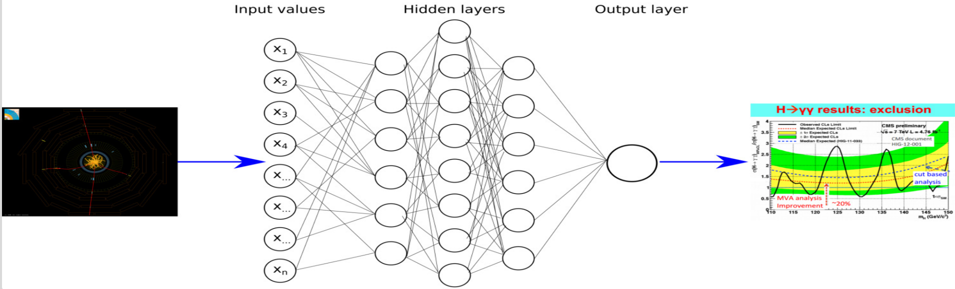
Machine learning in Higgs analyses

Higgs Couplings 2017

Marco A. Harrendorf on behalf of the CMS collaboration | October 9th, 2017

INSTITUT FÜR EXPERIMENTELLE TEILCHENPHYSIK (ETP)



From AlphaGo to AlphaGo Zero[1]

The game Go

Go being complexer than chess: $10^{10^{48-170}}$ vs. 10^{43-50} possible positions
$\Rightarrow$ Excellent playing field for machine learning

Difference between AlphaGo and AlphaZero

- AlphaGo: Supervised learning from human expert moves and reinforcement learning from self-play
- AlphaGo Zero: Solely based on reinforcement learning and knowledge of game rules

[1] Nature 550 (2017) 354, published on 19th of October

AlphaGo Zero: Learning and playing superhuman moves

Discovery of new corner sequences (joseki) by neural network

- Corner sequences (joseki): Important move sequences in the opening phase of Go
- AlphaGo Zero: Learned the already known corner sequences, but later discovered and played a new variation.

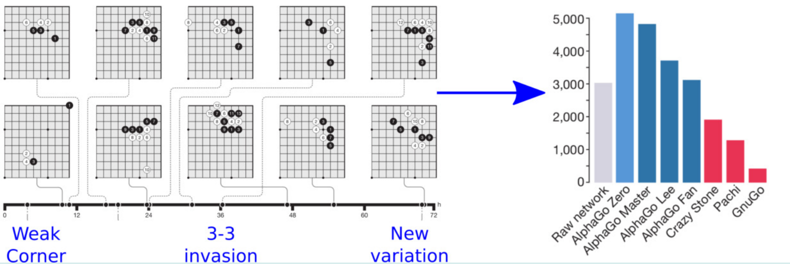
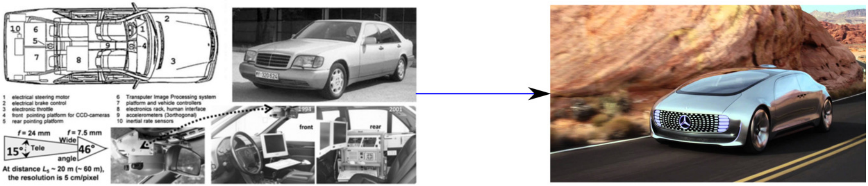


Fig.: Adopted from Nature Vol. 550

Further examples for late-breaking applications of neural networks



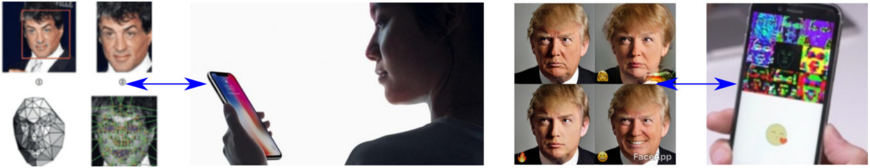
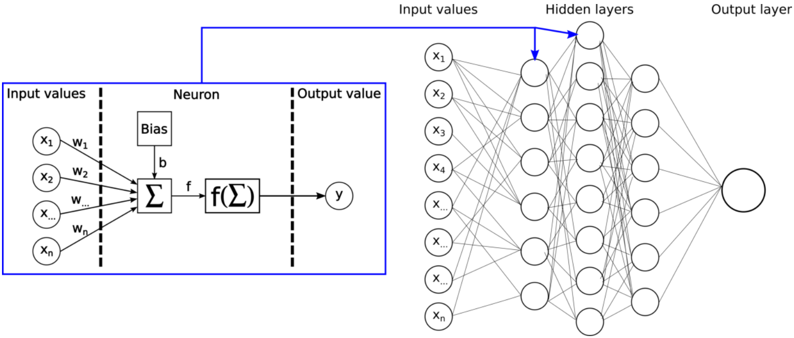


