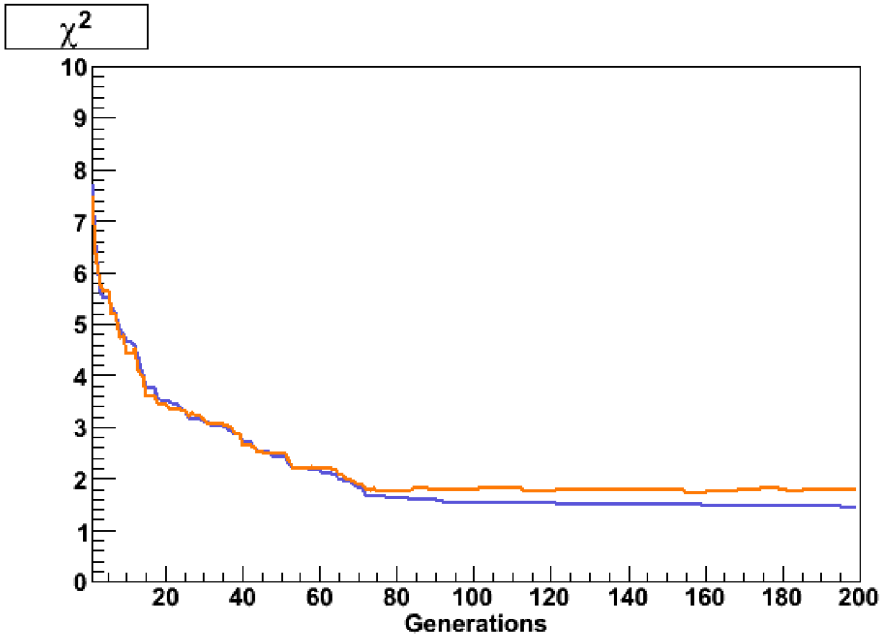


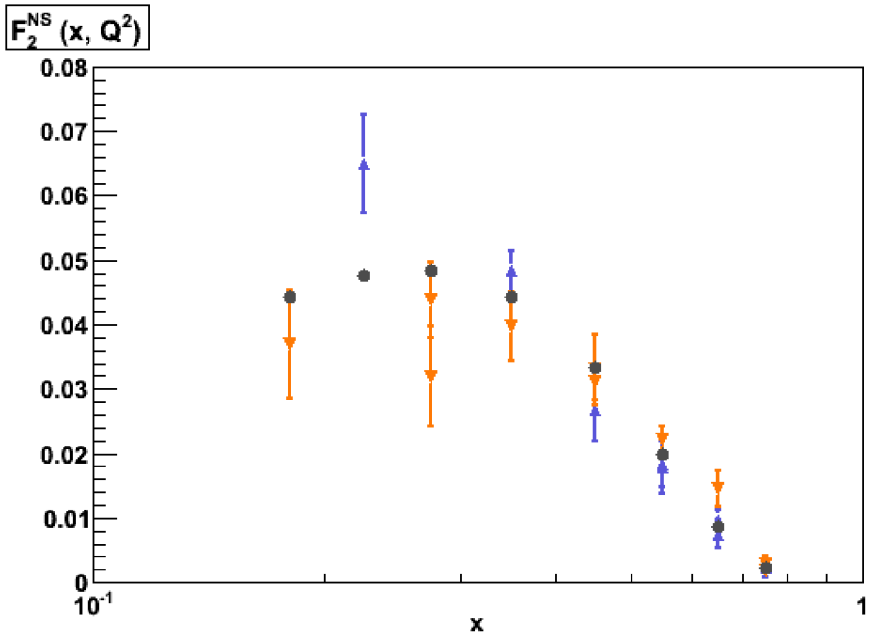
CONVALIDA INCROCIATA!

