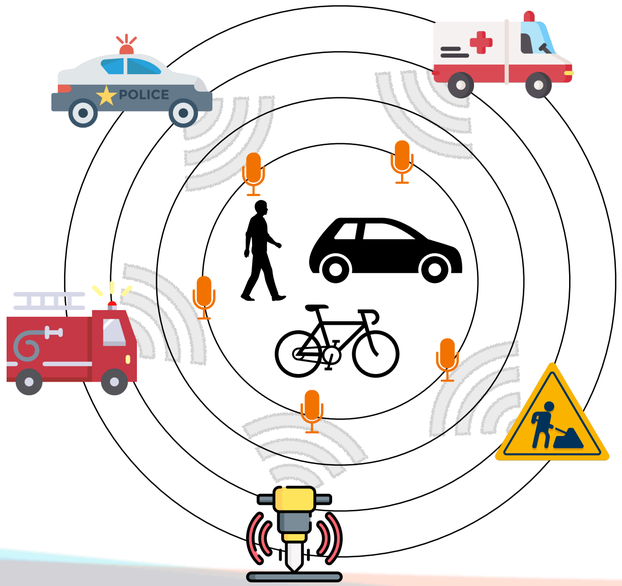
Data Mining

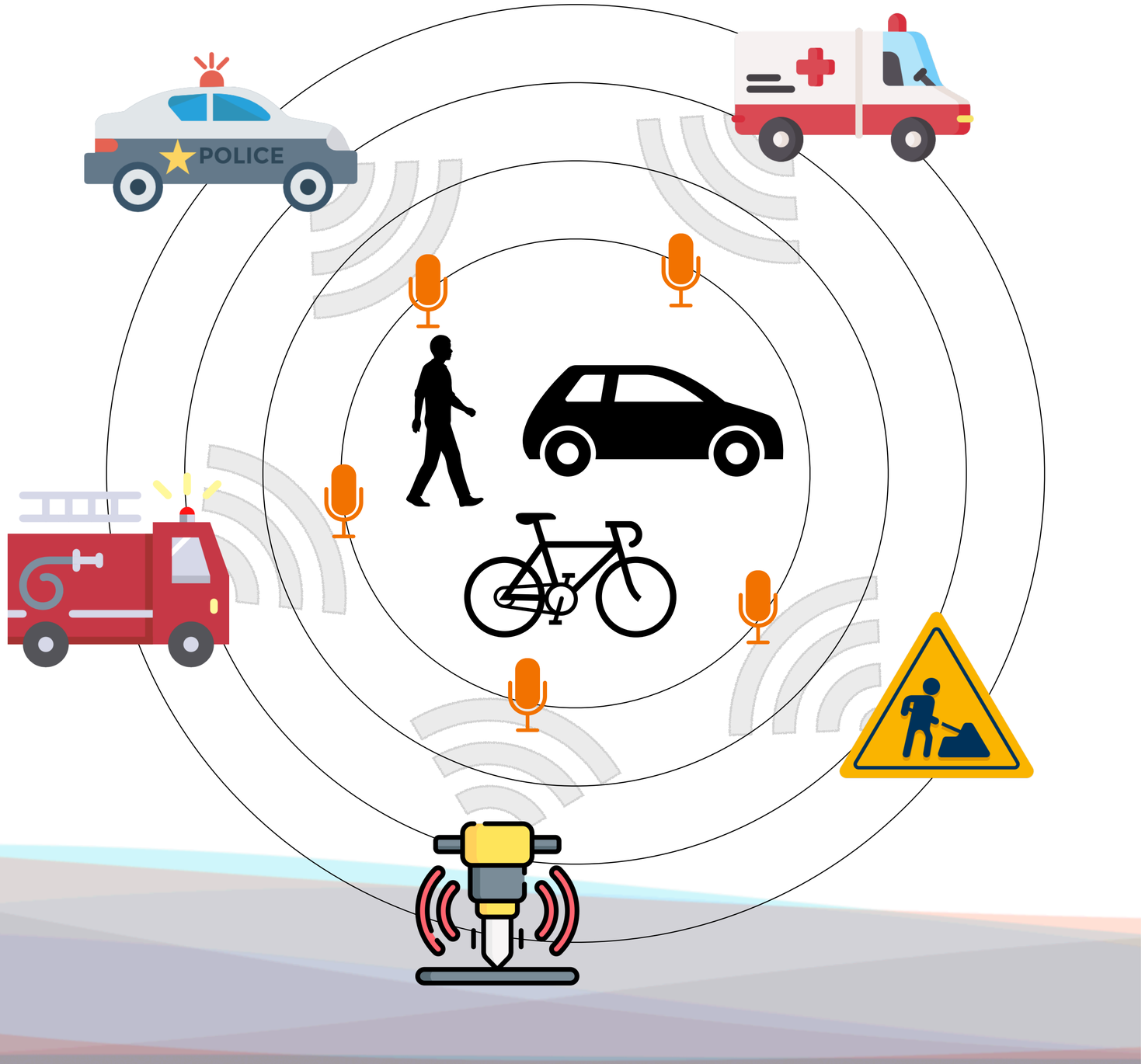
UNIVERSITÀ DI PISA

Artificial Intelligence and Data Engineering
