Natural Image Classification on CIFAR-10

Abhijeet Verma, *B19CSE001* Abhinav Kashyap, *B19CSE002,* and Adityaraje Devade, *B19CSE005*

Abstract

Natural Image Classification on CIFAR-10 dataset.

In Image classification we classify image into one of the predefined classes. In conventional way, different computer vision techniques are used to extract features from images and different machine learning algorithms use these extracted features to classify the images. Apart from various learning algorithms the accuracy and performance of the model mostly depends on the trained dataset and the algorithm used. In this paper we have used CIFAR-10 dataset for image classification. In this paper the proposed approaches for image classification makes essential use of machine learning methods.

The approach is using multiple classifiers and comparing them using cross validation.At last the best prediction and classifier is extracted.The approach is reliable since multiple classifiers are being used on single dataset and are compared appropriately.This method results into higher accuracy but takes ample amount of time.

Index Terms

Convolutional Neural Network,cross validation,data augmentation,principal component analysis.

I. Introduction

THE term image classification refers to the labelling of images into one of the predefined categories. Classification is a task to identify the class/category of new instance based on training set whose classes are known. In Image Classification, we classify an image into one of the predefined classes or multiple classes at the same time. In Multi Label Image classification, an image can have multiple classes present among the set of classes where as in simple Image classification an image contains only one class among the set of classes.

In general supervised learning, an object is represented by an instance (or feature vector) and it is represented with a class label. Initially, let X indicate the instance space (or feature space) and Y is the set of class labels. The task is to learn a function which is f: X $\rightarrow$ Y from a given data set. Even though the above method is existing and it is successful, there are many problems associated with real world where this work does not fit well. A real world problem may be related with a number of instances and labels simultaneously [1]. One of the main difficulties in applying the supervised learning is that we require large amount of training images. It becomes very difficult to label the images as it is expensive and also requires more time. Basically there are two different methods for solving such type of problems. First one is called as the problem transformation methods and other one is algorithm adaptation method. Image classification is one of the most widely studied subject in the field of Machine Learning which has developed many algorithms for it.This work focuses on implementing different algorithms and cross validating them.

Dataset - The CIFAR-10 dataset (Canadian Institute For Advanced Research) is a collection of images that are commonly used to train machine learning and computer vision algorithms. It is one of the most widely used datasets for machine learning research.The CIFAR-10 dataset contains 60,000 32x32 color images in 10 different classes.[3] The 10 different classes represent airplanes, cars, birds, cats, deer, dogs, frogs, horses, ships, and trucks. There are 6,000 images of each class. Computer algorithms for recognizing objects in photos often learn by example. CIFAR-10 is a set of images that can be used to teach a computer how to recognize objects. Since the images in CIFAR-10 are low-resolution (32x32), this dataset can allow researchers to quickly try different algorithms to see what works. Various kinds of convolutional neural networks tend to be the best at recognizing the images in CIFAR-10.CIFAR-10 is a labeled subset of the 80 million tiny images dataset.

By training the the dataset with four different classifiers their individual accuracies are calculated on both training and testing data.Cross validation is performed and the predictions are compared.

A. Preprocessing

The Cifar10 data has been preprocessed before approaching the classifiers. The preprocessing used includes conversion of int to float since for efficient training, the neural network inputs should be normalised to a specific roughly unit range (-1.0 to 1.0) or to mean 0, standard deviation 1.0. Both of these require float representation. The data is normalized where the values are centered around the mean with a unit standard deviation. This means that the mean of the attribute becomes zero and the resultant distribution has a unit standard deviation. One hot label encoding is used for converting each text category for the machine to process them using mathematical equations.

B. Classifiers Used

The classifiers used to train the given dataset are Logistic Regression,Random Forest Classifier, Bayes Classifiction and Convlolutional Neural Network(CNN).