ALGORITMO GENETICO:
AD OGNI GENERAZIONE, IL COSTO (χ^2) SCENDE O RESTA INVARIATO

- DIVIDI L'INPUT IN DUE: TRAINING CONVALIDA
- MINIMIZZA IL COSTO DELL'INSIEME DI TRAINING
- AD OGNI GENERAZIONE CALCOLA IL COSTO DELL'INSIEME DI CONVALIDA (NON USATO PER LA MINIMIZZAZIONE)
- QUANDO IL COSTO DI CONVALIDA NON SCENDE PIÙ L'APPRENDIMENTO È COMPLETO

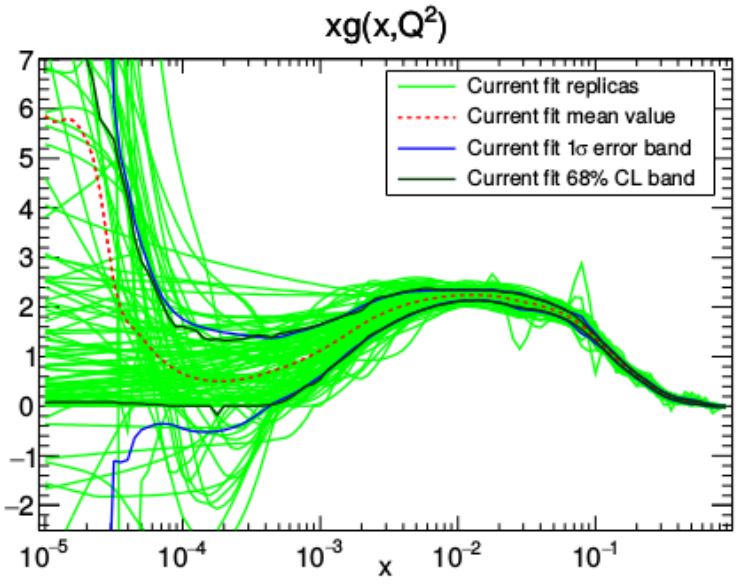
TROPPO TARDI!



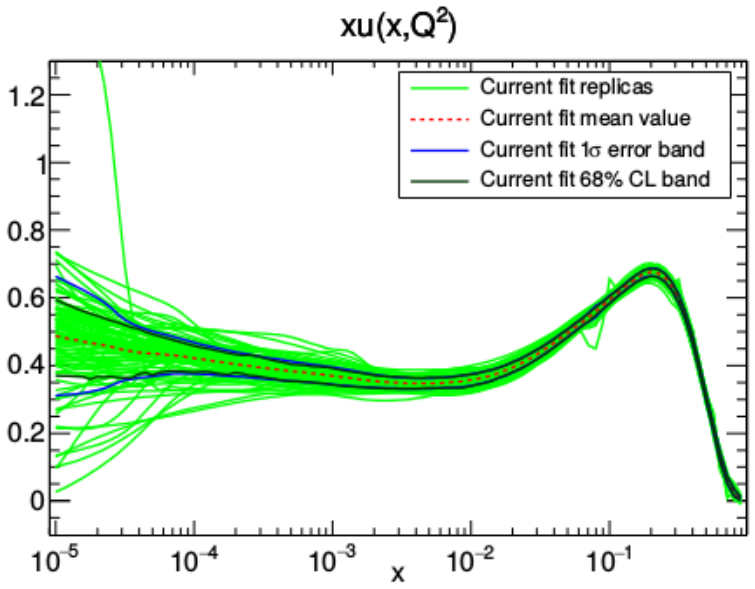


IL RISULTATO
IL RITRATTO DEL PROTONE (100 REPLICHE)