Fig.: Created using pictures from Heise, Mercedes Benz, Wikipedia

Neural networks in a nutshell

Artificial neuron

1. Input values x_i: Multiplied by weights w obtained from a training algorithm
2. Summing up all inputs and adding / subtracting additional bias b
3. Total sum Σ is modified by transfer function f to derive output value y



Foreword I: Arrival of neural networks

Machine learning applied in particle physics for a long time

- For example: Boosted decision trees and similar methods are and were used for many LHC Run I and Run II Higgs analyses [a]
- Also, neural networks were already tried in the 1990s[b]

---

[a] e.g. CMS $t\bar{t}H(bb)$ analysis, EPJC 75 (2015) 251
[b] e.g. JetNet package, Comp. Phys. Com. 81

Neural networks gaining traction by recent events

- Due to availability of GPU computing and large RAM ressources: Usage of elaborated neural networks becomes feasible
- Major step forward: Release of Tensorflow by Google in Nov. 2015
- Since this year: Tensorflow part of CMS software framework
  $\Rightarrow$ Many upcoming analysis with NN-methods in the next months, but only few published analyses so far.

Foreword II: Paradigm shift

So far: Human mind developing algorithms

Physics principles and mathematical laws together with physicists understanding of system considered: Basis for algorithm development
$\Rightarrow$ Algorithm as mapped thought process of physicist using known (and hopefully understood) observables

From now on: Machines developing algorithms

- Machines can derive algorithms on their own based on simple rules and principles but without the prior knowledge / bias of a physicist
- Machine algorithms exploit data deeper than physicist's algorithms
- Algorithms derived by machine can be a black box

$\Rightarrow$ Fundamental change: Superhuman insights require human's trust in algorithms obtained through machine learning
$\Rightarrow$ Establish new coping strategies and reliability measures

DeepCSV: Example of heavy flavour tagging

(Heavy) jet flavour tagging

- Trying to identify jets stemming from heavy flavour quarks (b, c) vs. jets from light quarks and gluons (u, d, s, g)
  $\Rightarrow$ Multiclassification problem
- Important for H$\rightarrow$ bb signal and various SM and BSM analyses with $t\bar{t}$ backgrounds (branching ratio (t$\rightarrow$ Wb)$\approx$ 100%)
  $\Rightarrow$ Many Higgs searches / analyses

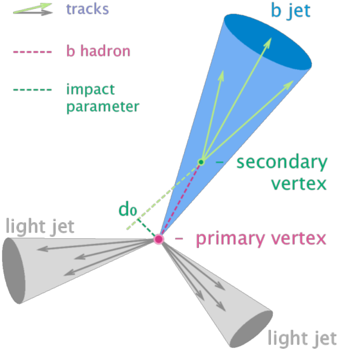


Fig.: Taken from Wikipedia

DeepCSV: Comparison with CSVv2 tagger

Karlsruher Institut für Technologie

CMS default b-tagging algorithm: CSVv2 (CMS DP-2017/012)

- Combination of secondary vertex and track-based lifetime information
- Updated version of Run I algorithm: Now combining the two sets of information with shallow neural networks instead of likelihood ratio
- Uses higher level features like masses of vertices and relatively raw information like significance of impact parameter per track.

