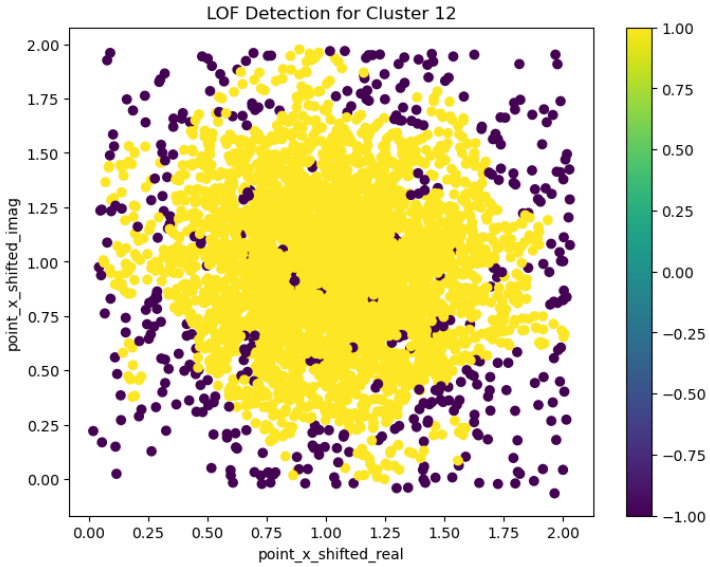
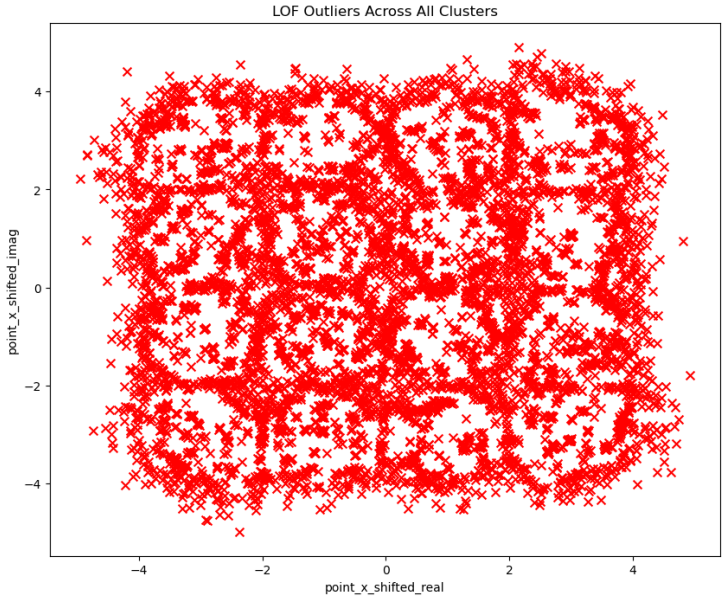
Summery Report

The data I received comprises 250 runs, each containing 65536 symbols. I initially utilized a Support Vector Classification (SVC), a specific implementation of the Support Vector Machine (SVM). An SVC is employed to classify data into one of two or more classes. In this problem, we have 16 symbols, equating to 16 classes. The machine was trained on run number 1 and subsequently tested on the remaining 249 runs. I achieved an average accuracy of 96.66%.

To enhance the accuracy, I had to identify outliers and overlapping points, so I treated each class as a cluster and then employed local outlier factor (LOF) detection. As an example, here is a sample of symbol 12:



Following that, I combined all the LOF points from the 16 symbol classes into a single dataset (about 6400 points):



Subsequently, I applied a decision tree classifier using the same inputs and target as used in the SVC but on the LOF points. Finally, I substituted the LOF values, which were previously predicted by the SVC, with the decision tree's predictions. The final accuracy of the decoded signal compared to the original one ranged from 97.2% to 99.5% for the 249 runs.