Report on Classification on Architectural Dataset using SVM Classifier

CSLM 625 - Soft Computing

Richa Baranwal

Student ID: 232211028

Department of Computer Science and Engineering,

National Institute of Technology

Delhi

November 2023

ABSTRACT

Architectural design involves a diverse array of styles, forms, and many structural elements that contribute to the richness of the built environment. Analyzing and classifying architectural data is crucial for various applications, such as style recognition, historical preservation, and urban planning.

The study utilizes a data set comprising architectural features, including but not limited to facade designs, building materials, and spatial layouts. Feature extraction techniques are employed to transform raw architectural data into meaningful numerical representations. SVM, a powerful machine learning algorithm, is applied to these features for effective classification.

Contents

1 Introduction

The main objective of the report is to classify datasets, which are picked up from the Kaggle platform, and showcase the outcomes received from running the support vector machine model.

2 Approach

To classify the architectural data/images with the Support Vector Machine model, follow these steps:

Collect a diverse dataset of architectural images.

If necessary, preprocess the data, to make it more useful(here we need to resize the data to get an absolute solution).

After resizing, Split it into training, validation, and test sets.

Select the SVM model to train the dataset. Evaluate the performance with confusion metrics.

3 Dataset

The dataset has been taken from Kaggle(here)[1] which contains a total of 25 folders of images from across the world, out of which I have chosen 'Byzantine Architecture' and 'Romanesque Architecture' to work on it. I am pasting an image of each style for the sampling purpose:

Figure 1: A Sample Image of Byzantine Architecture

Figure 2: A Sample Image of Romanesque Architecture

4 Data Preprocessing

Before training the SVM classifier, the data was preprocessed to ensure that it was in a format that the classifier could understand. The preprocessing steps included:

1. Image resizing: All images were resized to a standard size of 512 X 512 pixels.
2. Image normalization: The pixel values of all images were normalized to a range between 0 and 1.
3. Data splitting: Split the dataset into training, validation, and test sets. The training set was used to train the classifier, while the testing set was used to evaluate the classifier's performance.

5 SVM Classifier

What is SVM?

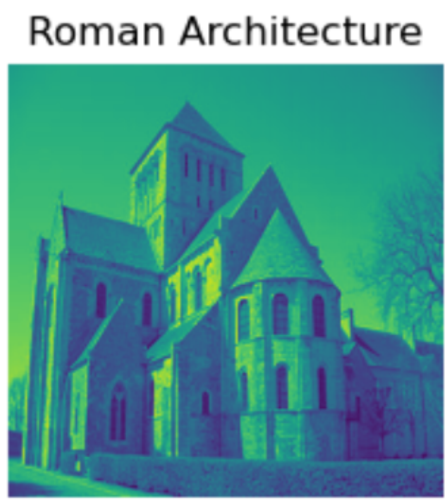
Support Vector Machines (SVMs) are a powerful supervised machine learning algorithm that can be used for a variety of tasks, including classification and regression. In the context of architectural design, SVMs can be used to classify architectural styles, identify building components, and even generate new architectural designs. It trains using the radial basis function (RBF) kernel.

6 Implemation done in code

- Import all the necessary libraries.
- Provide the path from where data is needed to be picked up
- Read the image and resize it to 512X512
- Split the data in training and testing. Now scale the feature between 0 and 1.
- Here, I have segregated the data 2 times. Once into 80-20 and another into 70-30 for training and testing respectively.
- Now train the data using a Support Vector Machine
- Get the required prediction and print
- Now testing the model with a few different datasets. Here I have taken 9 of each style(9 from Byzantine and 9 from Romanesque)
- At the end, create a confusion matrix to showcase the final output.

7 Prediction

This section evaluates the SVM classifier's predictive capacity on architectural datasets. It covers accuracy assessment, showcasing model performance across various architectural styles. Detailed analysis of feature importance and cross-validation ensures robustness.



8 Result Analysis

The performance of the SVM classifier was evaluated on the testing set. The classifier was evaluated using the accuracy metric. Accuracy is the proportion of predictions that are correct.

The confusion matrix [3] shows the true table and the achieved accuracy is 61%. This accuracy is received from 80-20 split of data.

The confusion matrix [4] shows the true table and the achieved accuracy is 59%. This accuracy is received from 80-20 split of data.