GLUON



UP



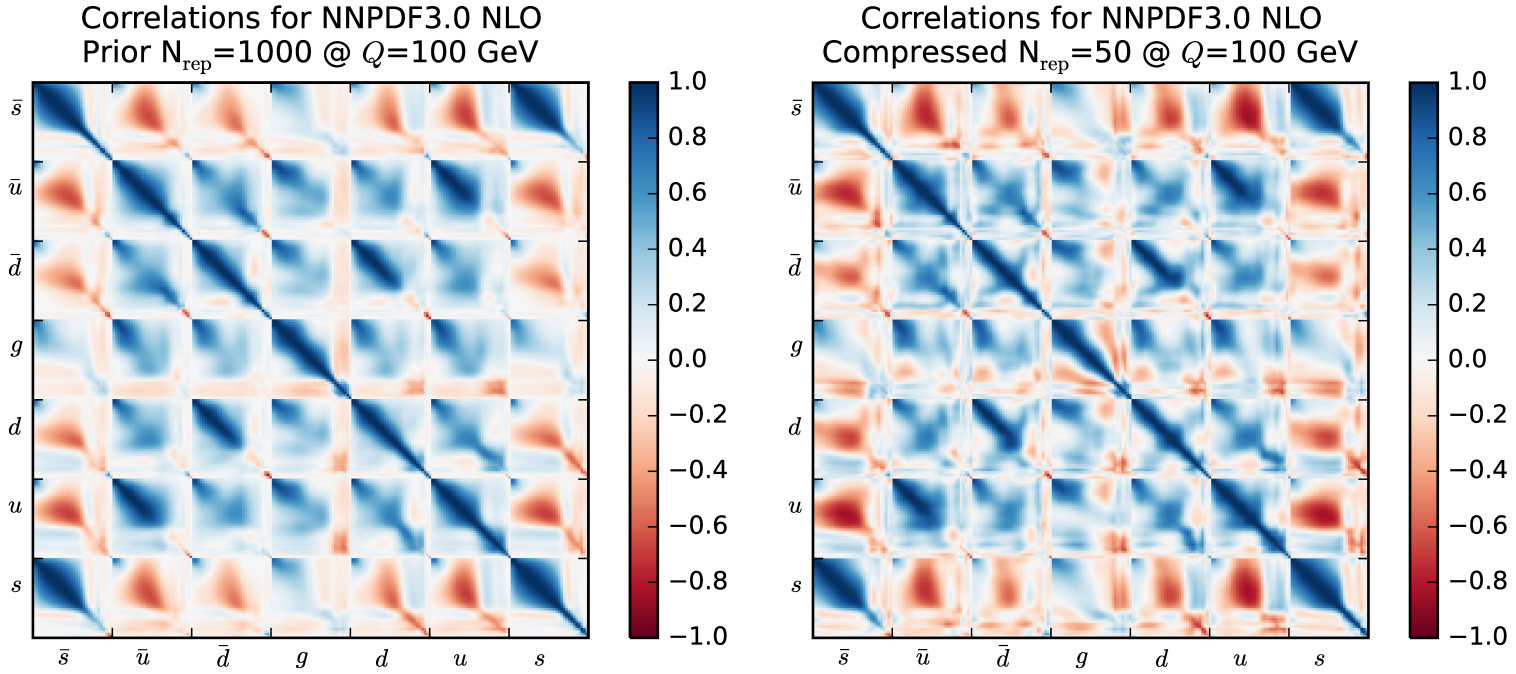
- DISTRIBUZIONE DI PROBABILITÀ ↔ REPLICHE
- RISULTATO PIÙ PROBABILE ↔ MEDIA
- INCERTEZZA ↔ DEVIAZIONE STANDARD
- OPPURE 68% C.L.
- MA BASTANO 100 REPLICHE?

