Gaussian Process for the estimation of theory uncertainties

Jesús Urtasun Elizari

Italian Physical Society - Milan, September 2020

UNIVERSITÀ
DEGLI STUDI
DI MILANO

N3PDF
Machine Learning • PDFs • QCD

erc
European
Research
Council

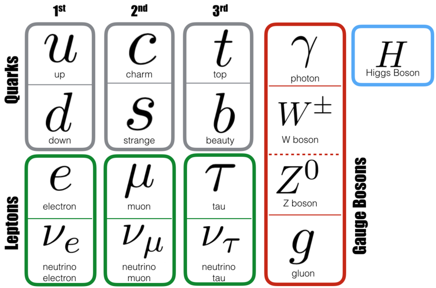
Outline

1. QCD in a nutshell
   - The Standard Model & strong interactions
   - Parton Distribution Functions
   - Factorization theorem
2. Machine Learning
   - Motivation for Machine Learning
   - Neural Networks & general strategy
3. Gaussian Process and theory uncertainties
   - The NNPDF methodology
   - Operator implementation in TensorFlow
   - Results & Conclusions

Quantum Chromodynamics in a nutshell

QCD in a nutshell

The Standard Model



Quantum Field Theory describing physics at the TeV scale

1. Fermions composing matter
2. Bosons mediating interactions
3. Scalar Higgs generating mass

QCD in a nutshell

Explore the strong interactions

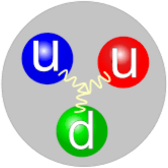
How to explore proton's inner structure?

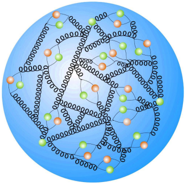
- Point-like projectile on the object $\longrightarrow$ DIS
- Smash the two objects $\longrightarrow$ LHC physics

"A way to analyze high energy collisions is to consider any hadron as a composition of point-like constituents $\longrightarrow$ **partons"** R.Feynman, 1969

QCD in a nutshell

Parton Distribution Functions





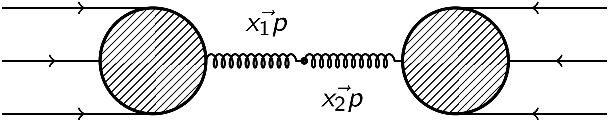
- Hadrons made of partonic objects $\longrightarrow$ non perturbative physics
- Interactions take place only at partonic level

Parton Distribution Functions: probability distribution of finding a particular parton (u, d, ..., g) carrying a fraction x of the proton's momentum

QCD in a nutshell

Factorization theorem

Observables in hadronic events $\longrightarrow$ σ is hard to compute

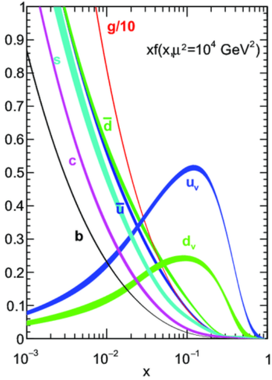


Factorize the problem $\longrightarrow$ Convolute the PDFs with the partonic $\hat{\sigma}_{ij}$

$$\sigma = \int_0^1 dx_1 \, dx_2 \, f_\alpha(x_1, \mu_F) * f_\beta(x_2, \mu_F) * \hat{\sigma}_{\alpha\beta}(\alpha_s(\mu_R), \mu_F)$$

- Partonic $\hat{\sigma}$ can be computed as perturbative series in α_s
- PDFs absorb the non perturbative effects, evaluated at μ_F

- Each parton has a different PDF $u(x), d(x), ..., g(x)$
- PDFs are not predicted, and can not be measured
- PDFs are extracted from data



Machine Learning

Machine learning

What is Machine Learning?

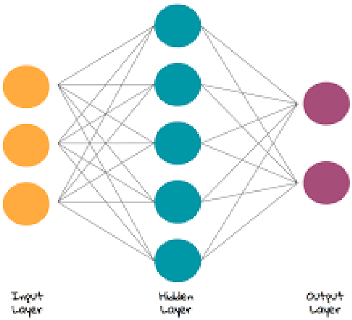
1. A subset of Artificial Intelligence (AI) algorithms
2. Used to solve *complex* tasks like classification and regression
3. Rely on comparison with data $\longrightarrow$ Learning

Machine learning

Building a ML model

Machine learning

Building a Ml model

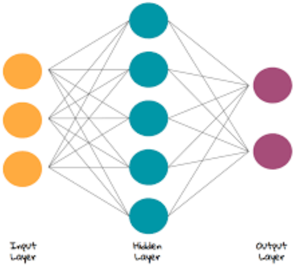


1. Turn the input set into an array and build a random prediction
2. Compare with truth and compute a Loss function
3. Update the parameters in a specific way

