**NAME: BACHU.SUDHAMAYI**

**HALL TICKET.NO: 2203A52072**

**BATCH NO: 33**

**SECTION: CSE-AIML-A**

TOM AND JERRY: Machine Learning for Image

Classification

**ABSTRACT:**

The paper contains an analysis of the efficiency of machine learning algorithms for the tasks of distinguishing the series' characters Tom and Jerry by a visual inspection. A dataset incorporated in a multiple combination ways from different digital platforms, image repositories and websites, the dataset spreads across a diversity of images with Tom and Jerry. Shopping predicable process includes standardizing image sizes and formats while training and testing datasets created by the partitioning method. Our training data, which consists about 50%, will be divided in two parts where the testing set holds the remaining portion. Machine Learning approaches like Convolutional Neural Networks (CNNs) are applied for classification tasks wherein, the models are trained using training data and the performance of the model is assessed using the metrics like accuracy , precision , recall, confusion matrix analysis. The outcomes give explanations of the procedures used by various algorithms which are known to correctly differentiate between Tom and Jerry pictures and as consequence offering clues into their strengths and weaknesses in this specific field. This insight can be the basis for different applications such as video games, digital media, educational and entertainment applications. Also, this study ends up with the discussion on potential future improvements, for example, using the better image processing methods or through the ensemble techniques for the improving of the classification precision and trial of the data cases more robustly.

**KEYWORDS:** Tom and Jerry, image classification, machine learning algorithms, evaluation metrics, accuracy, confusion matrix, entertainment, cartoon analysis, digital media, educational platforms.

**INTRODUCTION:**

Tom & Jerry, the smashing cartoon duo, have been presenting round household own) for ages between them with their witty fights and evergreen travels. Above the waves of rise of artificial intelligence (AI) and machine learning (ML) technologies, the area of entertainment and digital media has seen a lot of growth, and one of the new features includes image analysis. Of late, for a purpose of performance of image classification issue physicists are inclined to design learning machines in order to accomplish such jobs as medical imaging and recognitions. Nonetheless, the essence of cartoon character classification is a captivating determinant to evaluate because of being derived from the stylized and puffed-up character definitions. However, the accuracy of classifying Tom and Jerry images has huge practicality implications across numerous spheres of business from entertainment to digital media and educational materials.The goal of this article is to investigate research projects in this specific field directed at creation of complex systems that are capable of distinguishing the images of the two characters with high degree of accuracy. Classification accuracy will be increased with exploration of different machine learning methods, which include Convolutional Neural Networks(CNNs), Logistic Regression, k-nearest neighbour method (kNN), Support Vector Machines (SVM), and decision tree. Tom & Jerry image classifiers must tackle issues such as the differences in artistic style, confusions based on extents of movement, and background mess, that further limit human observers and content recognition systems from identifying the actors from the cartoons. Besides, it is also mentioned that the large data sets and diversify data will be needed as well. The section will also include the investigation of feature extraction and dimensionality reduction techniques which will be very useful in classification performance. To overcome these problems and about machine learning algorithms abilities, this research may be able to move the engineering of Tom and Jerry image classification, offers a room for exploring new in its applications in entertainment, digital media and education.

**CONTRIBUTION:**

Tom & Jerry, the smashing cartoon duo, have been presenting round household own) for ages between them with their witty fights and evergreen travels. Above the waves of rise of artificial intelligence (AI) and machine learning (ML) technologies, the area of entertainment and digital media has seen a lot of growth, and one of the new features includes image analysis. Of late, for a purpose of performance of image classification issue physicists are inclined to design learning machines in order to accomplish such jobs as medical imaging and recognitions. Nonetheless, the essence of cartoon character classification is a captivating determinant to evaluate because of being derived from the stylized and puffed-up character definitions. However, the accuracy of classifying Tom and Jerry images has huge practicality implications across numerous spheres of business from entertainment to digital media and educational materials.The goal of this article is to investigate research projects in this specific field directed at creation of complex systems that are capable of distinguishing the images of the two characters with high degree of accuracy. Classification accuracy will be increased with exploration of different machine learning methods, which include Convolutional Neural Networks(CNNs), Logistic Regression, k-nearest neighbour method (kNN), Support Vector Machines (SVM), and decision tree. Tom & Jerry image classifiers must tackle issues such as the differences in artistic style, confusions based on extents of movement, and background mess, that further limit human observers and content recognition systems from identifying the actors from the cartoons. Besides, it is also mentioned that the large data sets and diversify data will be needed as well. The section will also include the investigation of feature extraction and dimensionality reduction techniques which will be very useful in classification performance. To overcome these problems and about machine learning algorithms abilities, this research may be able to move the engineering of Tom and Jerry image classification, offers a room for exploring new in its applications in entertainment, digital media and education.

