172B

Some features I think are important include

MFCCs, it captures the important characteristics of the sound by analyzing the frequency of each tiny piece of an audio, it also rescales on how people perceive sound, so sounds that human ears are more sensitive to have a higher value.

Chroma stft is also an important one, it shows how pitches change over time, and because different vehicle sounds produce different pitches, so I think it might capture that pattern.

We also used root mean square, which is the average level of an audio.

We trained 3 lstm models, two models are for each dataframe, because those two dataframes are recorded in different ways, so we wanted to see if the model for one dataframe can make a good prediction on another dataframe, we will mention the result later, and also we trained one model after combining those two dataframes. Then we chose some parameters that might be influential to the result, and used grid search to find the optimal values. We tried two methods to do the grid search. The first one is that we defined a function, it is a bunch of for loops, and each for loop iterates one parameter. The parameters we chose include the number of neurons, batch\_size, activation function, dropout\_rate, etc. The return for this function is a dataframe with parameters and result for each trail, and sorted by accuracy. We used early stopping in the model, so epochs might be different. Another way we tried is using keras tuner, the parameters we search are almost the same, and beyond that we added the number of lstm layers into the tuning, for further improvement. After finding the best parameters for each model, we fit them with the corresponding dataframe, and the accuracy score for combined dataset is around 92.5%.

And then we did some data preprocessing, handling none values, particularly no-vehicle, almost half of the audio from Germany dataset are audio for no-vehicle, but they do not have label. Then used OneHotEncoder, train\_test\_split, and reshaped the dataframe to the shape required by lstm.