

ML Activities and plans in Göttingen

Arnulf Quadt

11 September 2020



GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Previous activities

Göttingen group involved in various activities using "machine learning as a tool":

- Contributions to *lwttn* project: light-weight re-implementation of neural networks into c++ analysis code [PhD thesis] [github]
- *Adaptive Network Based Fuzzy Inference System*: digital-service-availability modelling of the Göttingen Tier-2 centre for short-term and long-term availability forecasts [PhD thesis]
- Extensions of existing, established analysis tools:
 - BDT and NN optimisations for the KLFitter software for top reconstruction [PhD thesis] [BSc thesis]
 - Prompt Photon Tagger*: discriminator for prompt/fake photons in ATLAS, used in *tty* measurements [MSc thesis] [PhD thesis] [PhD thesis] [paper]
- ML techniques used in various analysis projects:
 - *ttH(H→bb)* searches with NN for event classification [PhD thesis] [PhD thesis] [paper]
 - *H→ττ* searches with BDT for event classification [PhD thesis]
 - Multiple *tty* measurements with NN binary and multiclass classifiers [MSc thesis] [PhD thesis] [paper]
 - *ttW/ttZ* measurement with NN for event classification [PhD thesis] [paper]
- MC generation (SHERPA)

Future plans

- Intense activities in teaching (lectures, hands-on tutorials, data science and ML summer schools, ...)
- Existing infrastructure (Tier-2 GoeGRID, HLRN, ...) for ML already in place, to be expanded, plans to add new structures e.g. GPU farms
- Campus-Institut Data Science (CIDAS), Göttingen – a collaboration well beyond particle physics: fundamentals of data science, applied topics in science and humanities, digitisation, ...
 - Collaboration with group Kneib (Centre for Statistics, Economics) – joint expertise for algorithm development
 - Collaboration with group Ecker (Data Science, Bernstein Center for Comp. Neuroscience (BCCN)) – many computational similarities between MC simulation used in HEP and to model nerve cell activities
 - W2 prof., advanced recruitment process
CIDAS prof. with focus on high-speed/high-throughput readout (FPGA) + ML techniques
- Continued usage of ML for analysis work, ML development in regression, classification, big data analysis/parallel hardware usage in training, evaluation of systematics, MC modelling via GANs, MC generation, time series analysis, error detection (detector operation, radiation damage, grid performance, ...)

