Machine Learning Techniques to Search for 2νββ decay of ^{136}Xe to the excited state of ^{136}Ba in EXO-200

George Halal[1]

[1]*georgech@stanford.edu, Gratta Group, Physics Department, Stanford University*

Stanford
CS229 (Fall 2019)

Introduction

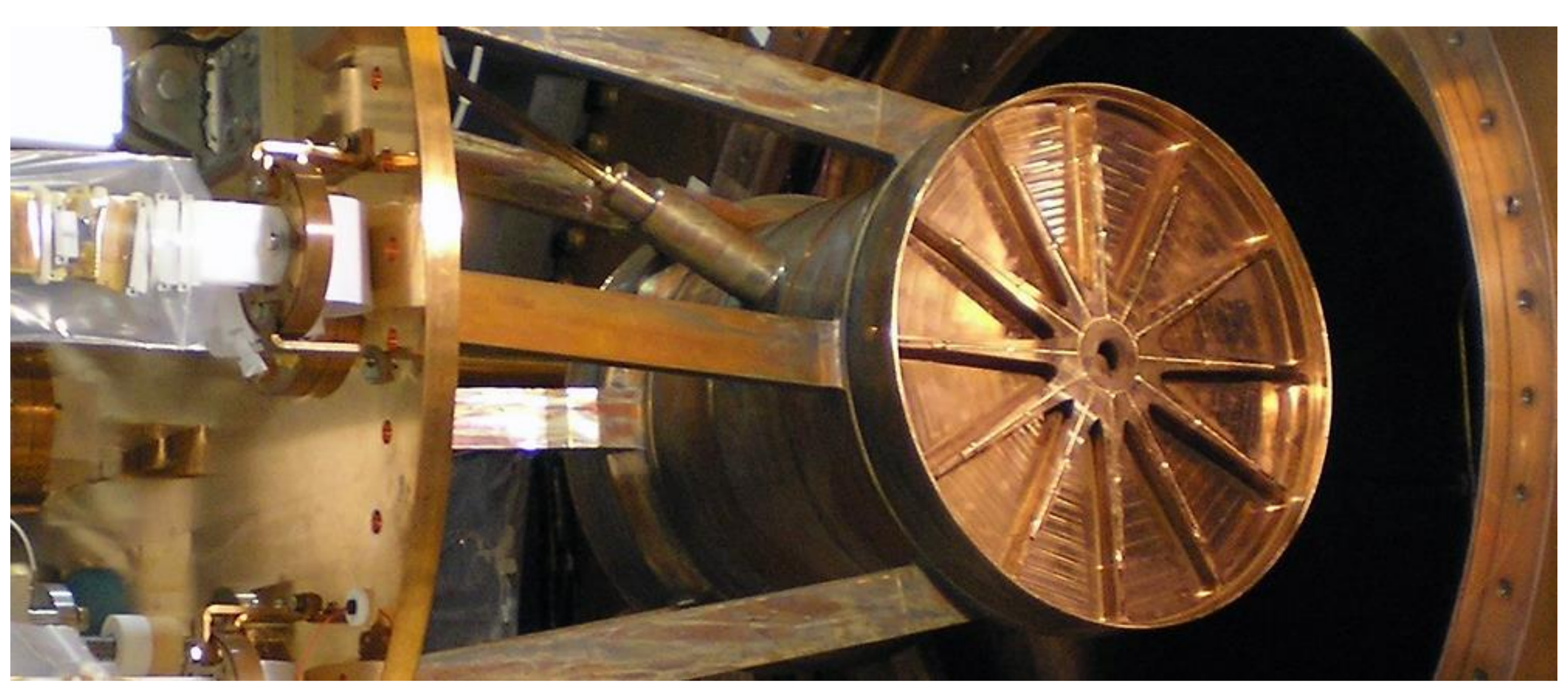
Measuring the 2νββ decay of ^{136}Xe to the excited state of ^{136}Ba is exciting because it's expected to introduce constraints to the Nuclear Matrix Elements, which would allow us to determine the effective Majorana neutrino mass more precisely[1].

Therefore, in this project, we use BDT, MLP, and LSTM models with simulations of the EXO-200 detector, to increase the sensitivity of this search and set a lower limit on the half-life of this decay. We find that the LSTM performs the best.

Data

The EXO-200 detector, pictured below, is a cylindrical barrel filled with enriched ^{136}Xe, which undergoes decays into lighter particles that are detected by sensors encompassing the walls of the detector.

We run Monte Carlo simulations of these decays that take the geometry of the detector into account. Each event is described by a number of clusters spanning the 3-dimensional volume of the detector. Each cluster measures the energy deposited in the volume of its position in the detector. We use 3,803,128 events for training, 950,782 for validation, and 1,188,477 for testing.

