COMP6721 Applied Artificial Intelligence Project Proposal

Chaima jaziri : 40305743 Yajing Liu 40268996 Taranjeet Kaur 40263787

A. Introduction

In this project, we aim to classify venue images into five different categories using a supervised decision tree, a semi-supervised decision tree, and a supervised CNN. Our approach involves analyzing images, processing data, training models, and evaluating their performance. Key challenges include improving accuracy, which can be addressed by using data augmentation and adjusting hyperparameters.

B. Data Preprocess

B.1. Data Description

We use 2,960 colorful images categorized into five datasets: bar (604 images), beach (622 images), Bookstore(746 images), restaurant (449 images), and subway (539 images). We eliminated parts of the datasets for having specific classes and combined datasets from different sources: image.cv [1] and mit.edu [2].

B.2. Data Exploration

To identify the class imbalance, we plot the number of images in each class in a bar chart, shown in figure 1.

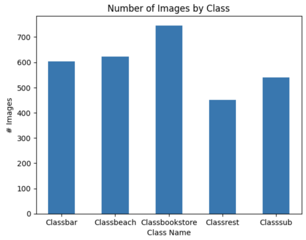

Figure 1. Identifying Class Imbalance

B.3. Data Cleaning

The acquired dataset has good quality. No duplicated images are present; results were shown in the notebook.

B.4. Data resizing

In this project, we did two data resizing processes. First, it was to the bar images to make its dimensions 256*256 pixels like the rest of the classes. Second, we found that each took over 10 minutes to load. Therefore, we resized them to 64x64 pixels.

B.5. Data Augmentation

To cater to class imbalance, around 200 new images are added to the restaurant class by augmentation. The augmentations include horizontal and vertical flips, randomly rotates images, brightness adjustment and Gaussian blur. In the figure 2, we share the new distribution of our dataset.

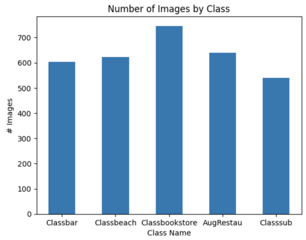

Figure 2. Augmented data

C. Methodology

C.1. Supervised Decision Tree

The supervised decision tree is mainly implemented through the scikit-learn library. Our solution contains three major steps: data preparation, model training, and evaluation.

- Data Preparation: Images are loaded via PIL and transformed with torchvision, resized to 64x64 pixels, and converted into tensors. Pixel values are normalized to a range of 0 to 1. The resulting 3D matrix is flattened into a 1D array to meet scikit-learn's input requirements.
- Model Training: In a first approach, the original dataset is then split into 80% training and 20% testing sets and later we worked on the augmented dataset. The Decision Tree Classifier is trained first with specified parameters: entropy criterion, maximum depth, minimum samples split, and minimum samples leaf and later it was choosen with a Hyperparameter tuning using Grid Search.
- Evaluation: Evaluation is performed on the testing set, computing accuracy, precision, recall, and F1 score metrics. Additionally, a confusion matrix offers detailed insight into classification performance across various classes.

C.2. Semi-supervised Decision Tree

The semi-supervised decision tree has similar data preparation and evaluation processes to the supervised decision tree. But model training shows different.

- Model Training: The dataset is split into 80% training and 20% testing sets. The labels are removed from

80% of the training data. The remaining labeled data is used to train the decision tree model. The model is trained in several iterations where in each iteration the top 10% of the high-confidence predictions from the unlabeled data are added to the labeled set. The training is done until all the data is labeled. The hyperparameters used for classification are entropy criterion, maximum depth, minimum samples split, and minimum samples leaf.

D. Results and discussion

D.1. Supervised Decision Tree

Originally, we used the following value of hyperparameters: $criterion(entropy)$, $max\_depth(20)$, $min\_samples\_split(20)$, $min\_samples\_leaf(10)$. And we got the results: accuracy (47.63%), precision (46.17%), recall (46.74%), and F1-score (46.28%). We experimented with different hyper-parameters manually and chose the following value: $criterion(entropy)$, $max\_depth(35)$, $min\_samples\_split(20)$, $min\_samples\_leaf(5)$. We got the improved results: accuracy (51.52%), precision (50.81%), recall (50.55%), and F1-score (50.51%) (Figure 3).

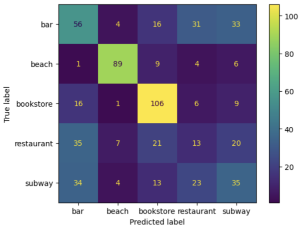

Figure 3. Supervised Learning original data

We kept the hyperparameters unchanged and trained the model with our new augmented datasets obtained through data augmentation teechniques. We got the improved results: accuracy(62.12%), precision(62.71%), recall(63.13%), and F1-score(62.73%) (Figure 4).

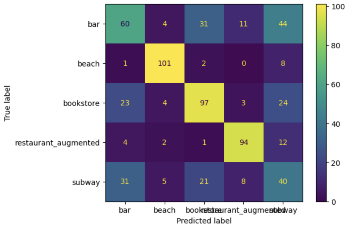

Figure 4. Supervised Learning augmented data

We chose the better value of hyperparameters with GridSearch and trained on the augmented dataset. The best hyperparameters were the following value: $criterion(entropy)$, $max\_depth(48)$, $min\_samples\_split(2)$, $min\_samples\_leaf(2)$. We got the improved results: accuracy(67.03%), precision(66.15%), recall(67.23%), and F1-score(66.28%) (Figure 5).

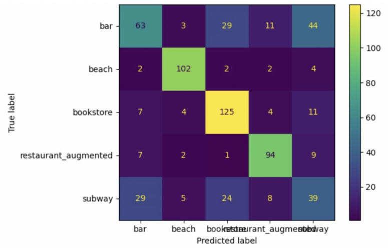

Figure 5. Supervised Learning with GridSearch

D.2. Semi-supervised Decision Tree

We experimented with different hyper-parameters manually and chose the following value: $criterion(entropy)$, $max\_depth(35)$, $min\_samples\_split(20)$, $min\_samples\_leaf(5)$. We got the improved results: accuracy (42.56%), precision (40.31%), recall (41.31%), and F1-score (40.82%) (Figure 6).

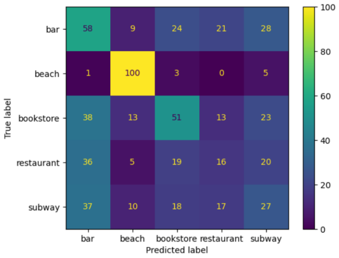

Figure 6. Semi-Supervised Confusion Matrix

E. Conclusion

In conclusion, we employed a supervised decision tree and a semi-supervised decision tree to classify 2,960 images into five categories. The semi-supervised model does not perform as well as supervised one. The achieved value of semi-supervised decision tree is 43% whereas supervised algorithm achieved 67% accuracy. To improve accuracy in the future, we will explore training models using supervised convolutional neural networks (CNNs).

References

[1] images.cv. https://images.cv/.

[2] MIT. https://web.mit.edu/torralba/www/indoor.html.