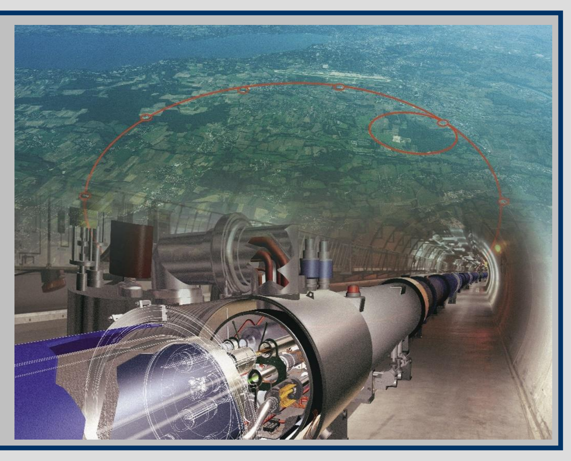
An Analysis of Standard Model and Supersymmetric Sub-Atomic Particle Events through Deep Neural Networks

Shyam Ravichandran, Ryan Rushing
Purdue University



What is the Standard Model and Supersymmetry?

The Standard Model is a widely accepted theory that classifies and categorizes elementary particles and forces, and it has successfully survived many rigorous experiments over time. These elementary particles possess 3 intrinsic properties, Mass, Charge, and Spin, and categorize particles by what they compose in a standard atom. First, Quarks form all the solid matter in the universe because they arrange to form protons and neutrons. Second, Leptons and their correlating neutrinos arrange to form electrons. Next, Force Carriers represent 3 of the 4 fundamental forces of the universe, where the 4th is the theorized Graviton. Finally, the Higgs Boson represents Scalar mass but is highly unstable and decays rapidly thus causing it to be difficult to study. The Standard Model, although very good, is an insufficient theory. The theories would expect an extremely large Higgs Boson mass and Neutrinos to be massless, which have both been observed to be false. One theoretical solution addresses these issues by introducing Supersymmetric partners for all elementary particles in the Standard Model. Supersymmetry provides an extension to the Standard Model, which alleviates some differences between theoretical predictions and experimental events. Theoretical calculations regarding Supersymmetry predict that angular distributions of the particles differ when traveling through the CMS detectors at the LHC in Geneva.