1) Logistic Regression: Logistic regression is one of the most important analytic tools in the social and natural sciences. In natural language processing, logistic regression is the baseline supervised machine learning algorithm for classification, and also has a very close relationship with neural networks.In multiclass case like this one, the training algorithm uses the multinomial logistic regression.Multinomial Logistic Regression is the regression analysis to conduct when the dependent variable is nominal with more than two levels. Similar to multiple linear regression, the multinomial regression is a predictive analysis. Multinomial regression is used to explain the relationship between one nominal dependent variable and one or more independent variables.

Standard linear regression requires the dependent variable to be measured on a continuous (interval or ratio) scale. Binary logistic regression assumes that the dependent variable is a stochastic event. The dependent variable describes the outcome of this stochastic event with a density function. A cut point can be used to determine which outcome is predicted by the model based on the values of the predictors.Object-based classification considers not only the individual pixels but also neighboring pixels for grouping individual pixels into information classes. With this approach, spatially contiguous pixels are usually first grouped into spectrally homogeneous objects, and then classification is conducted with objects as the minimum processing units.The optimized and inbuilt function LogisticRegression() of sklearn library of python is used to train the dataset.

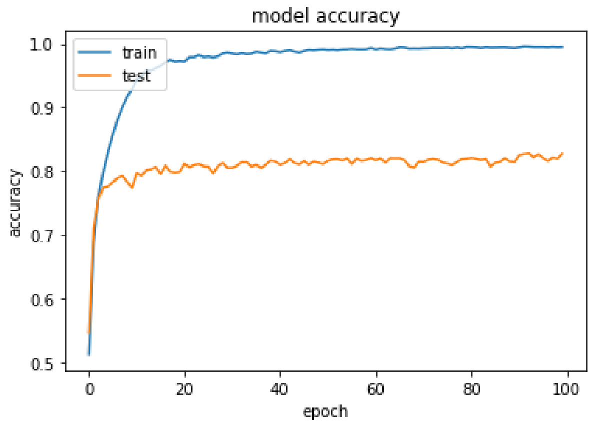
The testing and training accuracies using logistic regression are found to be Without PCA =¿Training Accuracy : 0.46326 Testing Accuracy : 0.3971 With PCA =¿ Training Accuracy : 0.4276 Testing Accuracy : 0.2251

2) Random Forest: Random Forest consists of a large number of individual decision trees that operate as an ensemble. Each individual tree in the random forest spits out a class prediction and the class with the most votes becomes our model's prediction. The random forest is used due to its advantage of using a large number of relatively uncorrelated models operating as a committee will outperform any of the individual constituent models. It fits a number of decision tree classifies on various sub samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The testing and training accuracies using Random Forest are found to be -

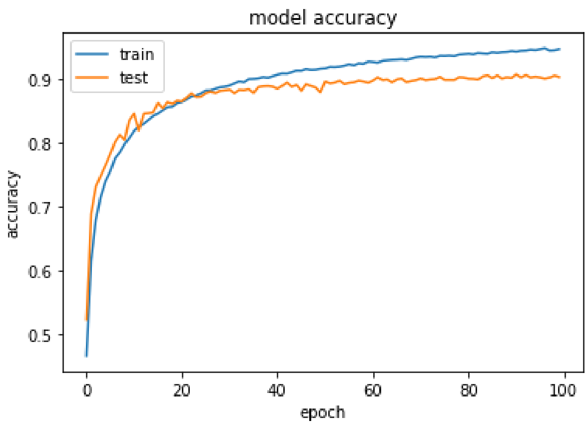
Without PCA =¿ Training Accuracy : 1 Testing Accuracy : 0.4705 With PCA =¿ Training Accuracy : 1 Testing Accuracy : 0.287