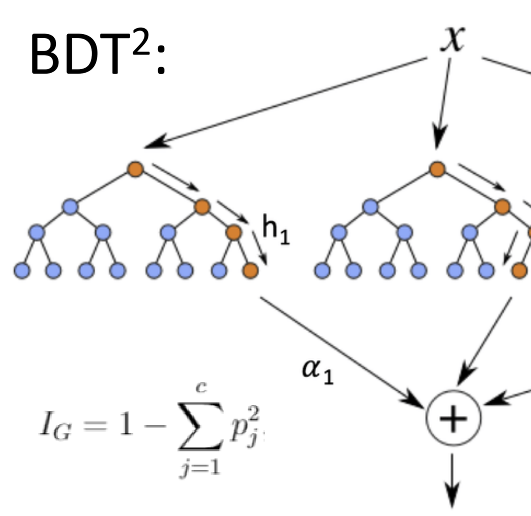


Models & Features

Boosted Decision Tree (BDT) & Multilayer Perceptron (MLP) features:

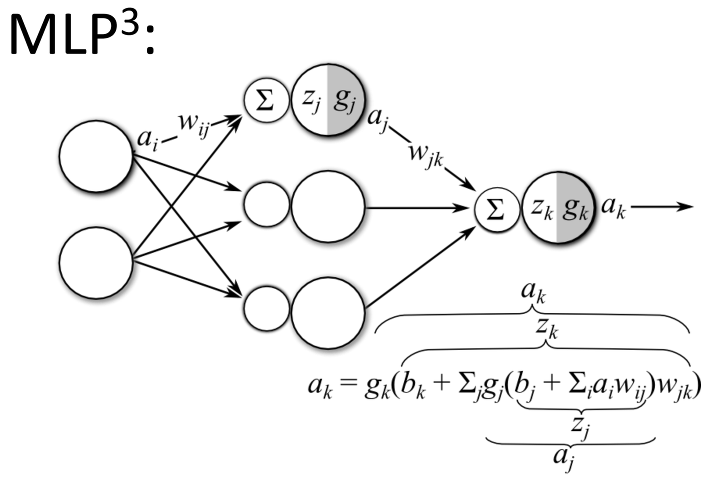
- **Sum of cluster energies / event** (normalized)
- **Number of clusters / event** (normalized)
- **Standoff distance** (normalized): minimum distance between a cluster with non-zero energy and the walls of the detector.
- γ_1, γ_2, γ_{sum} (normalized):, $\gamma_i \equiv \min_{j \in S} |E_j - \epsilon_i|$, where E_j is the energy at cluster j, ϵ_i is the theoretical de-excitation energy of photon i produced during the decay, and S is the set of clusters.
- $\Delta r_{1,2}$ (normalized): The radial distance between the positions of the clusters with energies closest to the energies of the de-excitation photons.

BDT[2]:



AdaBoost
Bagged boosting

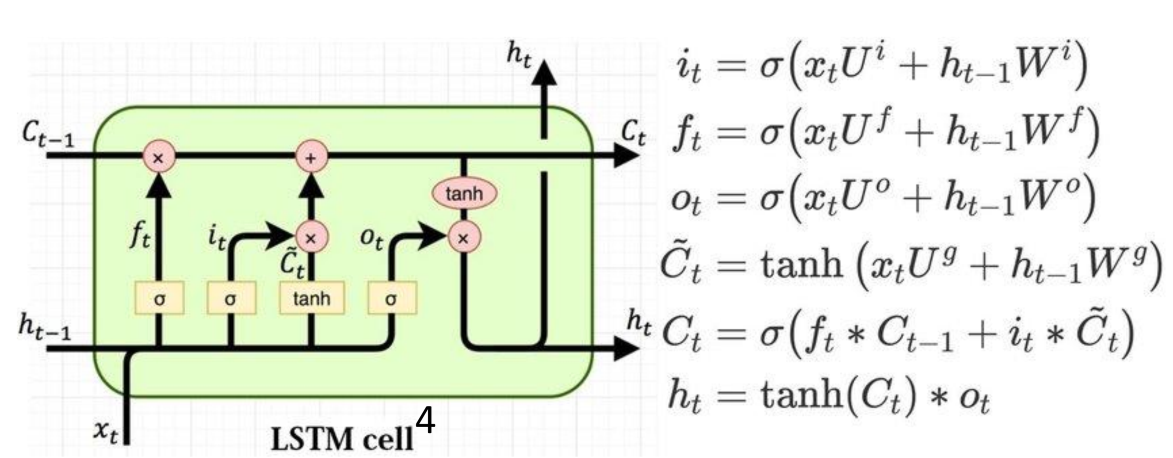
MLP[3]:



+ dropout layers

Long-Short Term Memory (LSTM) features:

We order the clusters in each event based on decreasing energy deposited in them. At each time-step, we feed the network a vector containing a cluster's **energy** and **x**, **y**, and **z positions** after normalization. The number of clusters per event varies, so we pad the sequence and mask the padded values while training.