Machine learning

Building a ML model

$\longrightarrow$

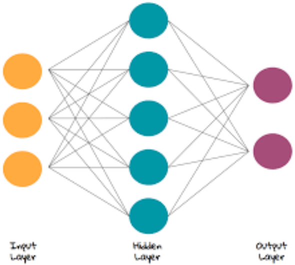


$\longrightarrow$ Tree ✓

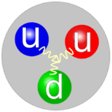
Machine learning

Building a ML model

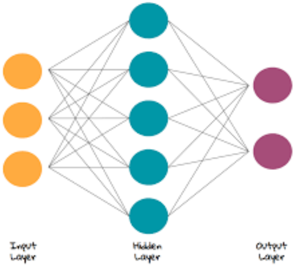
$\longrightarrow$



$\longrightarrow$ Tree ✓



$\longrightarrow$

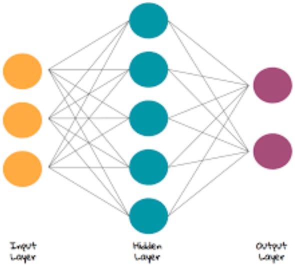


$\longrightarrow$ Not tree X

Machine learning

Building a ML model

$\longrightarrow$



$\longrightarrow$ Tree ✓

$$Loss = \begin{Bmatrix} y_1^{true} \\ y_2^{true} \\ \vdots \\ y_n^{true} \end{Bmatrix} - \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{Bmatrix}$$

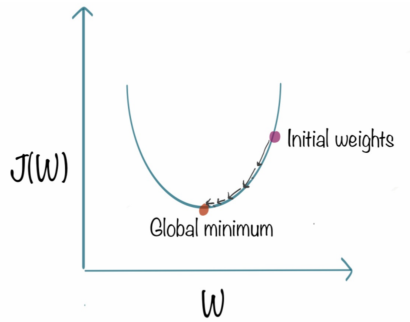
$$\underset{y_{true}}{} \qquad\qquad \underset{y_{pred}}{}$$

1. Compute a loss function

$$L = \sum_{i=1}^{N} \left(y_i^{true} - y_i^{pred}\right)^2$$

2. Perfect prediction will mean $L = 0$

1. Loss is a function of weights and bias $L = L(w, b)$
2. Compute gradient $\nabla_{w_{ij}} L$ to look for the minimum of L

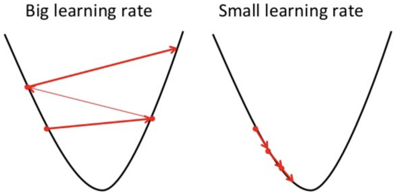


Update rule

Update the parameters of the network following the gradient descend

$$w_{ij} \longrightarrow w_{ij} - \alpha \nabla L$$
$$b_i \longrightarrow b_i - \alpha \nabla L$$

Where α is the learning rate



The N3PDF project

Machine Learning for the precision determination of PDFs

Factorize the problem $\longrightarrow$ Convolute the PDFs with the partonic $\hat{\sigma}_{ij}$

$$\sigma = \int_0^1 dx_1\, dx_2\, f_\alpha(x_1, \mu_F) * f_\beta(x_2, \mu_F) * \hat{\sigma}_{\alpha\beta}(\alpha_s(\mu_R), \mu_F)$$

- Partonic $\hat{\sigma}_{\alpha\beta}$ is computed perturbatively. Hadronic σ is measured.
- Use a Neural Networks to generate (*fit*) the PDFs
- Generate a vector of observables σ_N to be compared with data

$$\sigma_N = \sum_{i,j,\alpha,\beta} f_\alpha(x_i)\, f_\beta(x_j)\, \hat{\sigma}_{Nij\alpha\beta}$$

The N3PDF project

General structure of n3fit

- Use TensorFlow and Keras to determine the PDFs
- See paper by S.Carraza - J.Cruz-Martinez
  "Towards a new generation of parton densities
  with deep learning models",
  https://arxiv.org/abs/1907.05075

The N3PDF project

General structure of n3fit

σ_{pred}

$$\left\{\begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix}\right\} \longrightarrow \boxed{\text{PDF}} \longrightarrow \left\{\begin{matrix} f_{1\alpha} \\ f_{2\alpha} \\ \vdots \\ f_{n\alpha} \end{matrix}\right\} \longrightarrow \boxed{\text{Convolution}} \longrightarrow \left\{\begin{matrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_n \end{matrix}\right\}$$

x_{grid}

1. Build a NN to compute σ_{pred} observables from a grid x_i
2. Compute χ^2 loss function by comparing with data
3. Update values of PDF $\longrightarrow$ Fit

Summary & Conclusions

1. PDFs are required to have accurate predictions in high energy physics
2. ML provides a new way of determine the PDFs
3. Operator implementation leads to memory saving by taking full control on the computation

Thank you!

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