Data mining and machine learning project presentation

Recognition system for emergency vehicle sirens and road noises



Massimo Valentino Caroti - Simone Landi

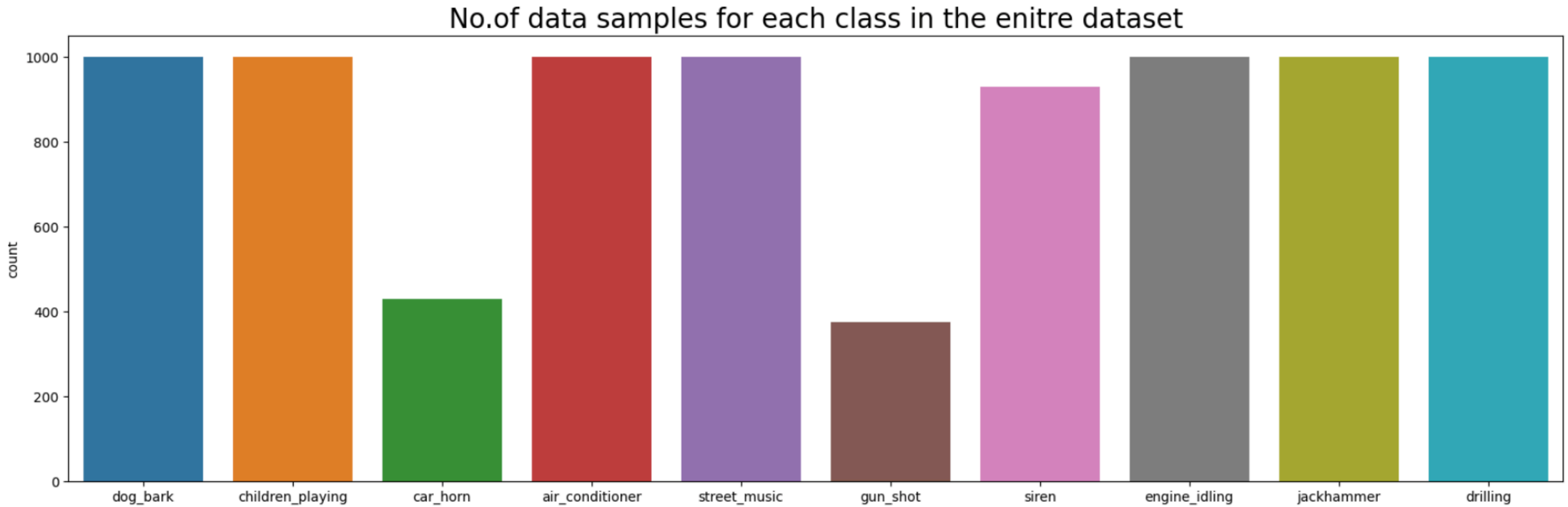


IDEA !

