

Monitoring, Profiling And Classification Of Urban Environmental Noise Using Sound Characteristics, KNN Algorithm, Decision Tree, Logistic Regression And Random Forest

INTRODUCTION

Noise is an unwanted sound has negative impact upon human activities. And in this work, we have investigated the aspect of urban noise as a factor of environmental pollution and its negative impact upon human activities and tried to classify different types of sounds. In this work we used "UrbanSound8K" dataset from Kaggle.

DATA SET

This dataset consists of many audio clips with different sizes (in terms of time). Data contains a total of 10 different types (categories) of noises. All these audio files are recorded in different environments of urban regions.

ANALYSIS

1. The investigation and profiling of urban noise using selected features.
2. The selected features are used to perform noise classification using KNN (K-Nearest Neighbors), Random Forest, Decision Tree, Logistic Regression.
3. Evaluation of the results by employing noise samples from a public database.
4. Comparison between the optimal machine learning algorithm

EXISTING WORK SHORTCOMINGS

Hyperparameter Tuning: The research paper used an algorithm that has been configured to allow from 1 to 3 neighbours. Comparison to Baselines: The research paper did not establish a baseline on its performance against other relevant algorithms to demonstrate superiority far more convincing. Evaluation Metrics: The choice of evaluation metrics was limited to only Euclidean, Chebyshev and cosine

CONTRIBUTIONS:

Here in this work, we compared multiple machine learning classification models with respect to accuracy levels of predictions. These machine learning models are,

1. Linear Regression model
2. Decision Tree model
3. Random Forest model
4. KNN classification model

And for KNN classification model we have evaluated efficiencies for different KNN models, which are formed by selecting number of neighbors ranging from 1 to 9 and using distance metrics Chebyshev distance, Hamming distance, Cosine distance, Euclidean distance and Manhattan distance. Total of 45 KNN models were evaluated, and the best result was considered.

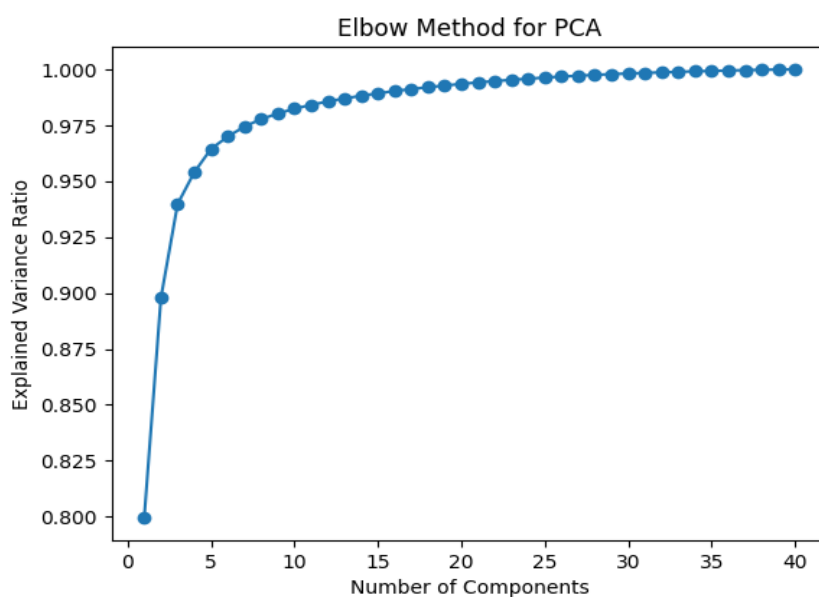
PROCESS

First, in this work, mfcc features were extracted from the audio files present in the chosen data set. Total of 40 features were extracted from each default frame using "librosa" library mfcc feature extraction method. Single feature was calculated for each feature, by averaging out all the corresponding features from all the frames of audio clip. Hence, for each audio file 40 features were generated.

