Machine Learning for the precision determination of Parton Distribution Functions

Jesus Urtasun Elizari

PhD Seminar, Milan, September 2019

UNIVERSITÀ DEGLI STUDI DI MILANO

N3PDF
Machine Learning • PDFs • QCD

European Research Council

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

Outline

1. Machine Learning, an introduction
2. Quantum Chromodynamics in a nutshell
   - The Standard Model
   - Parton Distribution Functions
   - Factorization theorem
3. The N3PDF project
   - Motivation for PDFs determination.
   - Operator implementation in TensorFlow.
   - Results & Conclusions.

Machine Learning, an introduction

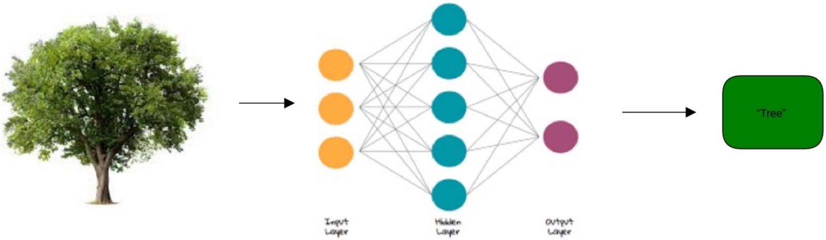
Machine Learning

An example: Image recognition

Identify a particular object $\longrightarrow$ hard problem for computers

Machine Learning

Introducing neural networks



1. Build a model that takes that array x and outputs $y(x)$ such that for a tree outputs "tree"
2. $y(x) = \sigma\{\mathbf{w} \cdot x + b\} \longrightarrow$ Need for training $\mathbf{w}$, b

1. Multiply by a matrix of random weights and apply an activation function σ

$$y(x) = \sigma\left\{ \begin{pmatrix} w_{00} & \dots & w_{0n} \\ \vdots & \ddots & \vdots \\ w_{n0} & \dots & w_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \right\}$$

2. First prediction will not be correct $\longrightarrow$ Compute a loss function by comparing with truth

$$L = \sum_{i=1}^{N} (y(x) - y_i^{true})^2$$

1. A perfect prediction will mean $L = 0$ Notice $L = L(w_{ij}, b_i)$

$$\nabla_w L = \frac{\partial L}{\partial w_{ij}}$$
$$\nabla_b L = \frac{\partial L}{\partial b_i}$$

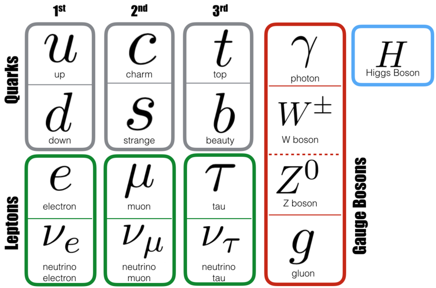
2. Update w and b such that they minimize the loss function

$$w_{ij} \to w_{ij} + \alpha \nabla_w L$$
$$b_i \to b_i + \alpha \nabla_b L$$

Quantum Chromodynamics in a nutshell

Quantum Chromodynamics

The Standard Model, a quick review



Quantum Field Theory describing physics at the TeV scale

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

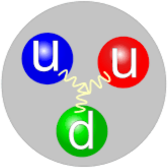
Quantum Chromodynamics

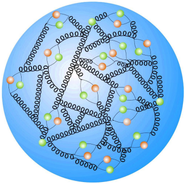
Explore the strong interactions

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





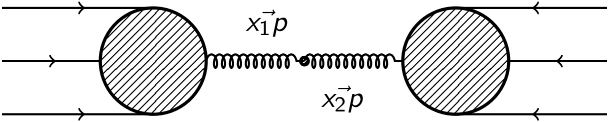
- Hadrons made of partonic objects $\longrightarrow$ non perturbative physics
- Interactions take place only at partonic level

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

Quantum Chromodynamics

Factorization theorem

Compute the cross section σ of a given process



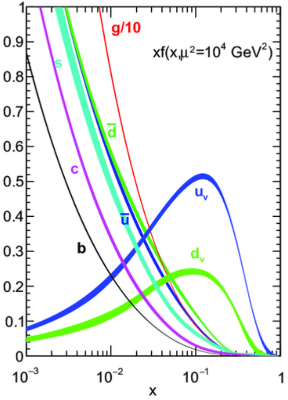
Factorize the problem into the partonic $\hat{\sigma}_{ij}$ and the PDFs at scale μ_F.

