Sha’Rise Griggs

**ITAI 1378**

**Department of Science, Technology, Engineering & Math**

**Patricia Mcmanus**

**September 26, 2024**

# **Reflective Journal on Image Classification and SVM**

## **My Initial Thoughts and Feelings**

When I first encountered the exercise on image classification using SVM, I wasn’t sure how to feel. There was a part of me that was excited because I had heard about machine learning before, but I didn’t know much about the details. At the same time, I was intimidated because I had no prior hands-on experience with SVM or image classification, and I wasn’t sure what to expect when running the code. It felt a little like jumping into the deep end of the pool.

## **Prior Knowledge and Its Influence**

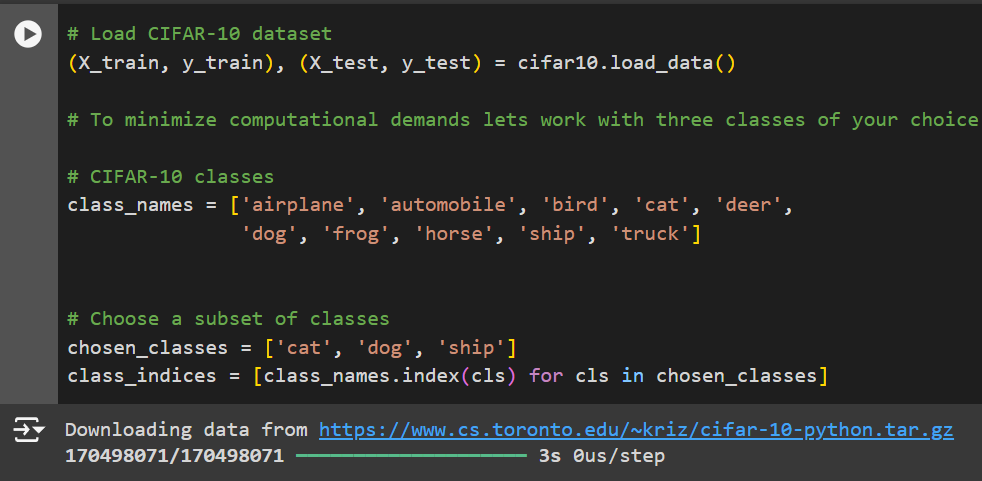
My experience with machine learning was quite limited before this exercise. I had some basic understanding of how classification models work in theory, like how you could train a computer to recognize objects in pictures, but I never worked with actual code or datasets like CIFAR-10 before. This lack of practical knowledge made me hesitant at first, but I knew this was a great opportunity to dive in and learn by doing. I just followed the steps in the Jupyter Notebook, hoping that everything would work as expected.

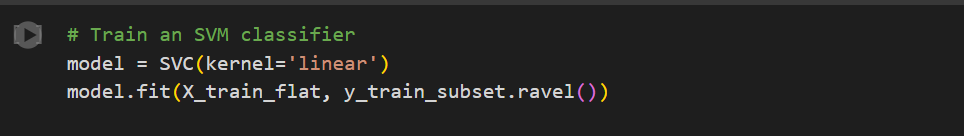
## **Understanding of the SVM Algorithm**

From what I gathered, SVM is an algorithm that helps classify data by finding the best boundary between different classes. In image classification, this means the algorithm tries to separate different objects, like cars, cats, or dogs, into the right categories. The key idea is to find a "hyperplane" that divides the data into two distinct classes, maximizing the margin between them. While the concept seemed simple enough, I realized that applying it to something like image data is much more complex because images contain a lot of intricate information.

## **Data Preparation, Model Training, and Evaluation**

One thing that stood out to me during this exercise was how important data preparation is. The CIFAR-10 dataset, which contains thousands of small images, had to be processed and organized before the model could start learning from it. This step required converting the images into a format the SVM model could understand, which involved reshaping and normalizing the data. Without this, the model wouldn’t perform well.



Training the model also required time and computational power. One challenge I faced was how long the process took. Since the CIFAR-10 dataset is relatively large, it wasn’t something that could be completed in seconds. Waiting for the model to train taught me

Training progress of the SVM model on CIFAR-10 dataset.

patience and made me appreciate just how much goes on behind the scenes in machine learning.

When it came to evaluating the model, it was fascinating to see the results. The model didn’t always get everything right, which made me reflect on how these algorithms are not perfect. They depend heavily on the quality of the data and the tuning of the model's parameters.