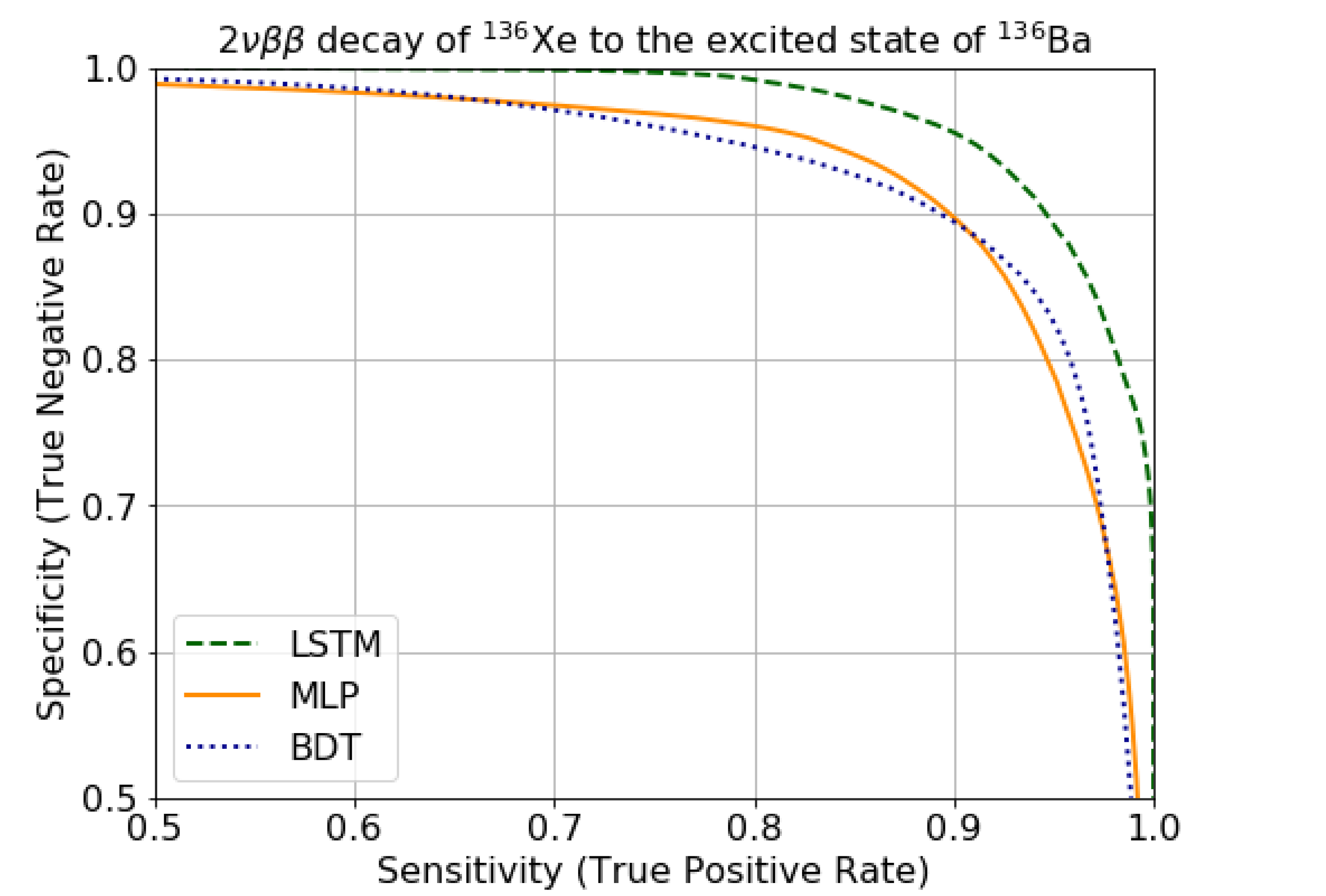


+ dropout and feedforward layers

Binary output for all models:

Signal (2νββ excited state decay) or background (all other decays)

Results & Discussion



	BDT	MLP	LSTM
Area under curve	0.9587	0.9592	0.9846

The LSTM performs the best. This is interesting, because it implies that the higher-level features used in the BDT and MLP disregard important information from the raw data used in the LSTM. Also, it implies that the order of the clusters based on decreasing energy may contain sequential information.

Future Work

Moving forward, we will apply the LSTM model to real data collected from EXO-200 to increase the sensitivity of this search and set a lower limit on the half-life of this decay.

References

1. Michael Duerr, Manfred Lindner, and Kai Zuber. Consistency test of neutrinoless double beta decay with one isotope. Physical Review D, 84(9), Nov 2011.
2. https://www.researchgate.net/figure/Schematic-diagram-of-a-boosted-ensemble-of-decision-trees_fig2_325632132
3. https://theclevermachine.wordpress.com/2014/09/06/derivation-error-backpropagation-gradient-descent-for-neural-networks/
4. https://www.researchgate.net/figure/Structure-of-the-LSTM-cell-and-equations-that-describe-the-gates-of-an-LSTM-cell_fig5_329362532