The Standard Model

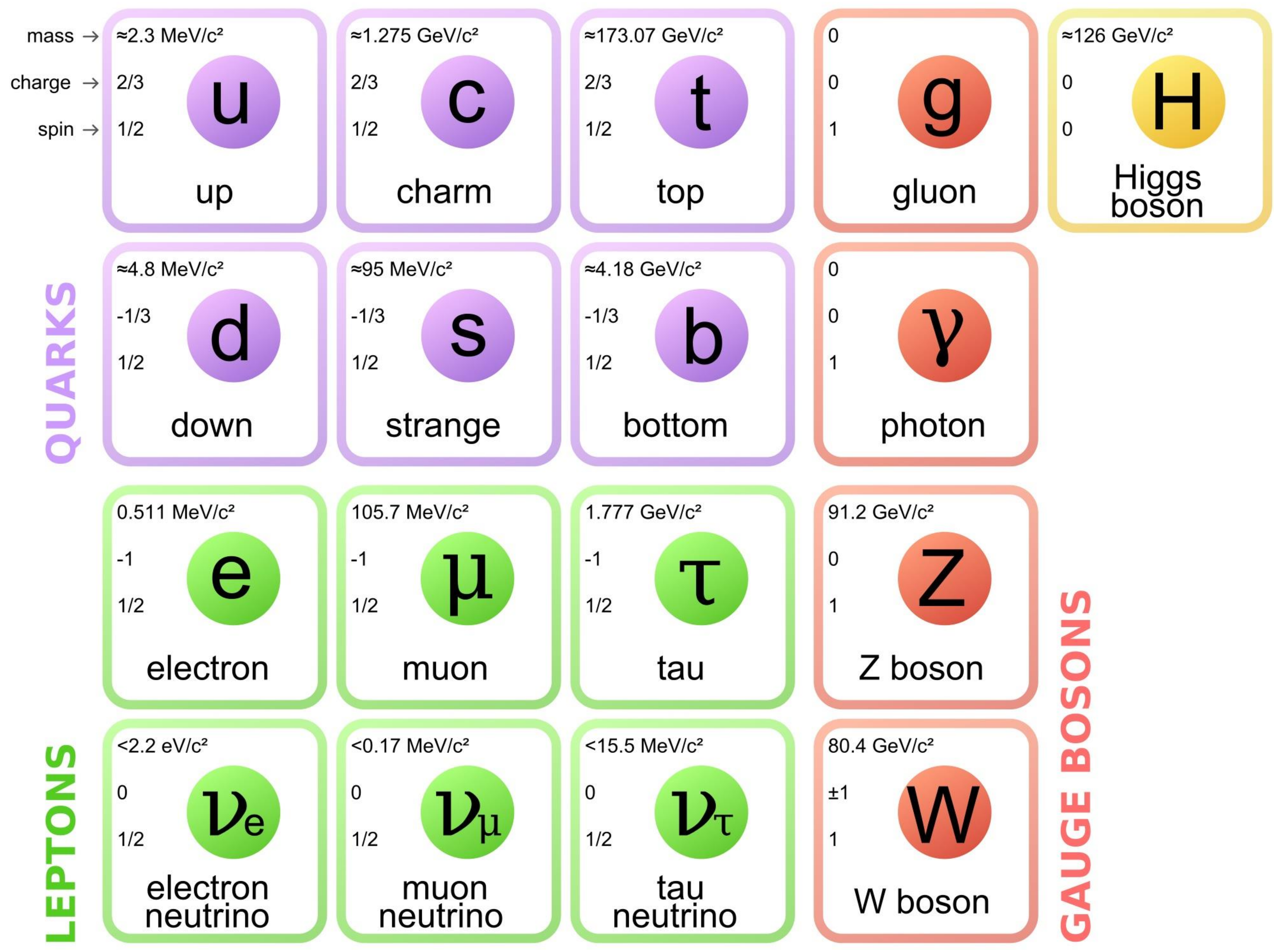


Figure 1: Graphical representation of the Standard Model.

Use of Deep Neural Networks

Differences between Standard Model and Supersymmetric events are undetectable through human testing and building a systematic approach by hand would be impossible. This led to the use of Deep Neural Networks to create a dynamic training model to search for differences through binary classification and simulated data created by Monte-Carlo simulations that theorize the properties of Supersymmetric events. While effective, Deep Neural Networks need optimization to their architecture and hyperparameters to achieve high accuracy that can be applied and generalized to new data that will be detected and recorded by the CMS detector at the LHC in Geneva.

