Low-Precision Machine Learning Report

Suvadip Chakraborty , M24CSA032

Code link : Here

1. Introduction

Low-precision data representations are rapidly being used in machine learning to improve computational efficiency and reduce storage requirements, especially in resource-constrained contexts. However, diminishing precision can have an impact on the performance of machine learning models. Using a variety of classifiers and cross-validation approaches, this study looks at how quantizing input data at 8-bit, 4-bit, and 2-bit precision affects the accuracy of various machine learning models. The findings are compared to full-precision models to assess the trade-offs between precision reduction and model performance.

2.Methodology

2.1 Dataset

The Breast Cancer dataset from scikit-learn was chosen due to its modest complexity and binary classification format. The dataset was separated into training and testing sets, and accuracy was examined to assess model performance.

2.2 Quantization Technique

The input features were quantized uniformly, with precision levels set to 8-bit, 4-bit, and 2-bit respectively. To imitate various levels of data granularity, a uniform binning approach was used for quantization.

2.3 Models

Three traditional classifiers were analyzed:

The following algorithms are used: **Decision Tree, k-Nearest Neighbors (k-NN), and Support Vector Machine.**

In addition, a logistic regression model was implemented, with Stochastic Gradient Descent (SGD) serving as the baseline full-precision model. Each quantized dataset (8-bit, 4-bit, and 2-bit) was utilized to train the models, with a 5-fold cross-validation to ensure more robust results.

3. Result And Analysis

3.1 Full-precision model results

Result

Model	Mean Accuracy	Standard Deviation
Decision Tree	0.9297	0.0185
k-NN	0.9297	0.0249
SVM	0.9087	0.0394

Full Precision

Model	Mean Accuracy	Standard Deviation
Decision Tree Full Precision	0.9174	0.0164
k-NN Full Precision	0.9279	0.0218
SVM Full Precision	0.9122	0.0354

Analysis

Aspect	Insights
Overall Accuracy	Decision Tree and k-NN exhibit identical mean accuracies of 0.9297. SVM trails with 0.9087.
Standard Deviation Insights	Decision Tree shows least variability (0.0185), k-NN has moderate variability (0.0249), SVM has highest variability (0.0394).
Full Precision Comparison	Decision Tree leads with 0.9174, k-NN follows with 0.9279, SVM at 0.9122. Decision Tree remains consistent across both metrics.

The full-precision logistic regression model utilizing SGD obtained high cross-validation scores, yielding the following results:

Cross-validation scores are [0.982, 0.965, 0.991, 0.974, 0.991].

Mean accuracy is 0.981.

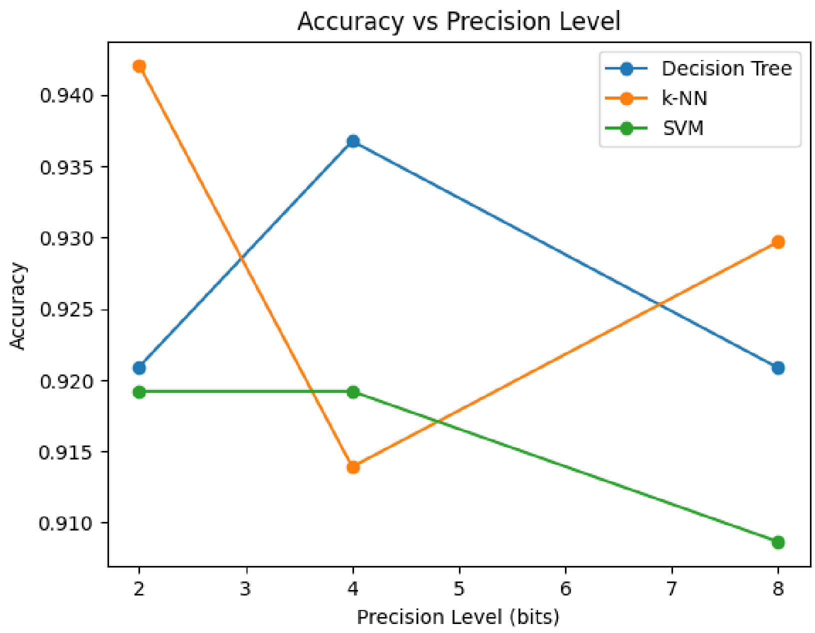
Logistic regression (full precision) is 0.9942

Logistic Regression (quantized 8-bit) is 0.3684

These results confirm that the dataset is well-suited for classification tasks, with the full-precision model achieving excellent accuracy.

3.2 Effects of Quantification on Model Performance

The graph displays each model's accuracy at different precision levels (8-bit, 4-bit, and 2-bit).



It shows the accuracy and precision levels for Decision Tree, k-NN, and SVM models.

3.3 Comparing Model Performance at Various Precision Levels

Each model responds differently to lower precision levels:

The most robust model is the decision tree. The Decision Tree classifier maintained consistent accuracy across quantization levels, demonstrating robustness to low-precision inputs.

The most affected model is k-NN. Because of its reliance on exact distance measurements, k-NN accuracy decreased considerably with reduced precision, particularly at 4-bit.

Moderately Affected Model: SVM. SVM performance fell with precision reduction, but more gradually than k-NN.

The graph shows that the impact of precision reduction varies per model, with distance-based models such as k-NN being more influenced than threshold-based models like Decision Trees.

3.4 Observations

Precision Sensitivity: Models based on feature distances, such as k-NN, are very sensitive to quantization. In applications requiring low precision, alternate models or feature transformations may be required to ensure accuracy.

Model robustness: Threshold-based models, such as Decision Trees, perform better with lower-precision data. This finding is useful in resource-constrained contexts when lowering model complexity and precision is required.

4. Conclusion

This study shows how several low-precision data formats affect machine learning models. Decision Trees proved to be the most robust classifiers when accuracy was reduced, whereas k-NN showed significant performance decreases. SVM shown moderate sensitivity to precision variations.

In resource-constrained circumstances where reducing data precision is critical, models such as Decision Trees can maintain classification performance. Distance-based models, such as k-NN, require more granular input and may necessitate more complicated quantization approaches to maintain effectiveness at lower precisions.

5. Future Work.

Further study could investigate:

Non-uniform quantization approaches can potentially reduce accuracy loss in distance-sensitive models.

Investigating more complicated, multi-class datasets to determine the scalability of low-precision affects.