3) Gaussian Naive Bayes: The Naive Bayes classifier is based on Bayes' theorem of probability. In Bayes' theorem, the conditional probability that an event belongs to a class can be calculated from the conditional probabilities of finding particular events in each class and the unconditional probability of the event in each class. That is, for given data, , and classes, where denotes a random variable, the conditional probability that an event belongs to a class can be calculated by using the following equation: $P(c_k|\boldsymbol{x}) = P(c_k)\frac{P(\boldsymbol{x}|c_k)}{P(x)}$ Equation (1) shows that the calculation of P(c—x) is a pattern classification problem since it finds the probability that the given data belongs to class and we can decide the optimum class by choosing the class with the highest probability among all possible classes, , which can minimize the classification error. For doing so, we need to estimate P(x—c) and assume that any particular value of vector conditional on is statistically independent of each dimension can be written as follows - $P(\boldsymbol{x}|c_k) = \prod_{i=0}^{n} P(\boldsymbol{x}|c_k)$ By combining equation (1) and equation (2), the Naive Bayes classifier can be summarized as the following equation - $k = argmax_k\, P(c_k)\prod_{i=0}^{n} P(x_i|c_k)$ The inbuilt naive bayes classifier is implemented and accuracies are found to be - Without PCA - Training Accuracy : 0.28472 Testing Accuracy : 0.2955 With PCA - Training Accuracy : 0.31382 Testing Accuracy : 0.2309 We see that the accuracy is not upto expectations and hence we proceed towards next classifier.

4) CNN: A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. It is used to develop a robust test harness for estimating the performance model. In the project CNN is used with and without augmentation. The hypertuning is done by increasing the number of Conv2D layers to build a deeper model, also increased number of filters to learn more features, also added Dropout for regularization and also added more Dense layers. Accuracies : without Augmentation : Training : 0.9993 Testing : 0.8883 with Augmentation : Training : 0.9025 Testing : 0.9850



Without Image Augmentation,



With Image Augmentation

C. Feature Extraction : PCA

PCA is a dimensionality reduction that identifies important relationships in our data, transforms the existing data based on these relationships, and then quantifies the importance of these relationships so we can keep the most important relationships and drop the others. PCA was applied to all the models escept CNN. It was obeserved that PCA was useful in increasing test accuracy in the case of Naive Bayes Classifier.

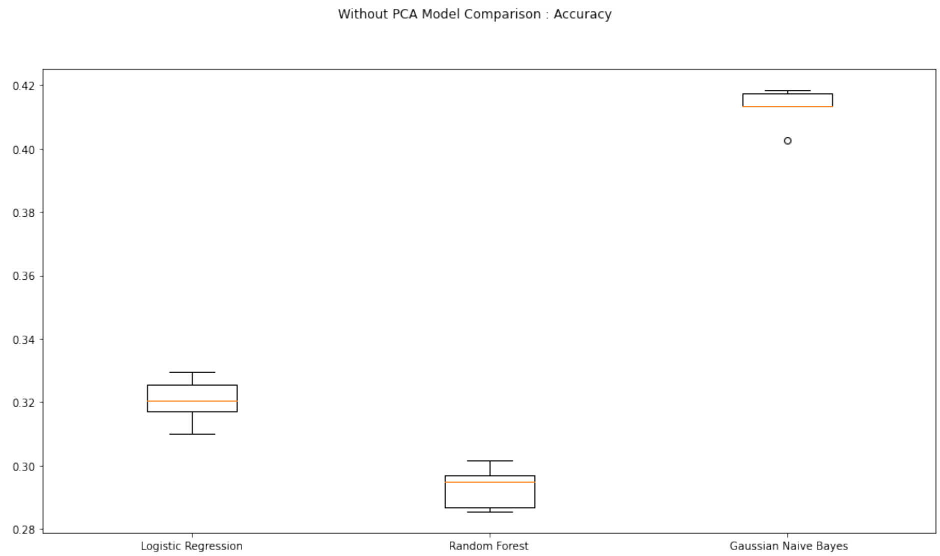
D. Comparison of models and Cross Validation

The fires three classifiers that are logistic regression, random forest and Gaussian Naive Bayes are cross validated using k-fold cross validation by keeping k as 5. The result obtained for accuracies when cross validated without PCA was - Logistic Regression CVS accuracy Scores: [0.31 0.3255 0.3295 0.317 0.3205] Accuracy: Mean: 0.321 , std deviation 0.008

Gaussian Naive Bayes CVS accuracy Scores: [0.2855 0.287 0.297 0.295 0.3015] Accuracy: Mean: 0.293 , std deviation 0.007

Random Forest CVS accuracy Scores: [0.4025 0.4135 0.4175 0.4185 0.4135] Accuracy: Mean: 0.413 , std deviation 0.006.