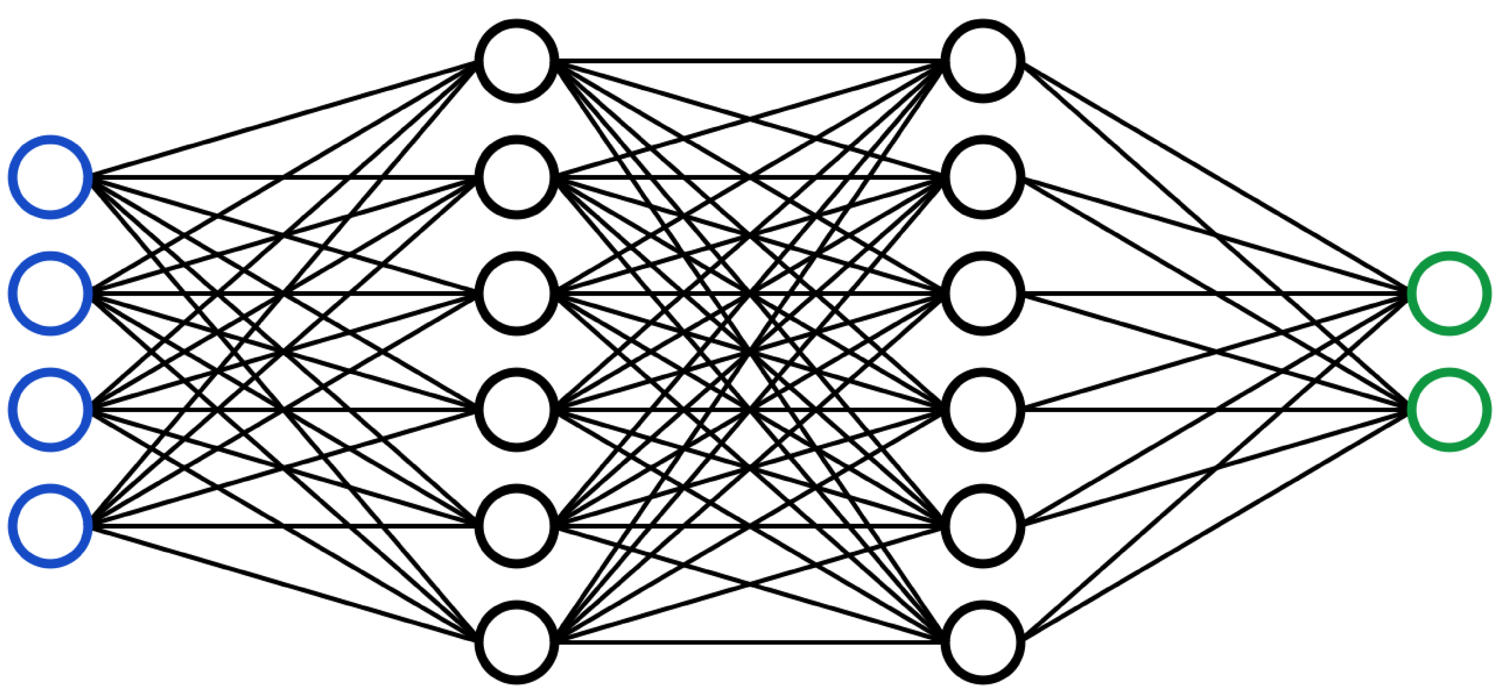


Figure 2: Simple Graphical Depiction of a Fully Connected Deep Neural Network with 4 Inputs, 2 Hidden Layers, and 2 outputs

