Classification of wireless transmitters

Alessandro Amella, Marian Abuziloaie

March 27, 2024

1 Introduction

The aim of this project is to classify wireless transmitters based on observed hardware imperfections. This involves analyzing a dataset of 19,200 samples, each described by nine features that represent various radio frequency impairments. This report outlines our approach to preprocessing the data, selecting and training a model, and discussing the results obtained.

2 Problem Description

Wireless transmitters often exhibit hardware imperfections that can impact their performance. The dataset provided includes measurements of nine such imperfections, including Carrier Frequency Offset (CFO), gain imbalance, and Error Vector Magnitude (EVM). Our task is to use this data to classify the transmitters, giving insight into the extent of their impairments.

3 Methodology

3.1 Data Examination and Preprocessing

Upon loading the dataset, preliminary examination revealed that it consists of floating-point and integer values, with a few columns not directly related to the classification task, such as the first (in Google Colab: `Unnamed: 0`) and `time [s]`. These were removed to focus on the nine features of interest. The data was then standardized using Scikit-learn's StandardScaler to ensure all features contribute equally to the analysis.

3.2 Model Selection

Given the nature of the task — classifying transmitters based on their hardware imperfections — we opted for a KMeans clustering approach. This decision was supported by inspection of the silhouette scores for varying numbers of clusters, suggesting that six clusters provide a good balance between within-cluster similarity and between-cluster difference.

3.3 Model Training

With the number of clusters determined, we trained a KMeans model on the standardized data. The process involved initializing the model with six clusters and a random state for reproducibility, then fitting it to the data to perform the classification.

4 Results

Upon training the model, we classified the transmitters into six distinct groups based on their impairment characteristics. The mean values of the features for each cluster, as presented in the following table, reveal notable differences in hardware imperfections across clusters.

Cl.	cfo_m.	cfo_d.	gain_i.	iq_imb	or_off	qua._err	ph_err	mag_err	evm
1	-309.24	-309.57	0.08	-34.83	-29.50	2.13	1.22	1.64	2.68
2	-373.45	-373.19	-0.03	-29.82	-36.92	-3.69	1.41	3.10	3.86
3	-398.17	-397.41	0.07	-45.94	-27.44	0.35	1.07	0.52	1.91
4	459.20	459.02	0.07	-34.70	-28.55	2.06	1.09	1.56	2.44
5	-943.04	-943.62	0.08	-38.89	-28.01	1.18	1.12	0.98	2.17
6	-703.24	-703.28	0.04	-31.66	-32.31	2.98	1.36	2.21	3.24

Table 1: Mean values of features by cluster.

Figure 1 shows the silhouette score analysis used to determine the optimal number of clusters.

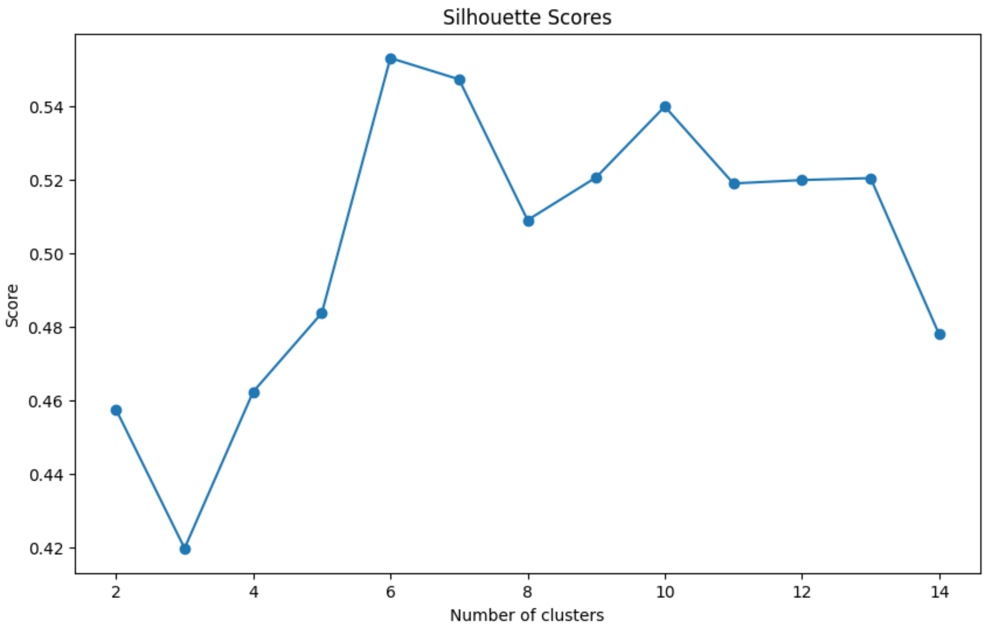

Figure 1: Silhouette score analysis for determining the number of clusters.

5 Conclusions

This project demonstrates the utility of machine learning techniques in classifying wireless transmitters based on data-driven insights into their hardware imperfections. The KMeans clustering model, chosen for its simplicity and effectiveness, successfully grouped transmitters into clusters based on the magnitude of their impairments.