UNSUPERVISED LEARNING

OTTIMIZZAZIONE

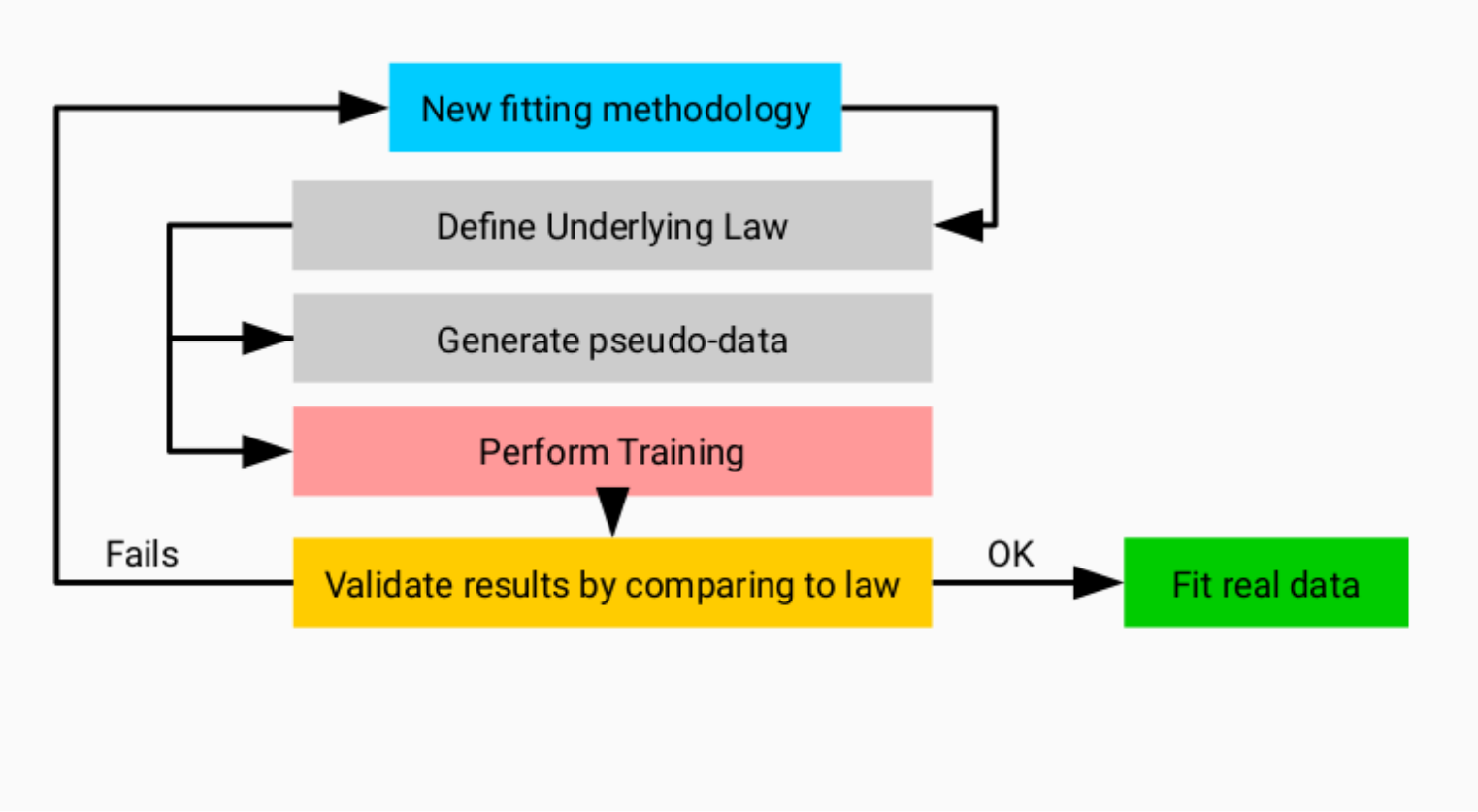
- COME MASSIMIZZARE L'ACCURATEZZA?
- INSIEME (PRIOR) DI MOLTE REPLICHE
- SELEZIONE GENETICA $\Rightarrow$ OTTIMIZZAZIONE INDICATORI STATISTICI (DIVERGENZA KULLBACK-LEIBLER)
- 50 REPLICHE OTTIMIZZATE $\Leftrightarrow$ 1000 REPLICHE DI PARTENZA

MATRICE DI CORRELAZIONE



COME FACCIAMO A SAPERE CHE IL RISULTATO È GIUSTO?