This data was combined into a data frame, where each column represents one feature. Next steps are as follow,

1. Data Preprocessing
2. PCA
3. Model preparation, Training and Testing
4. Comparison of results

ELBOW METHOD



After the preprocessing, to reduce correlated columns from the data, PCA dimensionality reduction method was implemented in this work. This feature selection using Principal Component Analysis (PCA) calculates the explained

variance ratio for different numbers of components (ranging from 1 to 40) and plots the results in an "Elbow Method" diagram. The elbow point on the plot helps identify the optimal number of components to retain for reducing the dataset's dimensionality while preserving relevant information. In this work, 20 components were chosen.

KNN MODEL METRICS EVALUATION

K-NN	Euclidean	Manhattan	Chebyshev	Cosine	Hamming
K=1	93.88%	94.10%	89.47%	94.68%	11.28%
K=2	91.13%	91.24%	84.77%	91.64%	11.28%
K=3	91.01%	91.01%	85.23%	91.76%	11.28%
K=4	90.15%	90.38%	83.86%	91.01%	10.59%
K=5	89.24%	89.07%	82.71%	90.56%	10.36%
K=6	88.09%	88.90%	81.91%	89.70%	10.59%
K=7	87.18%	87.52%	81.11%	89.12%	10.59%
K=8	86.61%	86.89%	80.94%	88.49%	10.59%
K=9	85.92%	85.92%	80.14%	87.81%	11.22%

In this evaluation of the metrics, the percentage results of Cosine proved to be the appropriate metrics which is essential to provide a comprehensive assessment of the model's effectiveness. And for KNN, we get the maximum accuracy when the number of neighbours is 1, however we concluded that it could be because of overfitting of the dataset. The maximum efficiency was obtained when the number of neighbours is 3, with an accuracy rate of 91.76% which is in decreasing order with increase in number of neighbours.

EVALUATION USING DIFFERENT ALGORITHMS

1. Accuracy for Decision tree = 11.22%
2. Accuracy Logistic regression = 54.04%
3. Accuracy for Random Forest (100) = 90.61%

Here, random forest gave out a higher percentage score greater than the algorithms tested other than KNN model, so we selected it to be compared with KNN model when the number of neighbours is 3.

FINAL EVALUATION COMPARISON

This is a Confusion matrix for 3-NN classification model with cosine metrics in comparison with Random Forest Model. KNN model achieved an accuracy of 91.76%, on the other hand, Random Forest Model achieved an accuracy of 90.61%.

True Labels \ Predicted Labels	air_conditioner	car_horn	children_playing	dog_bark	drilling	engine_idling	gun_shot	jackhammer	siren	street_music
air_conditioner	191	0	1	0	0	0	0	2	0	2
car_horn	0	64	3	6	3	0	0	1	0	0
children_playing	3	0	170	5	1	1	2	1	1	13
dog_bark	1	1	18	169	3	0	4	1	1	8
drilling	0	0	2	5	202	0	0	7	0	2
engine_idling	0	0	2	1	0	180	1	0	1	0
gun_shot	0	0	0	10	2	1	59	1	0	0
jackhammer	0	0	0	1	3	0	0	215	0	3
siren	1	0	5	2	0	0	0	0	183	1
street_music	0	0	15	5	0	3	0	6	3	149

True Labels \ Predicted Labels	air_conditioner	car_horn	children_playing	dog_bark	drilling	engine_idling	gun_shot	jackhammer	siren	street_music
air_conditioner	195	0	0	0	0	0	0	0	1	0
car_horn	1	67	3	2	2	0	0	1	0	1
children_playing	3	1	176	4	3	1	1	1	0	7
dog_bark	4	3	8	174	2	1	10	1	1	2
drilling	0	2	0	4	205	0	0	7	0	0
engine_idling	1	0	1	0	1	181	0	0	0	1
gun_shot	0	1	1	9	1	0	61	0	0	0
jackhammer	1	0	0	0	3	1	1	216	0	0
siren	1	1	1	2	0	3	0	0	184	0
street_music	4	2	14	7	2	1	0	3	4	144

We could conclude that based on the dataset, the matrices and the algorithm used in this machine learning project and the random forest model used as the baseline in comparison, KNN model proved to be much more efficient in terms of accuracy.