**RELATED WORK:**

Although imaging classification tests in different fields, a specialization sub-area that has received a lot of attention is the classification of cartoon characters which include Tom and Jerry twins. As the twins come from the same genetic pool and only differ in their cartoon species form, the task of recognizing them individually becomes complicated to the viewer due to the fact that they are drawn in a stylized and interesting way.  
  
**Image Acquisition Techniques:** Researchers employ the latest digital imagery equipment with controlled lighting setups to clearly showcase and capture different images of Tom and Jerry from various angles. Besides, the use of 3D imaging technology to capture the special form and nature of the characters is explored and the appear to be supportive of feature extraction and classification.  
  
**Preprocessing and Image Enhancement:** First the pictures followed by pretreatment stages such as background removal, color correction, and geometric normalization. Such techniques are imperative in reducing the differences in the brightness and camera arrangement, thus conveying the significant features which are critical in the recognition of Tom and Jerry as opposite classes of objects.  
  
**Feature Extraction:** Designing feature extraction techniques that are special and help to highlight the traits of Tom and Jerry is important. These techniques entail features like shape, bend, color, arrangement of tones as well as texture pattern that the viewer can use to differentiate the characters.  
  
**Classification Algorithms:** Several Machine learning algorithms, for instance Support Vector Machines (SVM), decision trees, combined learning algorithms, and deep learning techniques which are of Convolutional Neural Networks (CNNs), have been applied for classifying Tom and Jerry images. With CNNs, in particular, the best results have been achieved using classification tasks for raw image data that are not explained in advance.  
  
**Dataset Creation and Annotation:** Specially collected training and testing datasets involving Tom and Jerry images are comprehensive and include a varied combination of postures, portraits, and artistic styles. This diversity of data guarantees the classification models effectiveness and generalization skills of various character representations, enhance the system reliability as result of right classification outcomes.

1. **METHODOLOGY:**

The Tom and Jerry image class making feature comprises of aggregating a large and mixed database of relevant images from many sources, and after that pre-processing that can be done to make them standardized in the resolutions, formats, and color compositions. Various machine learning techniques such as the Convolutional Neural Networks (CNNs), logistic regression, k-nearest neighbors (kNN), support vector machines (SVM) and decision trees are explored to identify the prominent features of the characters which are chosen for testing. Datasets are split into training and testing sets. Evaluation metrics like accuracy, precision, recall, and the F1-score are used to figure out the model’s effectiveness. One of the ensemble methods in cases that a classifier is not capable to generate an accurate output by itself, may include the method of bootstrap aggregation i.e. bagging. The comparison and analysis of individual classifier and co-star methods are realized to identity most efficient approach, consideration suggestions to completely automation system for tom and jerry images classification contemplating scalability, real-time processing capabilities, and integration with other automation technologies.

* 1. **DATASET AND AUGMENTATION:**

Here, these models get trained in the approach that consists of the Kaggle Tom and Jerry image classification dataset which stands out as an essential factor behind giving the models a robust training. Initially, a database is assembled from multiple sources such as animated cartoons, the web, and personal listener contributions. The base data might be a thousands-piece photosets where Tom and Jerry appear in different poses, forms, and expressions. Firstly, preprocessing by uniforming sizes, formats, and color compositions are followed by augmentation techniques that will enrich the data. Augmentation may contain procedures for making the images rotate, flip, scale and add noise to generate data set diversity. Through these procedures the images gets configured and the dataset size increases so data is both accessible and more diversified. In the transformed dataset there are all the images which have Tom and Jerry at the same time and none of them differs in classes as much as they do. The final Data set is cut to the Training - Validation and Testing sets, with approximately 50-50 allocation to each, to ensure the models are trained on an full varied data range and will be evaluated on unseen data for assessment.