New CMS algorithm: DeepCSV (CMS DP-2017/005)

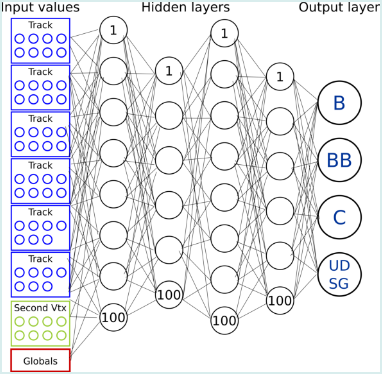
- Based on same set of variables as CSVv2 tagger, but using more charged particle tracks
- Based on “deep”[a] neural network with 4 hidden layers containing 100 neurons each (see next slide)

[a] NB: The concept of deep neural networks is not well-defined

DeepCSV: NN representation

66 inputs for neural network

- 6 strongly preselected charged particle tracks with 8 (7) properties
- 1 selected secondary vertex with 8 properties
- 12 global per-jet variables

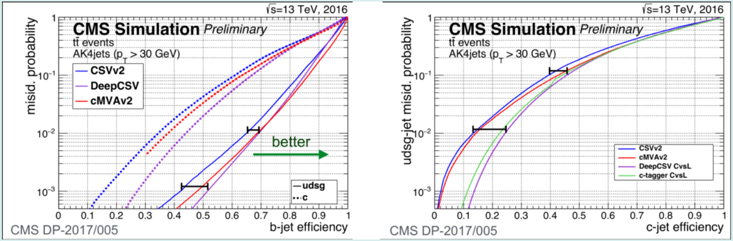


DeepCSV: Performance

DeepCSV outperforming other taggers

- In particular for high-purity selection (low misidentification probability)
- Also well-performing in terms of c-tagging

$\Rightarrow$ DeepCSV first deep-learned default tagger in CMS



Glimpse into the future: DeepFlavour tagger

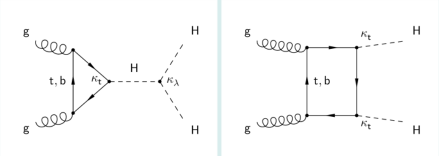
Using jet constituents / particle flow candidates as input for a convolutional neural network

Search for Higgs boson pair production in the $b\bar{b}l\nu l\nu$ final state

Analysis CMS-HIG-17-006

Search for resonant and nonresonant Higgs boson pair production in the $b\bar{b}l\nu l\nu$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV
arXiv:1708.04188 and submitted to J. High Energy Phys.
More details in Sebastien Wertz's talk on Tuesday

- Using the 35.9 fb^{-1} data of the 2016 LHC run
- One Higgs decaying into $b\bar{b}$ and the other in W($l\nu$)($l\nu$) (or either in off-shell Z(ll)Z($\nu\nu$))
- Invariant mass distribution of b-jet pairs in combination with neural network based on kinematic information as a discriminator



Search for Higgs boson pair production

Background processes

- Major bkgs: $t\bar{t}$, DY, and single top (decreasing order)
- $t\bar{t}$ as irreducible background
- Further bkgs: Diboson, triboson, $t\bar{t}V$, and SM Higgs production

Event selection (Rough summary)

- Events collected with dileptonic triggers
- Identification of jets originating from b-quarks (b-tagged jets) with combined multivariate algorithm. Algorithm applies boosted decision tree on inputs like CSVv2 tagger among other taggers / inputs.
- Final object selection:
  - 2 opposite sign leptons
  - 2 b-tagged jets

Search for Higgs boson pair production

Improvement of signal-to-background-separation with deep neural network

- (D)NN[a] based on Keras framework[b]
- Due to irreducible $t\bar{t}$ background: NN relies on kinematic information
- NN input variables: Exploitation of kinematics of the dilepton and dijet systems
  - Mass of the two lepton system
  - Transverse momentum of the two lepton or the two jet system
  - Minimal ΔR between one lepton and one jet
  - ΔR between the two leptons or two jets
  - $\Delta \Phi$ between lepton and dijet system
  - Overall transverse mass
- Furthermore, NN utilizes parameterized machine learning technique

---

[a] Since "deep" is not well-defined, I refrain from using DNN as a term
[b] see keras.io, High-level API running on top of Tensorflow and other NN providers

Search for Higgs boson pair production

Digression: Parameterized machine learning[1]

- Usual approach: Inputs to neural networks containing only measured / reconstructed features
- Parameterized ML: Expand inputs also with physics parameters