The aim of this project is to obtain a classifier of urban sounds to recognize the acoustic signature of each sound in order to distinguish sirens from other acoustic signals. This can be implemented on both means of transport and wearable devices to improve road safety

DATASET

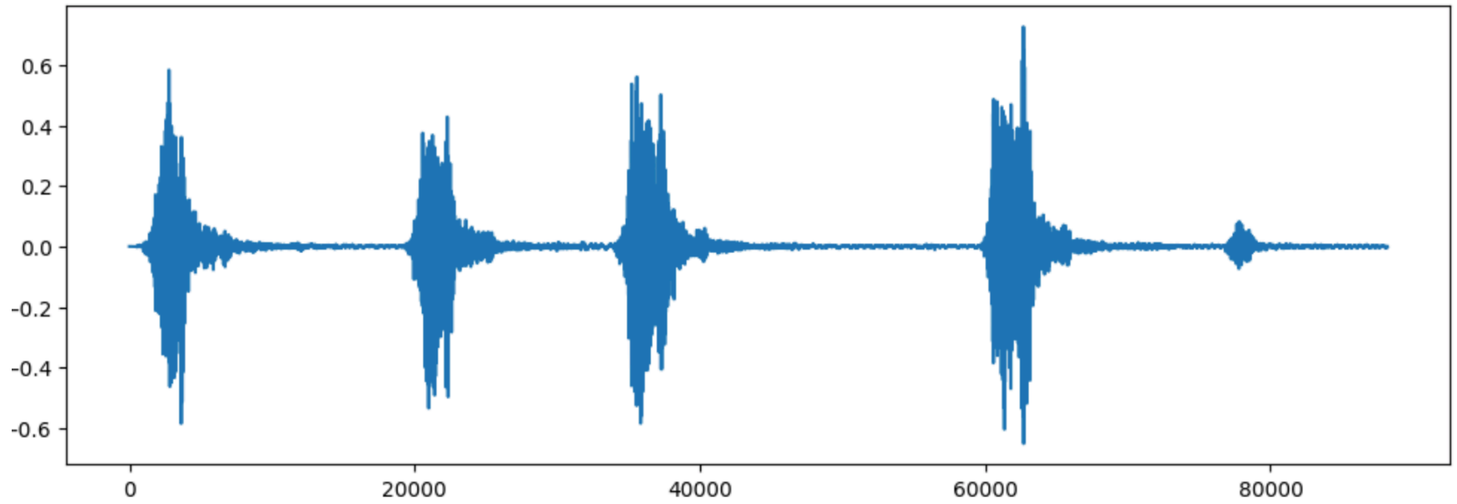
This dataset contains 8732 labeled sound excerpts (<=4s) of urban sounds from 10 classes:



The classes are drawn from the urban sound taxonomy.
All excerpts are taken from field recordings uploaded to www.freesound.org

PRE PROCESSING

we verified that the lengths of the audio samples were sufficient to be able to extract MFCC features



Librosa follows these steps to obtain the 40 features:

- Spectrograms, using the Short-Time-Fourier-Transform (STFT)
- The Mel spectrogram, from applying Mel scale filterbanks to the STFT
- Mel Frequency Cepstral Coefficients, from applying the DCT transform on the melspectrogram

REBALANCING

	Oversampling	Smote	Undersampling	No balance
KNeighbors	0.883761	0.876773	0.774392	0.877803
RandomForest	0.898188	0.896816	0.839440	0.896585
SupportVector	0.927966	0.926935	0.865322	0.926362
LogisticRegression	0.586923	0.586580	0.586923	0.590700
GaussianNB	0.457743	0.470342	0.457743	0.493359
Adaboost	0.835090	0.836920	0.835090	0.842763

The performances of the models are lower or equal than those without rebalancing.

