Machine Learning for Dark Matter and Neutrino Physics

Leila Kalhor

Department of Physics, Shahid Beheshti University

October 1, 2025

Abstract

Machine Learning (ML) is rapidly emerging as a transformative tool in fundamental physics, where rare signals, overwhelming backgrounds, and increasingly complex detectors pose unique challenges. Two fields that vividly illustrate these difficulties are the search for dark matter and the study of neutrinos. Both rely on extracting subtle signatures from vast, noisy, and high-dimensional datasets—precisely the domain where ML excels.

Astrophysical observations, from galactic rotation curves and gravitational lensing to the Bullet Cluster and the cosmic microwave background, provide compelling evidence for dark matter. Direct-detection experiments such as **XENONnT** and **LUX-ZEPLIN (LZ)**, together with complementary efforts like SuperCDMS and PICO, explore wide regions of WIMP parameter space with unprecedented sensitivity. ML enhances these searches by improving pulse-shape classification, separating electron recoils from nuclear recoils, denoising raw signals, detecting anomalies, and assisting in global statistical inference that combines multiple datasets.

Neutrino physics offers a parallel frontier. The discovery of oscillations established that neutrinos have mass, governed by the PMNS matrix. Today, long-baseline experiments (T2K, NOvA, DUNE), atmospheric detectors (Super-Kamiokande, Hyper-K, IceCube), and reactor experiments (Daya Bay, JUNO) address fundamental questions such as CP violation and the mass ordering. ML techniques are integral to these efforts: CNNs drive event classification in NOvA and DUNE, while GNNs and Transformers are pushing the boundaries of reconstruction in IceCube, from low-energy interactions to ultra-high-energy cosmic neutrinos.

A broad toolkit underpins these applications. Supervised learning methods such as boosted decision trees, CNNs, and GNNs remain central to classification and reconstruction tasks, while unsupervised approaches like autoencoders and normalizing flows enable anomaly detection and flexible density modeling. Simulation-based inference, which leverages neural ratio and score estimation, now allows direct extraction of physics parameters from Monte Carlo data. These developments are supported by modern ML frameworks (`PyTorch`, `TensorFlow`, `scikit-learn`) and HEP-specific tools (`uproot`, Awkward Arrays, `ROOT`).

Crucially, the deployment of ML in physics demands robustness, interpretability, and statistical rigor. Gaps between simulation and reality must be mitigated through domain adaptation and reweighting; classifier outputs require careful calibration and uncertainty quantification; interpretability tools such as saliency maps

and SHAP values build trust in model predictions; and blind analyses with rigorous cross-validation ensure unbiased scientific results. Practical experience also underscores the importance of starting with simple baselines, encoding detector geometry explicitly, co-designing ML with reconstruction algorithms, and enforcing strict reproducibility through careful logging of seeds, software versions, and hyperparameters.

Looking forward, the impact of ML in dark matter and neutrino physics will only deepen. Graph Neural Networks and simulation-based inference are poised to play leading roles, while cross-experiment synergies promise globally consistent analyses. Ultimately, the field is moving toward end-to-end differentiable pipelines that connect raw detector data directly to physics parameters. In this way, ML is evolving from a supporting tool into a central driver of discovery, accelerating our search for dark matter, advancing our understanding of neutrinos, and strengthening the bridge between data and theory in fundamental physics.