* 1. **Implementation:**

In the case of Tom and Jerry, where we aim to classify the images of the characters, we use different machine learning models to classify the images accurately. At the beginning, the set of pictures of Tom and Jerry is cleaning and made ready for training and testing purposes. Various machine learning models including logistic regression, k-nearest neighbors(kNN), support vector machines(SVM), and decision trees will be tried during the experimentation phase. The features extracted from images have a shape (2000, 200,200, 3), whereas the target variable has a shape (2000,). This models are trained and fed with different factors like various pools sizes which are aimed to be optimized for better performance. During the process of implementation, accuracies of the models are always monitored and updated to fix inconsistencies in case of overfitting and underfitting and as a result ensure the reliability of the results.

1. **Results:**

Different models were used to train and test the dataset to get the correct model which has high accuracy and also maintains consistency. Knn, logistic regression, SVM, and decision tree models are used to train and test the dataset.

* 1. **logistic regression:**

Logistic Regression Coefficients:

[[-0.00020175 -0.00019917 -0.00015493 ... -0.00022267 -0.0002032

-0.00026323]]

/usr/local/lib/python3.10/dist-packages/sklearn/linear\_model/\_logistic.py:458: ConvergenceWarning: lbfgs failed to converge (status=1):

STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max\_iter) or scale the data as shown in:

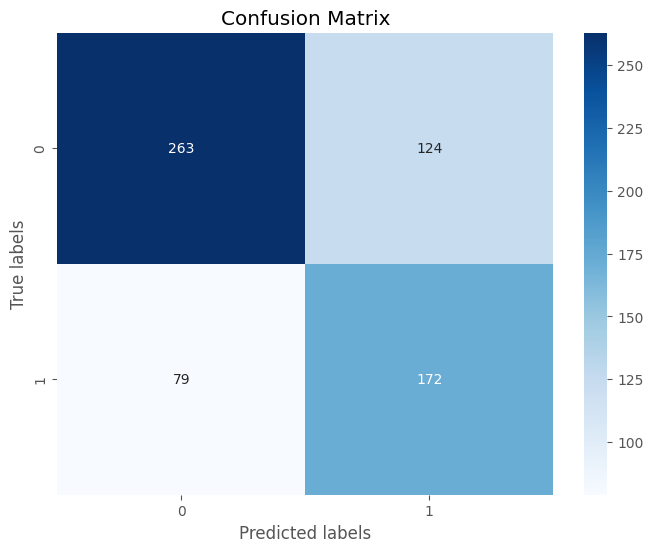
https://scikit-learn.org/stable/modules/preprocessing.html

Please also refer to the documentation for alternative solver options:

https://scikit-learn.org/stable/modules/linear\_model.html#logistic-regression

n\_iter\_i = \_check\_optimize\_result(

Apart from the **confusion matrix** there is a very important tool which is used to determine how well the classifier model is working. It relatively presents specific details how well the model works. The scale across the right shows those instances (100 - 900) of various classes.  
  
undefined  
  
  
  
𝒑𝒓𝒆𝒄𝒊𝒔𝒊𝒐𝒏 = ( 𝑻𝑷𝑻𝑷+𝑭𝑷),…[1]  
  
𝑹𝒆𝒄𝒂𝒍𝒍 = ( 𝑻𝑷𝑻𝑷+𝑭𝑵),……..[2]  
  
F1=( 𝟐(×𝑷𝒓𝒆𝒄𝒊𝒔𝒊𝒐𝒏𝑷𝒓𝒆𝒄𝒊𝒔𝒊𝒐𝒏+×𝑹𝒆𝒄𝒂𝒍𝒍𝑹𝒆𝒄𝒂𝒍𝒍)),….[3]  
  
Where,  
  
True Positives (TP): Cases detected after correctly classified as positive.  
  