The boxplot for the same -

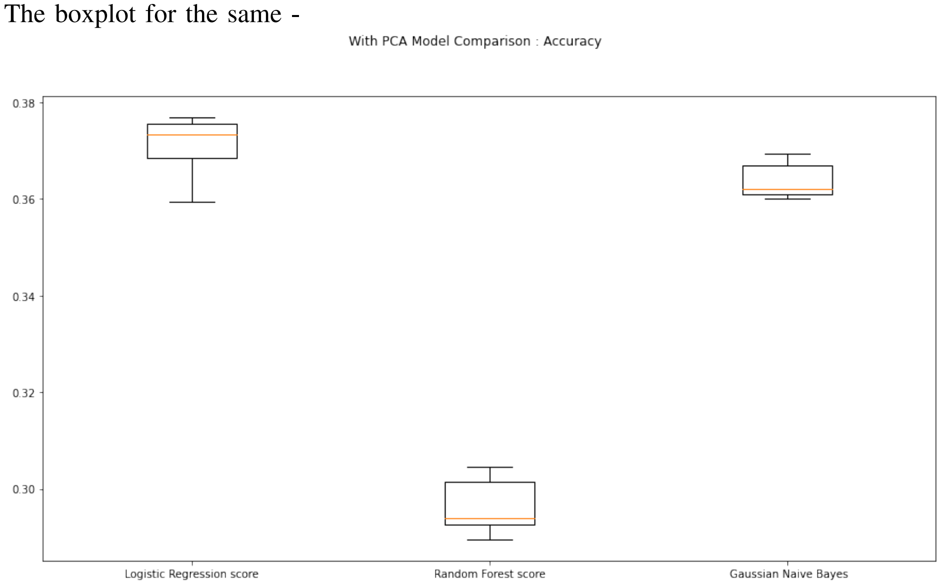


The result obtained for accuracies when cross validated with PCA was - Logistic Regression CVS accuracy Scores: [0.3685 0.3735 0.377 0.3755 0.3595] Accuracy: Mean: 0.371 , std deviation 0.007

Gaussian Naive Bayes CVS accuracy Scores: [0.2895 0.2925 0.3045 0.294 0.3015] Accuracy: Mean: 0.296 , std deviation 0.006

Random Forest CVS accuracy Scores: [0.367 0.361 0.362 0.3695 0.36] Accuracy: Mean: 0.364 , std deviation 0.004

The boxplot for the same -



Hence we notice that the accuracy of Gaussian naive bayes classifier is higher than both logistic regression and random forest.After applying PCA feature extraction we see higher change in the accuracy of logistic regression.

II. Conclusion

All four classifiers are used to train the dataset and the predictions based on these were obtained successfully.The comparison was done among 3 classifiers using cross validation and the box plots were plotted for the same.PCA technique was used for feature extraction and dimensionality reduction. We conclude that the CNN classifier with data augmentation gives the highest accuracy on both training and testing data and hence it is best among all used.

III. Contribution

Abhijeet Verma(B19CSE001) - Data Augmentation,Logistic Regression and Random Forest Classifier
Abhinav Kashyap(B19CSE002) - SVM and PCA
Adityaraje Devade(B19CSE005) - CNN

The report was made by combined efforts of all members.

Acknowledgment

The authors would like to thank Dr.Richa Singh and Dr.Yashaswi Verma for proper guidance throughout the journey of this project.

References

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2323, 1998.
[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks."
[3] "CIFAR-10 and CIFAR-100 datasets." [Online]. Available: https://www.cs.toronto.edu/ kriz/cifar.html. [Accessed: 20-Oct-2019].
[4] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014