Accuracy: 0.61

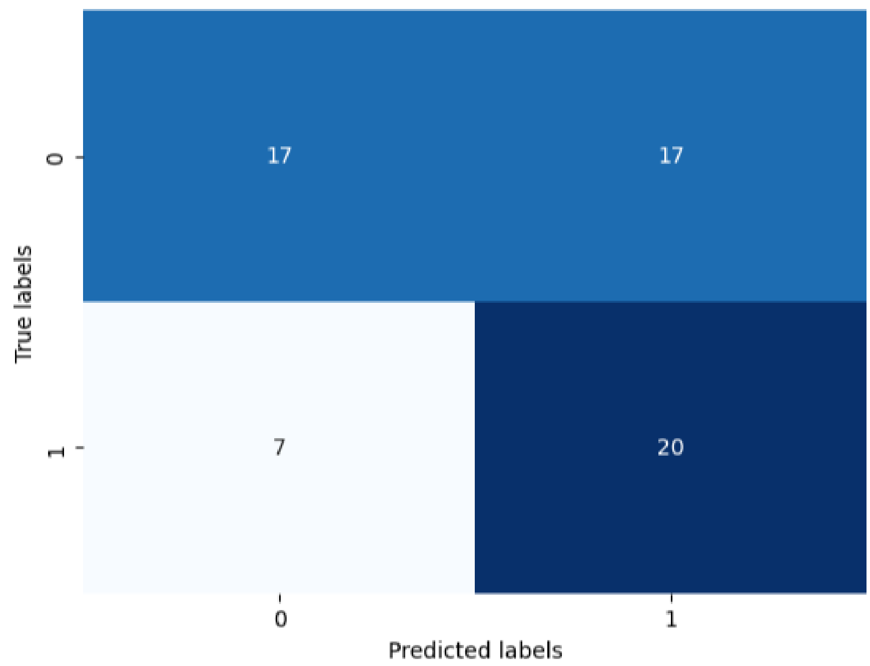


Figure 3: Confusion Matrix for the first split

Accuracy: 0.59

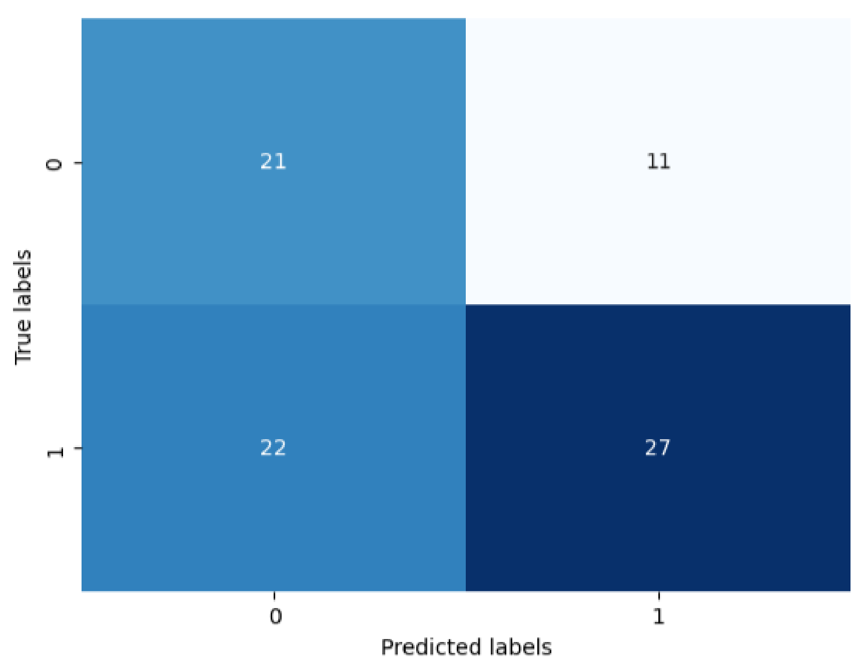


Figure 4: Confusion Matrix for second split

9 Conclusion

In conclusion, Given the obtained accuracies of 61% on the first split and 59% on the second split in classifying architectural data using SVM. it's evident that while SVM shows promise, the current model's performance is moderate. The application of SVM in classifying architectural data yields encouraging outcomes, showing its potential in automating the recognition of architectural styles, and benefiting urban planners and conservationists. Further refinement and investigation into feature selection methods could amplify the model's precision and adaptability across diverse architectural scenarios. While SVMs offer a strong foundation for architectural classification, additional enhancements are necessary to achieve higher accuracy levels and ensure their reliability in real-world architectural analysis and urban planning tasks.

10 Reference

[1] Dataset Reference:

https://www.kaggle.com/datasets/wwymak/architecture-dataset