CHAPTER 8

Conclusion

Top quark physics is one of the most active research fields in all particle physics. With a coupling to the Higgs close to unity and a resulting large mass, the top quark offers a great potential to observe physics beyond the Standard Model. The decay of four top quarks produced in the same event is a particularly rare process. Therefore is the cross-section of $t\bar{t}t\bar{t}$ very sensitive to additional BSM production processes. Its rarity makes the identification of $t\bar{t}t\bar{t}$ events particularly challenging and makes the need for more powerful and efficient techniques apparent. One promising perspective are Artificial Neural Networks, which outperform other state of the art machine learning techniques in many applications [???].

The analysis presented in this thesis uses Feedforward Neural Networks (FNNs) and Recurrent Neural Networks (RNNs) to identify $t\bar{t}t\bar{t}$ events in the same sign multilepton channel. The Neural Networks are trained on ATLAS simulated proton-proton collision data. All signal and background datasets were split into a training, a testing, and a validation dataset to ensure that the obtained separation between signal and background events generalizes well to unseen data. 18 input features were selected for the Feedforward Neural Networks. The input features obtained to be the most discriminating ones are the number of jets and a b-tagging weight. For Recurrent Neural Networks, 9 input features that contain the raw kinematic information for all particles in the event were combined with the three most discriminating features of the FNN study.

It was observed that the renormalization of event weights and the transformation of the input features to normal distributions is pivotal for the performance of the Neural Networks. The chosen number of trainable parameters in a Neural Network was observed to be important to handle overtraining and achieve the best possible performance. The performance gain for the different considered activation function and weight initializations was minor compared to the default choice (ReLU). The most impactful hyperparameter for both RNNs and FNNs was observed to be the learning rate of the optimization algorithm. In contrast to this, had the choice of the optimization algorithm itself no impact on the peak performance. The batch size of the events considered at on training iteration impacted the convergence speed of the training but not on the performance of the Neural Networks. The usage of polynomial and cyclic learning rate decay did not significantly improve the performance of FNNs.

The highest area under the ROC curve (AUC) reached for FNNs was 0.854 on the testing set and 0.852 ± 0.005 on the validation set. The best performing RNN has an AUC of 0.842 on the testing set and 0.838 ± 0.006 on the validation set. Therefore, it can be stated that in the considered case, constructing features using physical knowledge is more successful than presenting all information to the Neural Network. The signal efficiency in the defined signal region is 1.8 and was improved to 2.9 for FNNs and 2.7 by applying a optimized cut on the Neural Network score at 0.8 for FNNs and 0.85 for RNNs. In both cases, the dominant backgrounds were $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$. The background that was hardest to reduce is $t\bar{t}t$.

Truth level studies of $t\bar{t}t$ and $t\bar{t}t\bar{t}$ showed that the reconstruction of two hadronically decaying top quarks could be a promising prospect to discriminate $t\bar{t}t\bar{t}$ from the two processes. However, the performed χ^2 based reconstruction of two hadronically decaying top quarks was not able to distinguish between one and two hadronic top events. The main problem was identified to be the W boson produced in association with the $t\bar{t}t$.
The training of both FNNs and RNNs on CPUs and GPUs where investigated as a function of the batch size and number of the trainable parameters in a Neural Network. For FNNs and RNNs GPUs outperformed CPUs and decreased the training time upto a factor of 4. Therefore, GPU training of Neural Networks will become key for particle physics once the Large Hadron Collider is updated to high Luminosities.