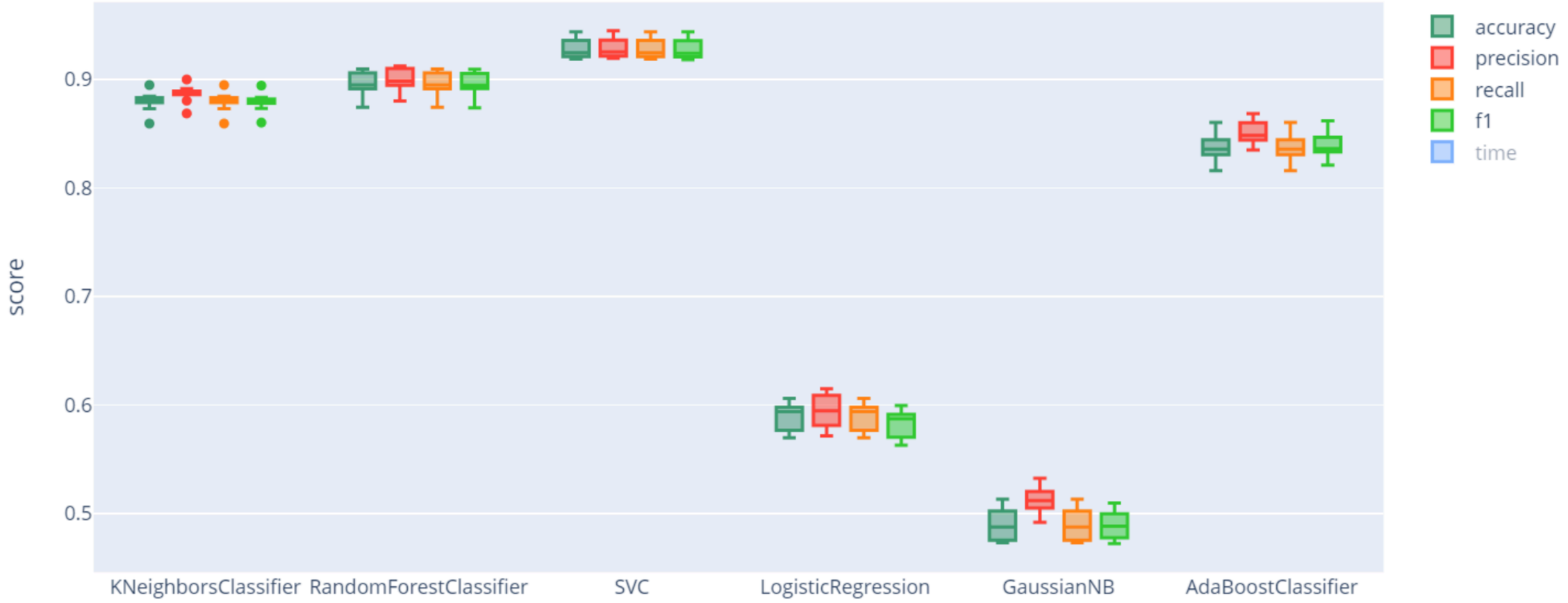
Without rebalancing, frequent cases will have fewer misclassifications.

MODEL'S EVALUATION

classifier	accuracy	precision	recall	f1	time
AdaBoostClassifier	0.837	0.850	0.837	0.839	4.088
GaussianNB	0.489	0.512	0.489	0.488	0.008
KNeighborsClassifier	0.880	0.887	0.880	0.879	0.008
LogisticRegression	0.590	0.594	0.590	0.583	3.764
RandomForestClassifier	0.896	0.900	0.896	0.896	3.089
SVC	0.928	0.929	0.928	0.928	1.748