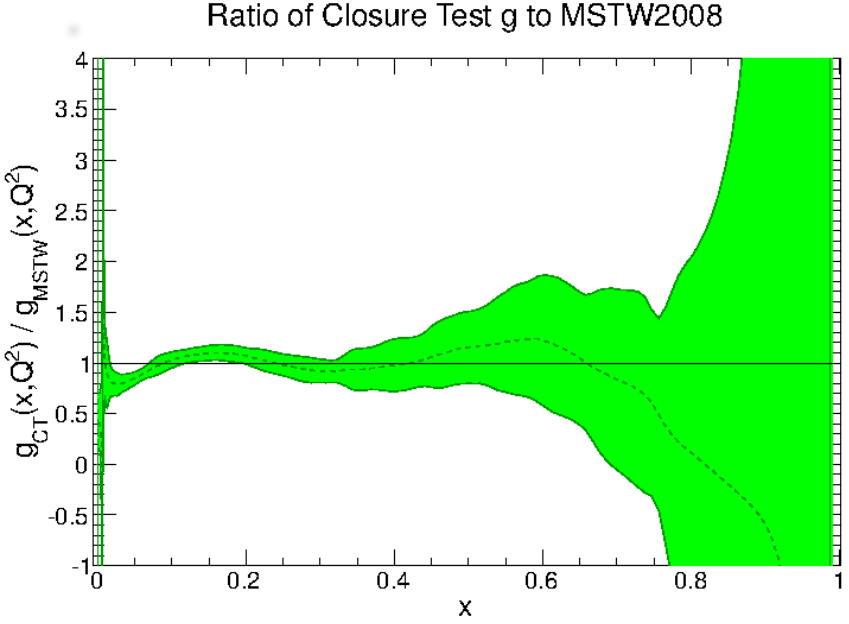
TEST DI CHIUSURA



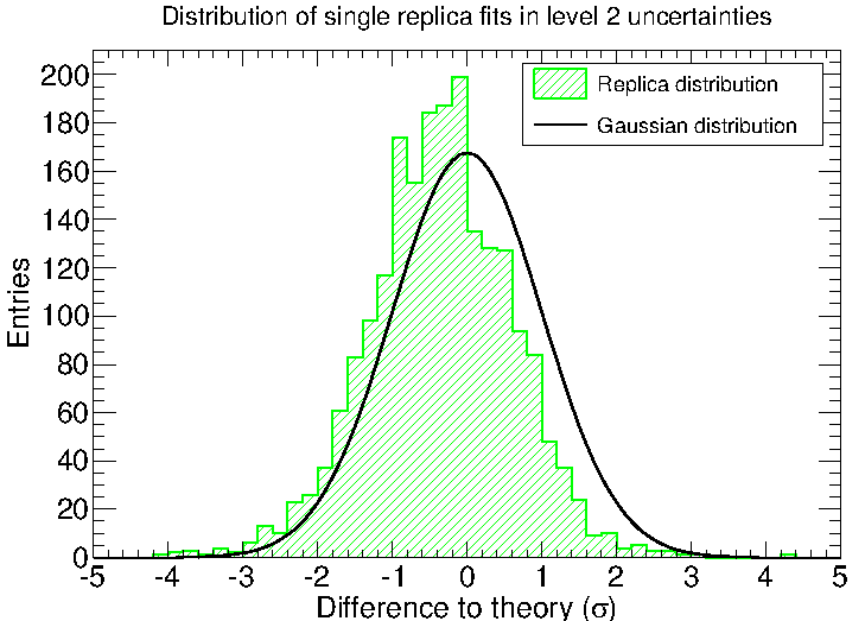
PRIMO TEST DI CHIUSURA

RISULTATI

IL GLUONE: RISULTATO/"VERITÀ"

