Brief Description of the Fashion-MNIST Dataset

The Fashion-MNIST dataset is a popular benchmark dataset widely used for testing and validating computer vision models, particularly for image classification tasks. It serves as an alternative to the classic MNIST dataset, replacing handwritten digits with grayscale images of fashion items. The dataset is designed to be a more challenging and realistic problem for machine learning models to solve.

Key Characteristics

Number of Classes: The Fashion-MNIST dataset consists of 10 classes, each representing a specific fashion item category. The classes include items such as T-shirts/tops, trousers, pullovers, dresses, coats, sandals, shirts, sneakers, bags, and ankle boots.

Image Size: All images in the dataset have a resolution of 28x28 pixels, similar to the original MNIST dataset. These images are grayscale, meaning they have only one channel, representing different shades of grey.

Number of Samples: The dataset contains 60,000 training images and 10,000 test images. These images are evenly distributed across the 10 classes, with 6,000 samples per class in the training set and 1,000 samples per class in the test set.

Label Encoding: Each image in the dataset is associated with a label (0 to 9) corresponding to the fashion item category it represents.

Purpose and Usage

The Fashion-MNIST dataset is widely used in machine learning research and education to evaluate and compare the performance of various image classification models. Its similarity to the original MNIST dataset allows researchers to assess the generalisation capabilities of their models on more complex and diverse images, which are representative of real-world scenarios.

Source: Fashion MNIST | Kaggle

Evaluating the Models: CNN vs. SVM

The report compares two models, Convolutional Neural Network (CNN) and Support Vector Machine (SVM), applied to the Fashion-MNIST dataset for image classification. The evaluation metrics used to assess the models' performance are accuracy, precision, recall, and F1-score.

CNN Model

- Test Accuracy: 0.881
- Precision: 0.8828
- Recall: 0.8810
- F1-score: 0.8811

SVM Model

- Test Accuracy: 0.8446
- Precision: 0.8436
- Recall: 0.8446
- F1-score: 0.8438

Discussion

The CNN model achieved superior performance across all evaluation metrics compared to the SVM model. Here are the key observations:

1. **Accuracy:** CNN outperformed SVM by approximately 4.5%. This indicates that the CNN model correctly classified more instances than the SVM model.
2. **Precision:** CNN's precision score was slightly higher than SVM's by approximately 0.39%. Precision measures the proportion of true positive predictions among all positive predictions. A higher precision score implies fewer false positives, which is crucial in tasks where misclassification is undesirable.
3. **Recall:** CNN's recall was marginally better by about 0.1%. Recall, also known as sensitivity, measures the proportion of true positive predictions among all actual positive instances. A higher recall indicates that the model can capture more positive instances correctly.
4. **F1-score:** CNN's F1-score was higher by approximately 0.36%. F1-score considers both precision and recall and provides a balance between the two metrics. A higher F1-score indicates that the model strikes a good balance between minimising false positives and false negatives.

Based on these evaluation metrics, it is evident that the CNN model outperformed the SVM model in image classification on the Fashion-MNIST dataset. CNN's ability to learn hierarchical features and patterns from images enables it to excel in such tasks, leading to higher accuracy and better overall performance. Therefore, for this particular classification problem, the CNN model is recommended as the superior choice due to its higher accuracy and balanced performance across precision and recall metrics.

CIFAR10 Dataset [CIFAR-10 and CIFAR-100 datasets]

The CIFAR-10 dataset is a popular benchmark dataset used in computer vision and machine learning research. It stands for the "Canadian Institute for Advanced Research" where it was created. The dataset consists of 60,000 color images in 10 classes, with each class having 6,000 images. The images are of size 32x32 pixels, and each pixel is represented by three color channels (red, green, and blue).