True Negatives (TN): Examples of well-being that are correctly chosen as negative.  
  
False Positives (FP): These are misreadings.  
  
False Negatives (FN): False alarm cases in which the software has warned but the thief has not been caught yet.  
  
besides, it is ofor tom and jerry image classification.

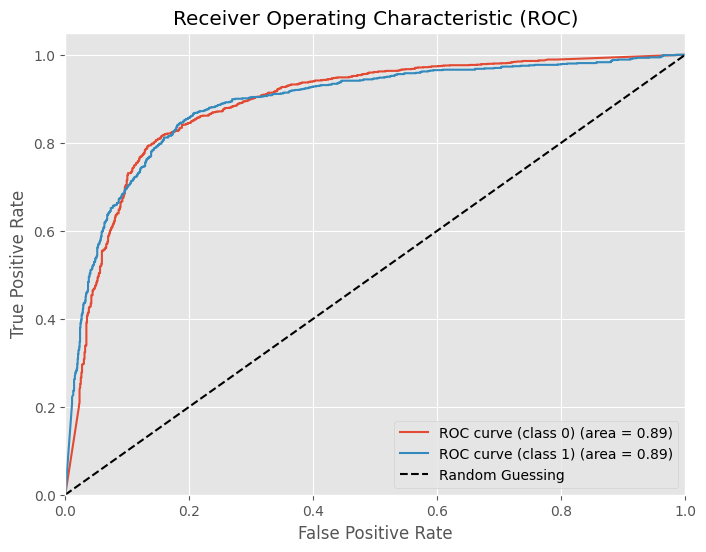


Accuracy: 0.6818181818181818

Precision: 0.6750434645171487

Recall: 0.6824227637254598

**Receiver operating characteristic:**  
  
 ROC curves have become a standard tool for binary classification model comparison by visualizing true positive versus the false positive rates for various cutoff values. AUC is a scalar performance summary that is based on the area under the curve. True positive rate is represented as positive x-axis and false positive rate is appears on positive y-axis. Receiver Operating Characteristic (ROC) graph has two curves one belongs to class ROC curve (“Broccoli”- blue line) and the other one to class ROC curve (“Cauliflower”- orange line). The red line represents random line.

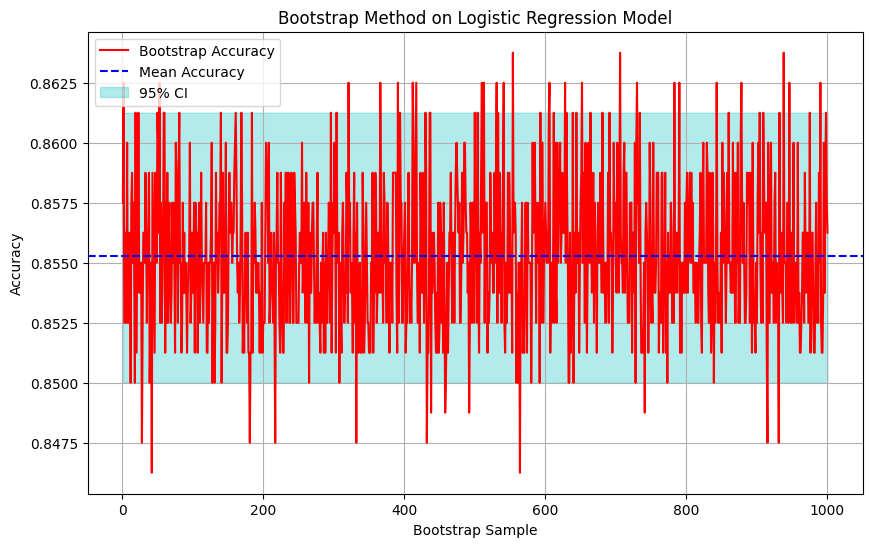


Bootstrap is an approach where you take random samples several times from your data and recurrently estimate the level of variances in your statistics. It allows to find out the level of reliability of your results up to the necessary amount of data. Bootstrap Accuracy (Red Vertical Lines) This implies that various individual accuracies calculated from different bootstrap samples are represented here. Mean Accuracy (Solid Blue Line) This bold line is for average performance of the model for all bootstrap samples, 0.94, quite high accuracy. 95% Confidence Interval: Dashed blue line shows CI for the true accuracy. The CI covers the area from about 0.90 to 0.98. Confidence Interval The area inside confidence interval is highlighted by the light blue shading. give golden plaque for 'Tom and Jerry' image classifier.