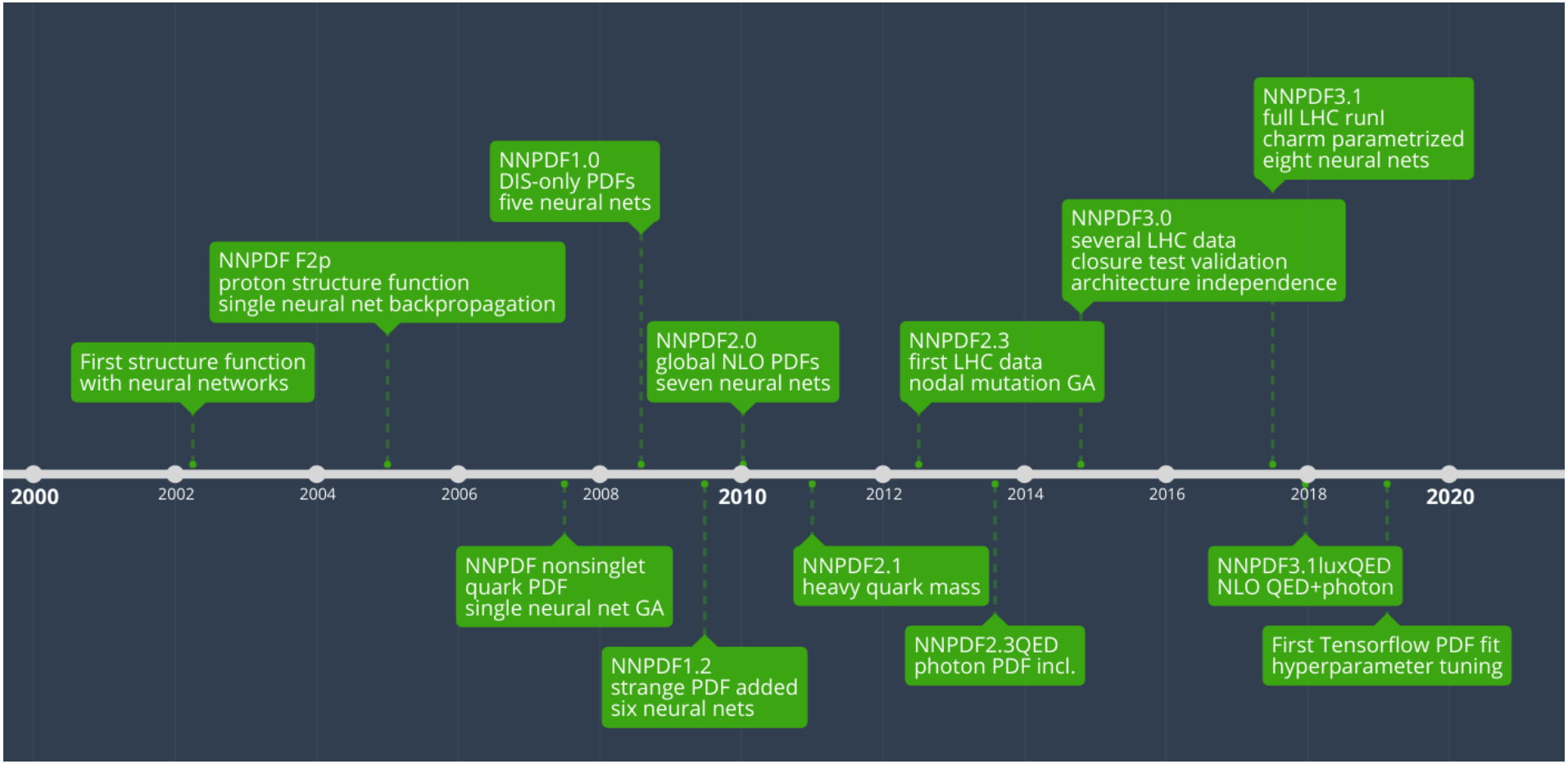


DISTRIBUZIONE NORMALE DELLE DEVIAZIONI



REINFORCEMENT LEARNING?

