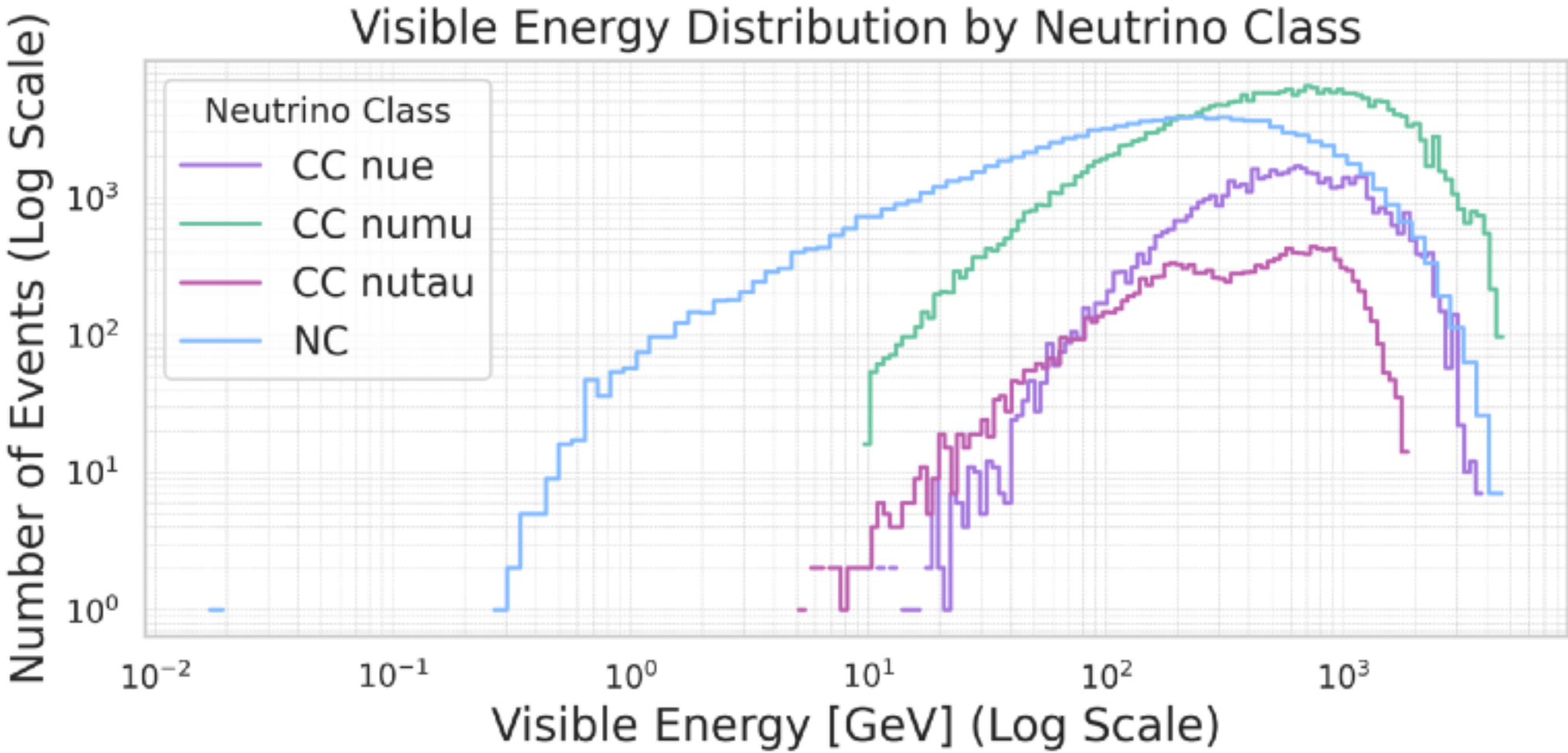# Dataset

FASERCal Simulated Data

- **Simulated Dataset:**

  – Training set: ~261k events.

  – Test set: ~131k events.

- **Energy Spectra:** interaction from ~GeV to ~TeV.

- **Class Imbalance Problem:**

  – ν_τ CC events are extremely rare.

  – *Imbalance*: central difficulty for identifying rare physics signals.

Visible Energy Distribution by Neutrino Class

Interaction Type	Events	%
ν_e CC	53,072	12.18
ν_μ CC	261,180	59.95
ν_τ CC	14,880	3.42
NC	106,502	24.45
Total	**435,634**	**100**

Methodology

Boosted Decision Trees