Student Name: Juan Holmes

Course: ITAI 1378

Reflective Journal on Image Classification Using Support Vector Machine and CIFAR-10

1. Reflection on Learning

* 1. Understanding SVM in Practice

Before this laboratory session, I did not understand Support Vector Machines but applying them to CIFAR-10 for picture classification proved to be highly enlightening. Although SVM is not typically employed for picture data, it can perform effectively with enough pretreatment.

1.2 Data Preparation and Its Importance

Flattening the photographs and converting them to grayscale prompted me to reflect on the significance of the data. The project's needs dictate the best resolution media to accomplish your goals. This experience highlighted the need to comprehend your data before engaging in model selection.

1.3 Training and Model Performance

Given the high dimensionality of the CIFAR-10 data set, SVM did take some time, even after only using a subset of the original dataset. The outputs allowed me to progress in the project, but I am interested in using different models to compare in the future.

1.4 Overcoming Challenges

The lab helped me understand that preprocessing is just as important as the

algorithm itself. I’m interested in working on more preprocessing tasks in the future.

The quality of the data you feed into the model can greatly affect its performance. Proper preprocessing, like cleaning and normalizing data, makes sure the algorithm works with the best possible information, leading to more accurate results.

2. Applying What I Learned

2.1 SVM Applications

SMV may also prove advantageous for tasks involving many features. Applying this model showed me that it is essential, especially when conserving computational resources is a priority.

2.2 Relevance to Future Projects

This lab demonstrates that algorithm selection during preprocessing is crucial. I am interested to learn what that selection process could look like. Based on my other courses, a framework could look like this:

- Identify the Problem Type: Determine if it's a classification or regression problem to choose the right algorithms.

- Examine the Dataset:

- Check the dataset size to match it with suitable algorithms.

- Identify if features are numerical, categorical, or a mix.

- Handle Missing and Imbalanced Data:

- Fill in missing values.

- Balance class distribution if needed.

- Check Data Dimensionality:

- Use algorithms like SVM for high-dimensional data.

- Try simpler models for low-dimensional data.

- Transform Data:

- Normalize/standardize data if needed.

- Encode categorical variables.

- Consider Computational Resources:

- Pick simpler algorithms if you have limited resources.

- Use complex models if you have more computing power.

- Start Simple:

- Test simple algorithms first to set a baseline.

- Evaluate and Adjust:

- Use cross-validation to measure performance and pick the best algorithm.

3. Key Takeaways

1. Preprocessing is Crucial: How you prepare your data significantly influences the outcomes. Converting photos to grayscale and flattening them simplified the process while preserving essential information.

2. Deep Learning Is Not Always Necessary: While CNNs excel in image processing, SVM remains effective for less complex or smaller datasets.

3. Understand the Constraints: SVM performs well with high-dimensional data; nevertheless, it has limitations with images, making the selection of the appropriate model for the situation essential.

4. Visuals

Grayscale vs. Color Images

A collage of images of cats

Description automatically generated

Conclusion

In summary, how you prepare your data is very important, as making images more straightforward, like turning them to grayscale, helps without losing essential details. While deep learning models like CNNs are great for images, simpler models like SVM can work well with easier tasks or smaller datasets. It’s also important to know that SVM has limits with images, so picking the right model for each problem is critical.

Works Cited

1. Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

2. Raj, A., Gupta, J., & Manning, C. D. (2020). *Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems.* O'Reilly Media.

3. Müller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media.