Neural Network Classification on Crystal Structures

Ieva

December 11, 2023

1 Introduction

In this document, we analyze a code snippet that implements a neural network-based classifier and compares its performance with two baseline classifiers: a random classifier and a most common class classifier. The code aims to classify crystal structures based on their features.

2 Code Overview

The provided code reads crystal structure data from a CSV file, where each row contains an identifier, space group, and a list of distances as features. It then splits the data into input features (`X`) and class labels (`y`). Three classifiers are defined:

1. **Neural Network (SimpleClassifier):** A simple neural network is implemented using PyTorch. It consists of four fully connected layers with ReLU activation functions. The model is trained using the Adam optimizer and cross-entropy loss.
2. **Random Classifier:** This classifier predicts class labels randomly.
3. **Most Common Class Classifier:** This classifier always predicts the most common class in the training data.

The neural network is trained and evaluated using k-fold cross-validation, and the accuracy of each model is recorded.

3 Results

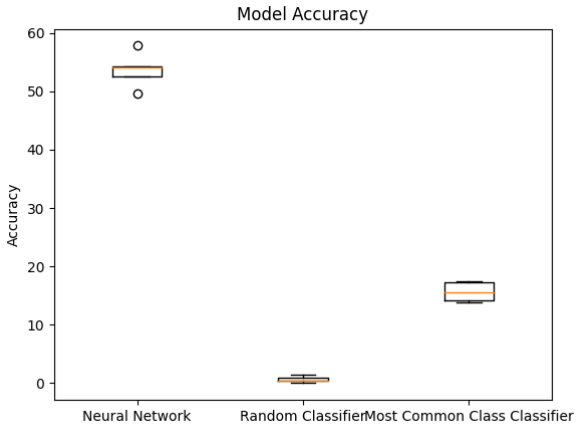

Figure 1: Model Accuracy Over k-Folds

Graph Neural Networks for Feature Learning Instead of manually designing features from crystal structures, an alternative approach is to represent the data as a graph and employ Graph Neural Networks (GNNs) for feature learning. GNNs can naturally capture relationships and dependencies between atoms or nodes in a crystal lattice. By constructing a graph object, where nodes represent atoms and edges capture spatial relationships, GNNs can effectively learn meaningful features directly from the graph structure. This approach eliminates the need for manual feature engineering and allows the model to adapt to the inherent complexity of the data. Popular libraries like PyTorch Geometric provide tools to work with graph data seamlessly, enabling the implementation of powerful GNN architectures for crystal structure classification.