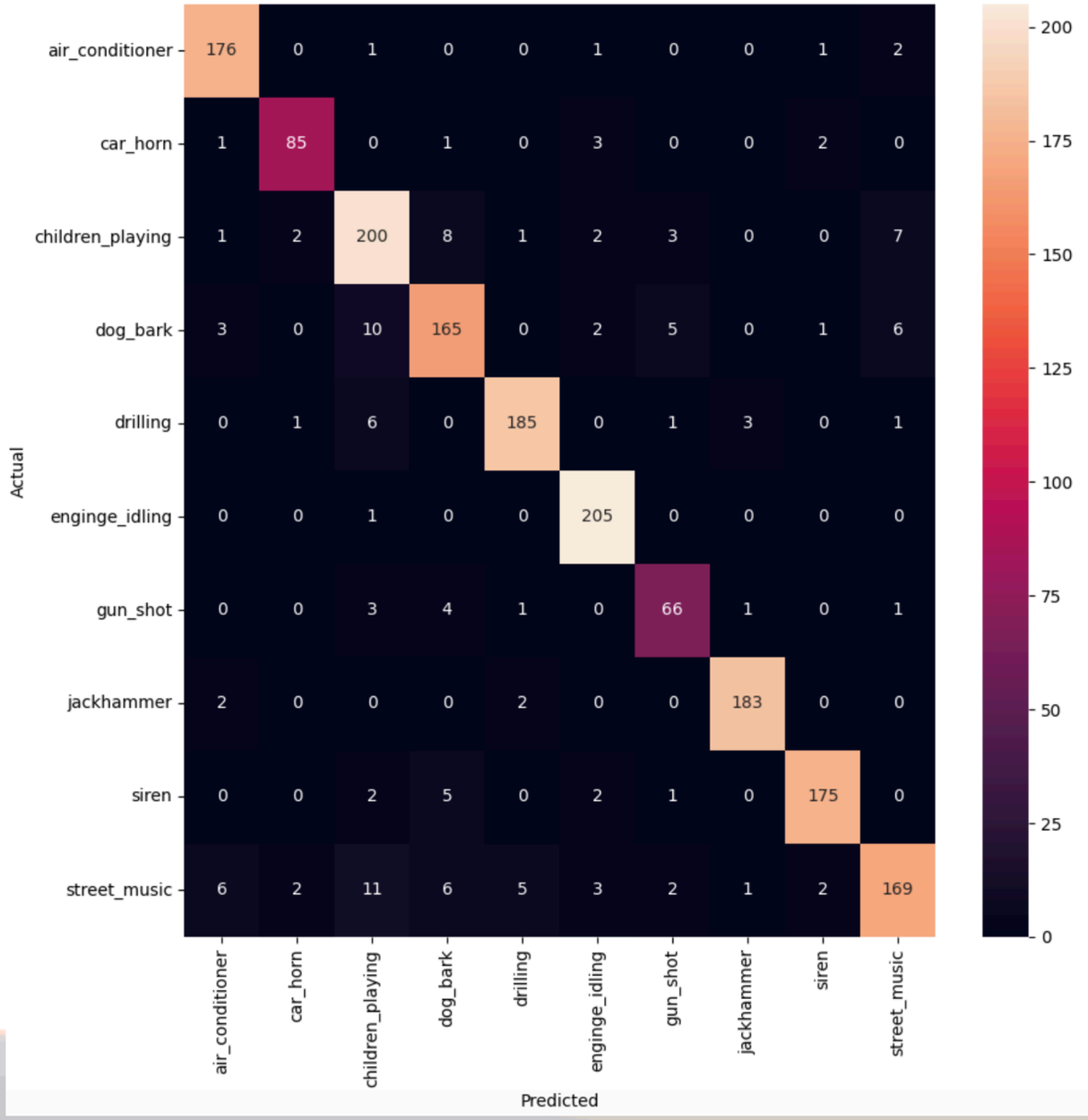
- Evaluation of the performance without applying feature selection techniques
- Evaluation of the algorithms using different score functions and a different number of selected features to find the best parameters.

MODEL'S EVALUATION



CONCLUSIONS

We chose to implement in the application the model created with SVC, because compared to other models has an optimum trade-off between F1 and training time



APPLICATION

Users can drag and drop the audio file. The file is processed by a Python server to extract features and classify the audio using an SVM model.