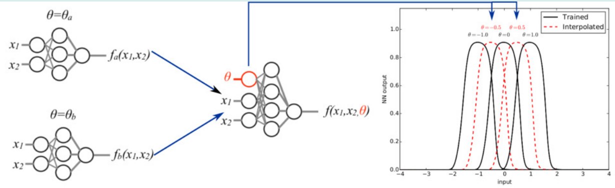


Fig.: Graphics adopted from arXiv:1601.07913

⇒ Resulting parameterized NN classifier can smoothly interpolate between physics parameters
⇒ Performance of single parameterized neural network similar to multiple individual networks

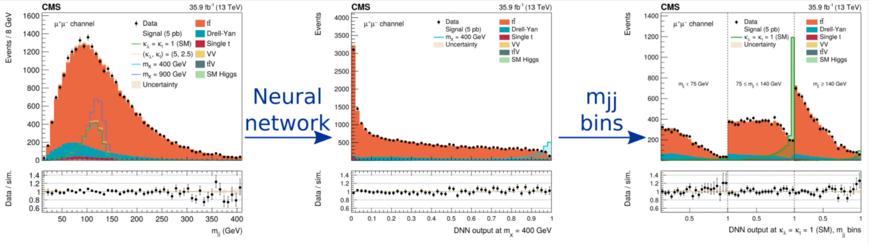
Search for Higgs boson pair production

Karlsruher Institut für Technologie

Parameterized machine learning in this analysis

- Resonant search: Adding the mass of the resonance m_X
- Nonresonant search: Adding the coupling modifiers κ_λ and κ_t
- Both cases: Adding the dilepton flavour channel (e^+e^-, $\mu^+\mu^-$, $e^\pm\mu^\mp$)

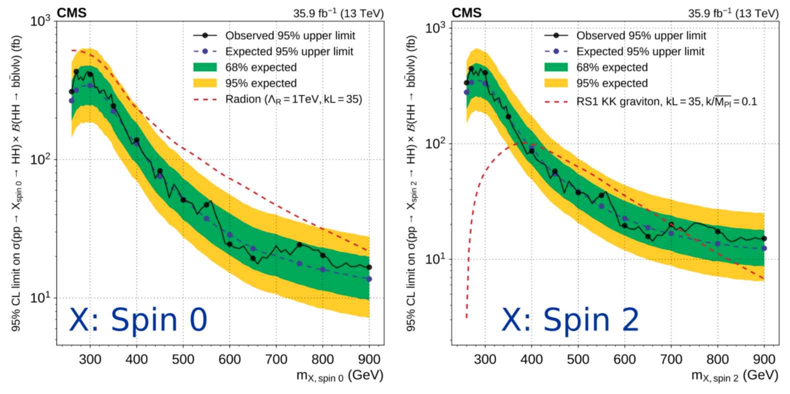
Analysis workflow for resonant search



Search for Higgs boson pair production

Result of resonant search

- Best fit signal cross sections obtained using maximum-likelihood fit
- No significant excess for X particle mass hypothesis observed between 260 and 900 GeV.



Summary

Neural networks and machine learning tools

- In many cases providing more insight in data than established tools
- Increasing spread not only in particle physics, but also in many other fields and industrial applications

My prediction for 2018 and beyond

- In 2018: Publication of several analyses showing the benefits of neural networks
  $\Rightarrow$ Stay tuned and look out for new results
- Run II and beyond: Neural networks becoming important tools to deal with large amount of data and increasing demand for precision measurements
  $\Rightarrow$ Better become acquainted with them now

Best type of neural network for your specific analysis / problem?

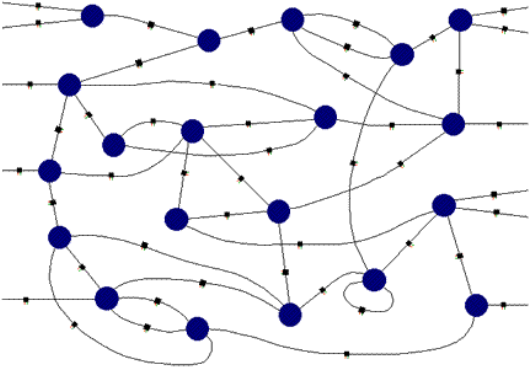


Fig.: Taken from http://www.alanturing.net/turing_archive/graphics/bigb.gif