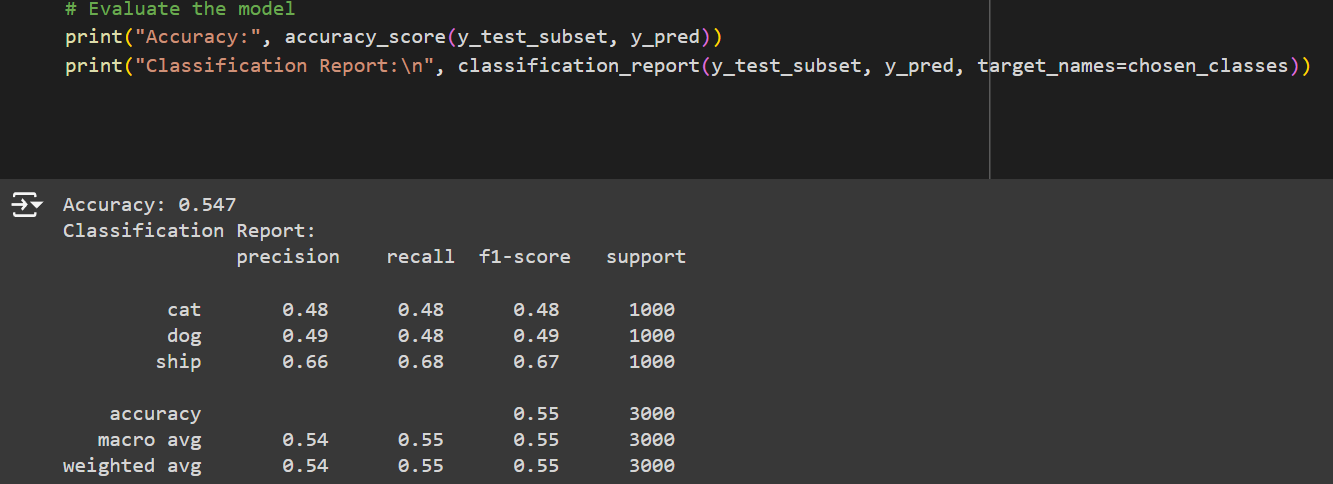
## **Challenges and Solutions**

One of the main challenges I faced was understanding the code in the notebook. Since I hadn’t worked with image classification or SVM before, it was easy to feel overwhelmed by the technical details. I got stuck a few times, especially during the data processing steps. To overcome this, I spent time Googling different terms and reading tutorials to get a better grasp of what was happening. While this slowed me down, it also helped me develop a deeper understanding of how things worked.

Another challenge was the long processing time due to the size of the CIFAR-10 dataset. My computer isn’t the fastest, so I had to be patient while the model trained. This experience made me think about the limitations of working with large datasets and the importance of having powerful hardware for machine learning tasks.

## **Insights from Model Performance**

Seeing the results of the model gave me a lot to think about. The accuracy wasn’t as high as I had hoped, which led me to reflect on what could be improved. I realized that adjusting certain parameters, like the kernel function in the SVM, could potentially lead to better

 results. Additionally, I began to understand that machine learning models don’t just “work” out of the box – they require fine-tuning and careful consideration of both the data and the algorithm.

## **Responses to Lab Questions**

**What is SVM?**

SVM stands for Support Vector Machine. It’s a classification algorithm that works by finding the best boundary between data points of different classes. In this lab, we used it to classify images from the CIFAR-10 dataset into categories like cars, trucks, or animals.

**How does the algorithm apply to image classification?**

In image classification, the algorithm tries to identify patterns or features within the image data and then classifies the image based on those patterns. With SVM, the goal is to draw a boundary that separates images into their correct categories by using the features extracted from the image data.

**What were the main challenges in working with CIFAR-10?**

The main challenge was dealing with the size and complexity of the CIFAR-10 dataset. The dataset contains 60,000 images, which means a lot of data needs to be processed and handled during training. This can lead to long processing times and requires a lot of memory and computational power.

## **Critical Analysis & Referencing**

This exercise helped me understand the importance of selecting the right algorithm for a given task. While SVM can be effective for smaller datasets or simple classification tasks, I realized that for more complex datasets like CIFAR-10, more advanced models like Convolutional Neural Networks might be better suited. SVM’s ability to handle large amounts of image data is limited, which became clear when the accuracy wasn’t as high as expected.

By looking at other machine learning approaches, I learned that combining SVM with other techniques, such as feature extraction using deep learning, could potentially improve performance. This shows that machine learning is not a one-size-fits-all field it requires experimentation and a good understanding of the data.

### **References**

* CIFAR-10 Dataset Documentation (<https://www.cs.toronto.edu/~kriz/cifar.html>)
* Scikit-learn: Support Vector Machines (<https://scikit-learn.org/stable/modules/svm.html>)
* Machine Learning Algorithms: Support Vector Machine (SVM) – A Tutorial (<https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/>)