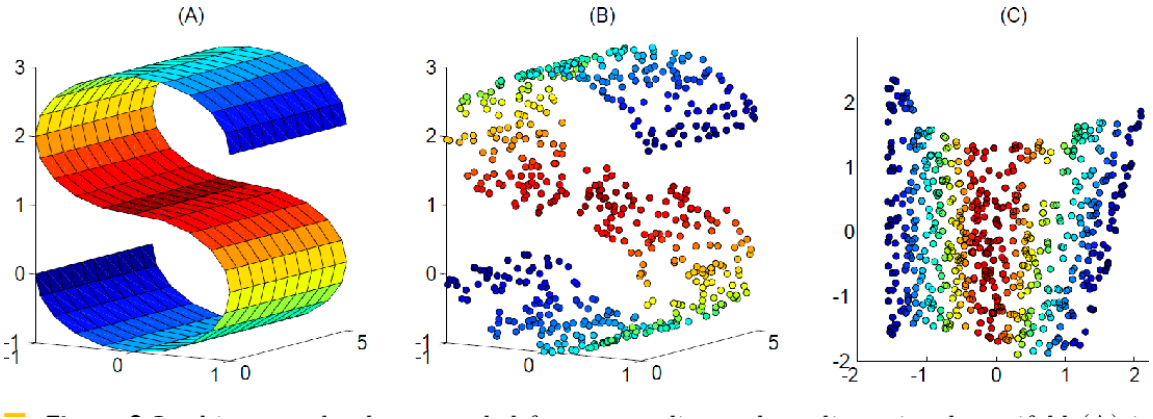
Dimensionality Reduction for Supersymmetry Dataset

Mentee: Noah Rodarte (Arizona State University)
Mentor: Sergei Gleyzer (University of Alabama)



The Abstract

Abstract:

Machine learning algorithms are a great ally in the search for new Physics. One of the many goals of Experimental High Energy Physics is to identify events that are evidence of Physics Beyond the Standard Model. Machine learning algorithms can be used to achieve better distinction between signal and background events in data analysis for the purposes of signal classification. Additionally, machine learning may be used to reduce the dimensionality of a dataset. A dimensionality reduction algorithm can infer correlations between variables and combine them into unified (reduced) variables, thereby decreasing the number of variables necessary to represent our information. This is of particular interest to new physics because these “reduced dimensions” may reveal information about underlying processes which are not given to us explicitly by detector signals. In this project, we seek to reduce the dimensionality of a supersymmetry classification dataset using machine learning to construct these reduced dimensions and to evaluate our information loss using a neural network classifier.

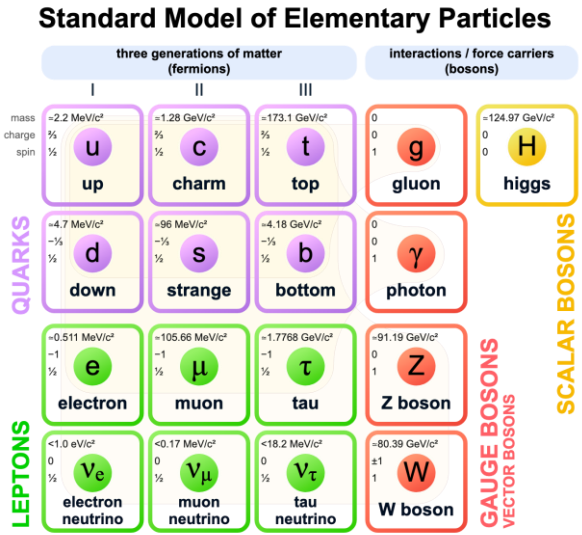
Overview

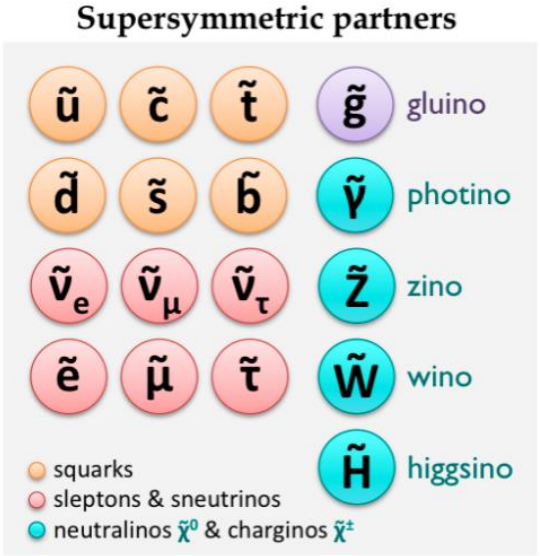
1. The Problem - Classifying SUSY particles
2. Machine Learning
3. Work to do
4. Dimensionality Reduction Explained
5. Step by Step