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

- Non perturbative physics hide inside the PDFs
- Partonic $\hat{\sigma}_{ij}$ can be written as a series in perturbative expansion in α_s

Quantum Chromodynamics

What PDFs look like



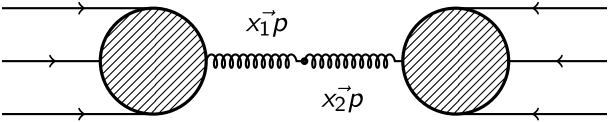
- Each parton has a different PDF $\longrightarrow u(x), d(x), ..., g(x)$.
- PDFs are not predicted, and can not be measured.
- PDFs are extracted from data.

The N3PDF project

Machine Learning for the precision determination of PDFs

The N3PDF project

What we actually measure



Convolute the partonic $\hat{\sigma}_{ij}$ with the PDFs $\longrightarrow$ Observable σ.

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

In our language $\longrightarrow$ vector of observables from a grid of x_i:

$$y_N = \sum_{i,j,\alpha,\beta} f_\alpha(x_i) f_\beta(x_j) FK_{Nij\alpha\beta}\ .$$

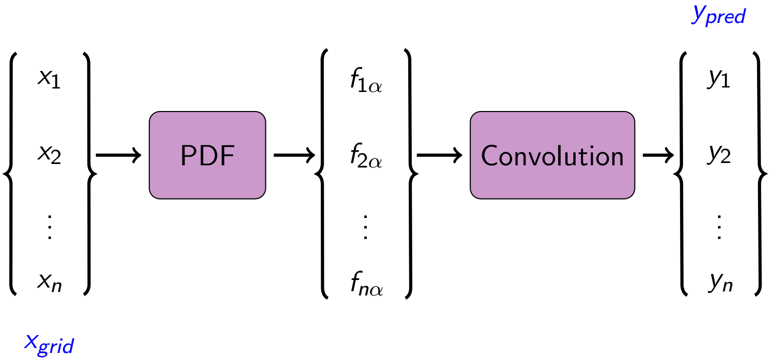
The N3PDF project

General structure of n3fit

- Use TensorFlow and Keras to determine 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

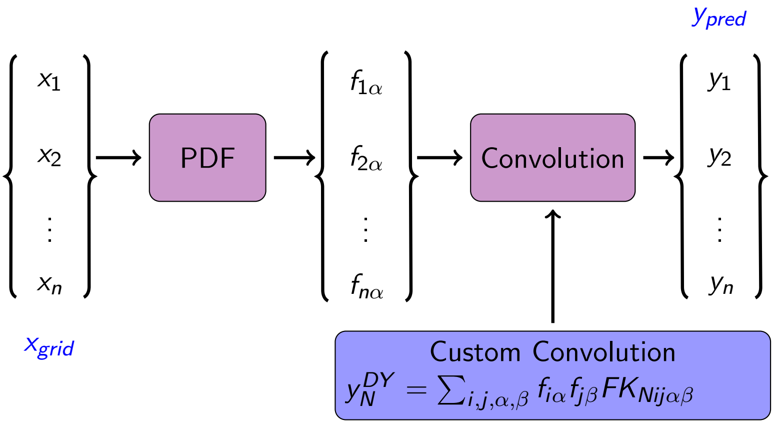
General structure



1. Build a NN to compute y_{pred} observables from a grid x_i
2. Compute gradient $\nabla\chi^2$ by comparing with data
3. Update values of PDF $\longrightarrow$ Fit

The N3PDF project

Operator implementation



1. TF relies in symbolic computation $\longrightarrow$ High memory usage
2. Implement c++ operator replacing the convolution

DIS only:

	TensorFlow	Custom	Ratio
Convolution	1.9207904	1.9207904	1.0000000
	2.4611666	2.4611664	0.9999999
	1.3516952	1.3516952	1.0000000
Gradient	1.8794115	1.8794115	1.0000000
	1.505316	1.505316	1.0000000
	2.866085	2.866085	1.0000000

Results

Checking computation

Hadronic:

	TensorFlow	Custom	Ratio
Convolution	8.142365	8.142366	1.0000001
	8.947762	8.947762	1.0000000
	7.4513326	7.4513316	0.9999999
Gradient	18.525095	18.525095	1.0000000
	19.182995	19.182993	0.9999999
	19.551006	19.551004	0.9999999

Results

Memory saving

Hadronic only:

	TensorFlow	Custom Convolution	Diff
Virtual	17.7 GB	13.8 GB	3.9 GB
RES	12.1 GB	8.39 GB	3.2 GB

Global:

	TensorFlow	Custom Convolution	Diff
Virtual	23.5 GB	19.7 GB	3.8 GB
RES	18.4 GB	12.5 GB	5.9 GB

Summary & Conclusions

1. PDFs are a crucial object in high energy physics.
2. ML provides a new way of precisely determine PDFs.
3. Operator implementation leads to saving memory 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.