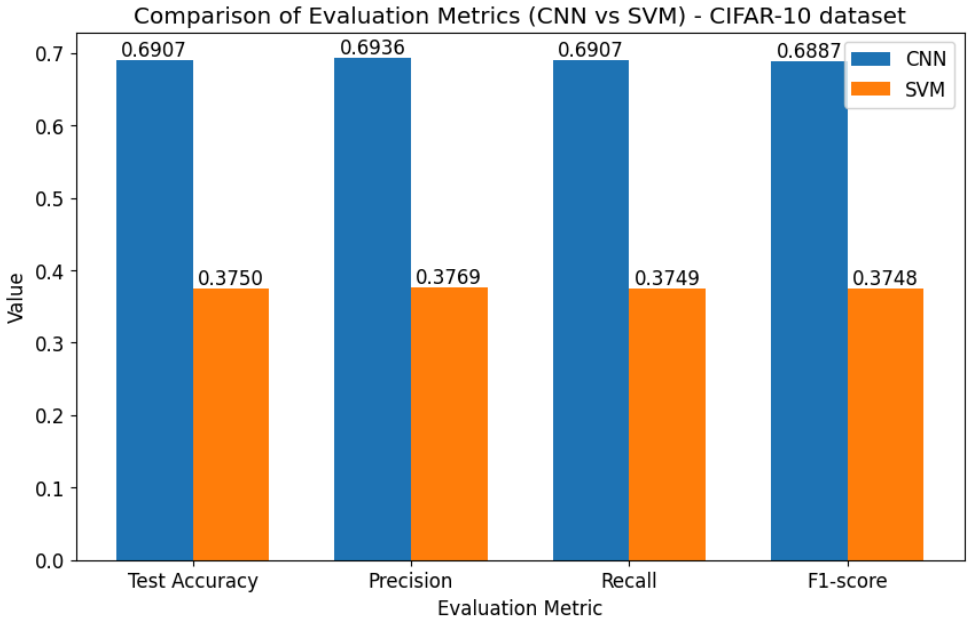
The 10 classes in the CIFAR-10 dataset are:

1. Airplane
2. Automobile
3. Bird
4. Cat
5. Deer
6. Dog
7. Frog
8. Horse
9. Ship
10. Truck

The dataset is divided into two subsets: a training set with 50,000 images and a test set with 10,000 images. Each class is equally represented in both subsets.

CIFAR-10 is widely used for image classification tasks and is a challenging dataset due to its small image size and the complexity of distinguishing between similar classes like cats and dogs or trucks and cars. Researchers and machine learning practitioners often use the CIFAR-10 dataset to evaluate and compare the performance of various image classification models and techniques. It serves as a standard benchmark to test the capabilities of deep learning models, such as convolutional neural networks (CNNs), in solving real-world computer vision problems. The dataset's relatively small size makes it suitable for rapid experimentation and prototyping of new models and algorithms.

SVM vs CNN



Based on the evaluation metrics of the two models (CNN and SVM) using the given dataset, we can observe the following:

Based on the evaluation metrics, we can observe the following comparisons between the CNN and SVM models:

Test Accuracy: The CNN model outperforms the SVM model significantly in terms of test accuracy. The CNN model achieved an accuracy of 69.07%, while the SVM model achieved only 37.5%.

Precision: The CNN model also outperforms the SVM model in terms of precision. The precision of the CNN model is 69.36%, while the SVM model achieved a precision of 37.69%. Higher precision indicates that the CNN model is better at correctly identifying positive samples.

Recall: The CNN model and SVM model have similar recall scores, but the CNN model has a slight edge. The recall of the CNN model is 69.07%, and the SVM model achieved a recall of 37.49%. Higher recall indicates that the CNN model is better at capturing positive samples.

F1-score: The CNN model has a higher F1-score compared to the SVM model. The F1-score of the CNN model is 68.87%, while the SVM model achieved an F1-score of 37.48%. The F1-score takes into account both precision and recall, making it a better metric for comparing models.

In summary, the CNN model significantly outperforms the SVM model in all evaluated metrics, indicating that it is a better model for the given task. The CNN model shows higher accuracy, precision, recall, and F1-score, demonstrating its superior performance in classifying the data. Therefore, we can conclude that the CNN model is the preferred choice between the two for this specific task.