The Problem

- One of the largest goals in all of high-energy particle physics is searching for new particles that fall outside the scope of the standard model of physics

  - How can we learn to classify SUSY particles efficiently?*





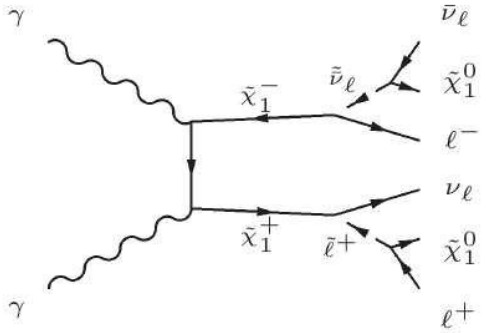
Machine Learning

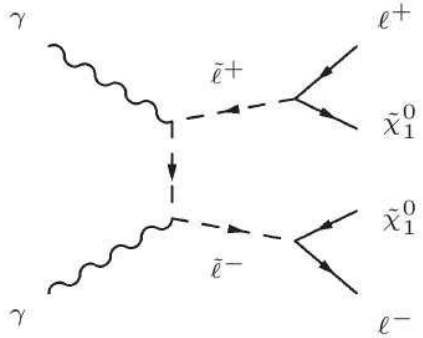
- Why use Machine Learning?
  - Machine Learning has the capability to imitate intelligent human behavior such as classification at dramatically increased speeds

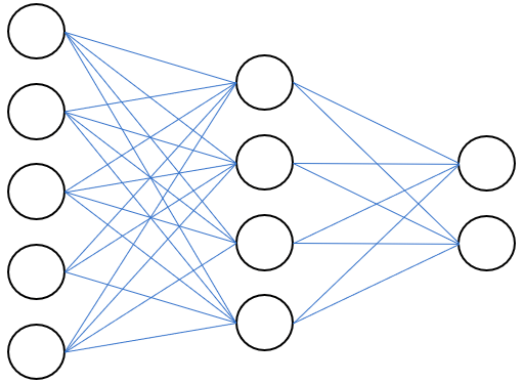
- Since the hypothetical world of SUSY particles and their complex variables keep increasing, why not try to use machine learning techniques to aid in the distinction of particle collisions?
  - Greatly increase efficiency and aid in the reduction of features necessary to identify as well as classify SUSY particles.

Work to do

- Construction of a deep neural network classifier using Python via libraries such as tensorflow and keras.
- Create, train, and test this neural network using the **SUSY dataset**
- Use **Dimensionality Reduction** to aid in the creation of a more efficient identification and classification process of SUSY particles





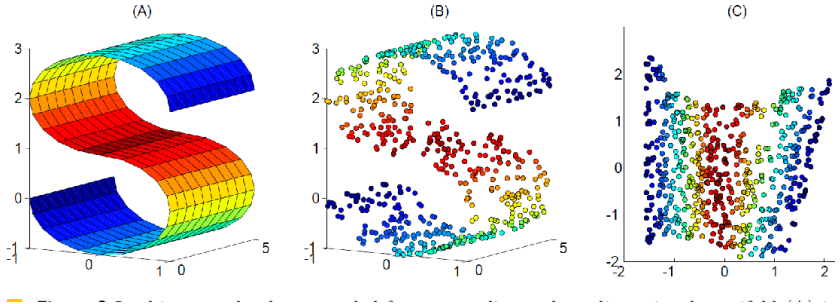


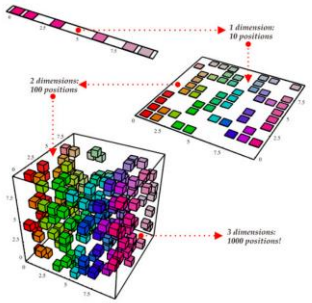
Dimensionality Reduction Explained

Dimensionality Reduction -

the transformation of data from a high-dimensional space into a low-dimensional space so that the low-dimensional representation retains some meaningful properties of the original data

- We will be using Dimensionality Reduction to hopefully reduce the need for complex variables whilst attempting to maintain or exceed the accuracy of the original variables
  - Potentially find a more efficient approach to detecting and classifying SUSY particles





Step by Step

1) Create a neural network classifier and train it on the full dataset (containing all variables/features). Adjust the architecture or training hyperparameters until a reasonably good performance is achieved.
2) Perform dimensionality reduction on the dataset to some specified number of reduced dimensions.
3) Create a new neural network using the exact same number of layers and the same numbers of neurons in each layer (i.e., using the same ”architecture” for each network).
4) Train this new network on the reduced dataset and evaluate its performance in comparison to the original.
5) Repeat for various numbers of reduced dimensions.