Construction of the Neural Network

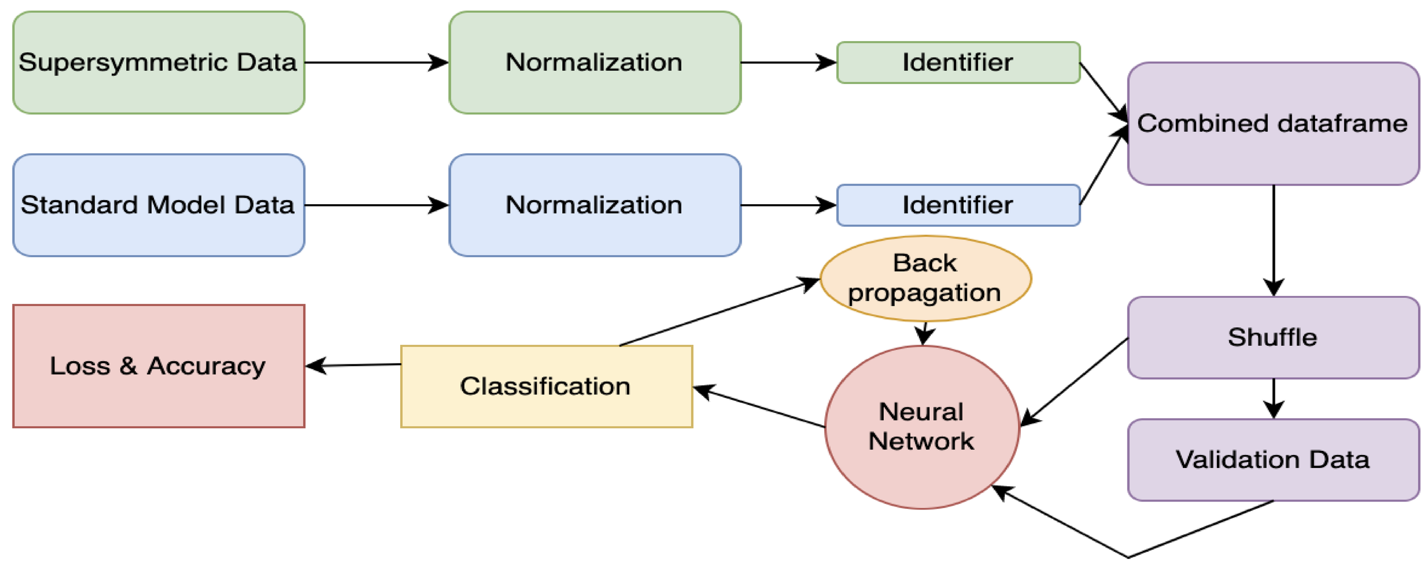


Figure 3: Flowchart showing the entire construction of the Neural Network from the data to the output values of Accuracy and Loss

Hyperparameter Optimization

The process of improving the architecture began with optimization of the various hyperparameters that influence the learning process of the Neural Network. First, this was performed manually by running different neural networks with different permutations of hyperparameters through a fixed number of epochs. After this manual hyperparameter optimization, it was noted that denser neural networks with a learning rate in the magnitude of .001 performed better yet lead to no significant changes in accuracy or loss with the final accuracy ranging between 53% and 54%. To further optimize the Neural Network, it was decided to try to use hyperparameter optimization algorithms. This was implemented through the SHERPA optimization package due to its capability with the Keras API and ease of implementing advanced algorithms such as Bayesian Optimization and Asynchronous Successive Halving Algorithm (ASHA). Using SHERPA and ASHA, the Neural Network was optimized more effectively, leading to a significant improvement in accuracy and a drop in loss.

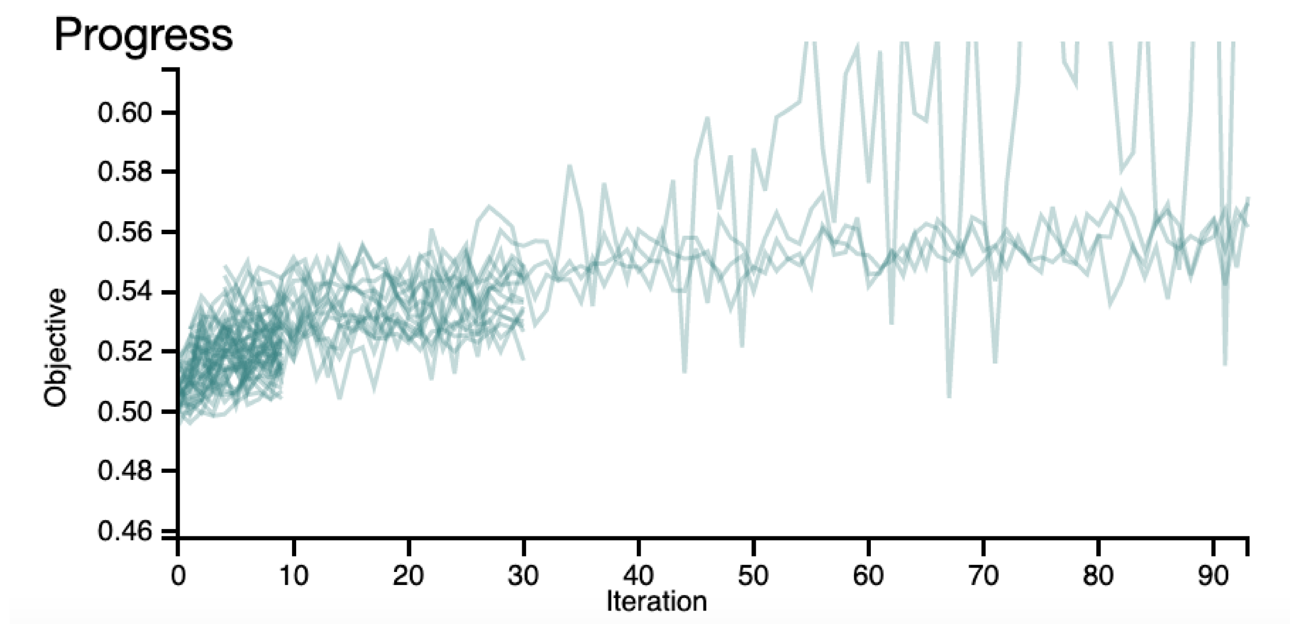


Figure 4: Plot showing the Objective over Epochs for each combination of hyperparameters tested by SHERPA using ASHA