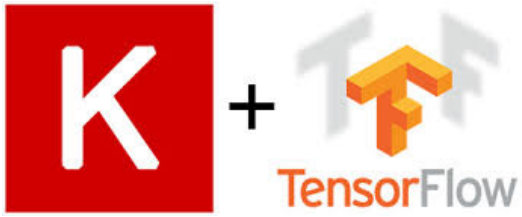
NNPDF: A CHE PUNTO SIAMO?



VERSO IL "DEEP LEARNING"

NUOVI STRUMENTI

- KERAS, TENSORFLOW,. . .
- LIBRERIE OPEN-SOURCE PER IL DEEP LEARNING, ELEVATA EFFICIENZA
- PROGRAMMAZIONE AD ALTO LIVELLO IN PYTHON



SET NNPDF: 15 ORE $\Rightarrow$ 10 MINUTI

OBIETTIVI:

- OTTIMIZZAZIONE DEI PARAMETRI
- "REINFORCEMENT LEARNING" AUTOMATICO DELLA METODOLGIA OTTIMALE

VERSO IL “DEEP LEARNING”

OTTIMIZZAZIONE DEI PARAMETRI



- VELOCITÀ $\Rightarrow$ POSSIBILE PROVARE STRATEGIE ALTERNATIVE
- SCAN DELLO SPAZIO DI PARAMETRI $\Rightarrow$ AUTOMATIZZAZIONE

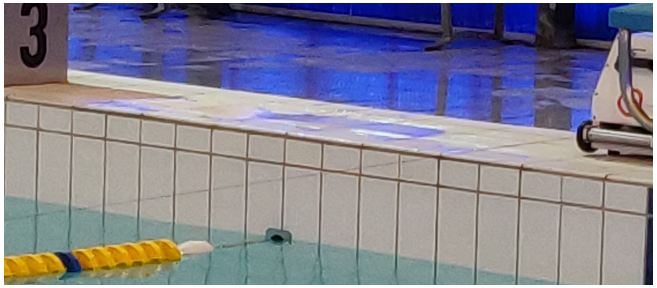
“DEEP LEARNING”

CHE COSA ABBIAMO IMPARATO?

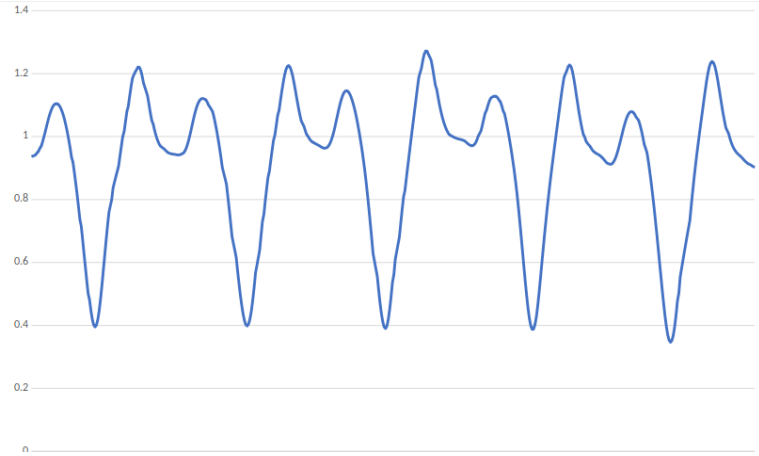
IMPARARE UNA LEGGE SCONOSCIUTA $\Leftrightarrow$ PROBABILITÀ

IL PROGETTO NUOTO









IMPARARE LA QUALITÀ DI UN NUOTATORE DAL PROFILO DI VELOCITÀ

$\Rightarrow$ STRATEGIA DI ALLENAMENTO OTTIMIZZATA

UN LAVORO DI MOLTI!

LA COLLABORAZIONE NNPDF



IL TEAM N^{3}PDF

Machine Learning • PDFs • QCD

European Research Council
Established by the European Commission