Results and Conclusions

After running the ASHA algorithm, the prediction that the Neural Network with a denser middle layer and a learning rate in the magnitude of 10^{-2} was correct, as the best performing neural network had a learning rate of 0.065 and a middle layer with 53 nodes. The final Neural Network has a validation accuracy of 58% and a validation loss of 0.67. These results show signs of being able to successfully identify events as either Standard Model or Supersymmetry; however, the Neural Network still needs further optimization and more data to confirm this research.

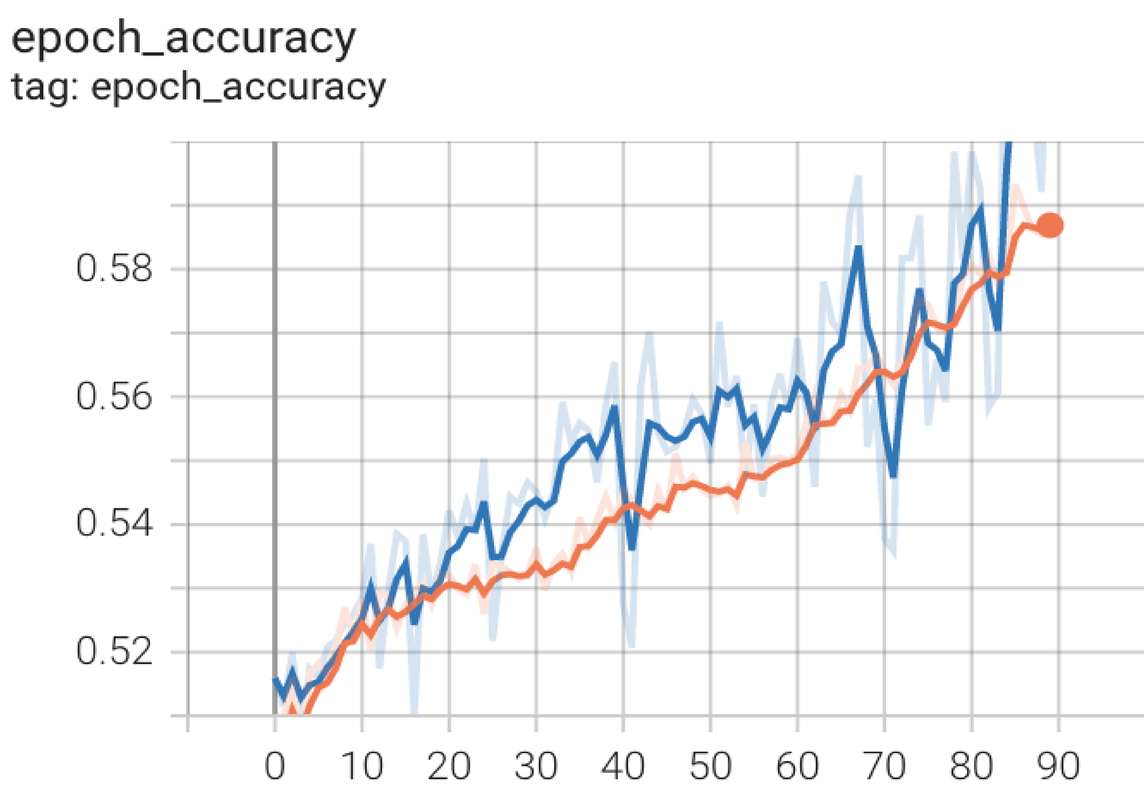


Figure 5: Plot showing Accuracy over Epochs for the final Neural Network