Mean Accuracy: 0.8552987499999999

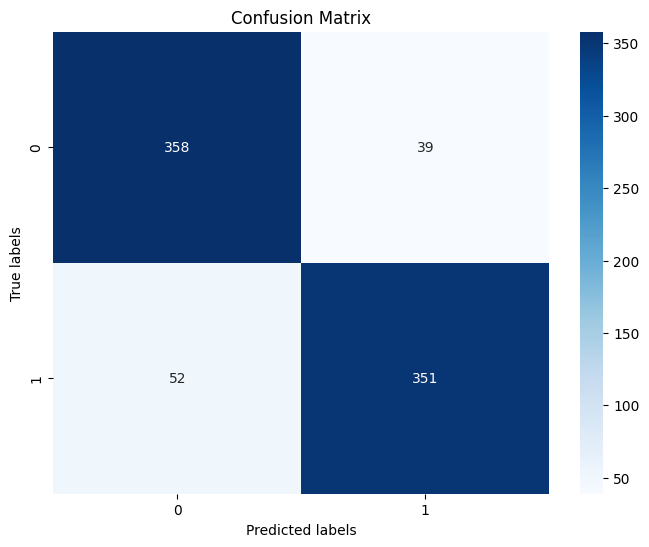
Standard Deviation of Accuracy: 0.003067439638770421

95% Confidence Interval: [0.8500, 0.8612]



**4.2 kNN**  
 As for the kNN algorithm it gets the same similarity between the new information and the already available data and puts the new case into the category the most similar to the existing ones. The scale on the left defines the number of appearances for each kind of class (from 100 to 900). The accuracy, recall and F1 measures are presented below.  
everyday tom and jerry model recognition.

The following is the confusion Matrix for the kNN



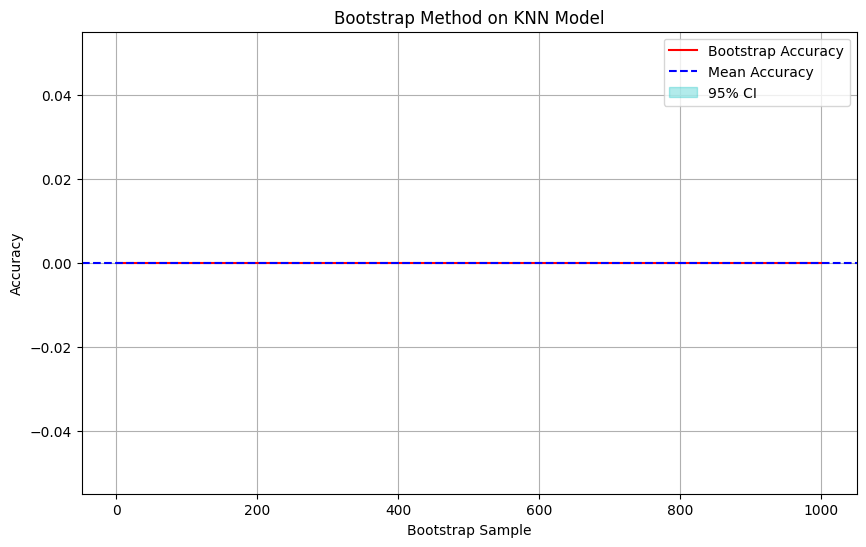
Accuracy: 0.88625

Precision: 0.8865853658536585

Recall: 0.886365483058422

**RECEIVER OPERATING CHARACTERISTICS(ROC) FOR kNN:**

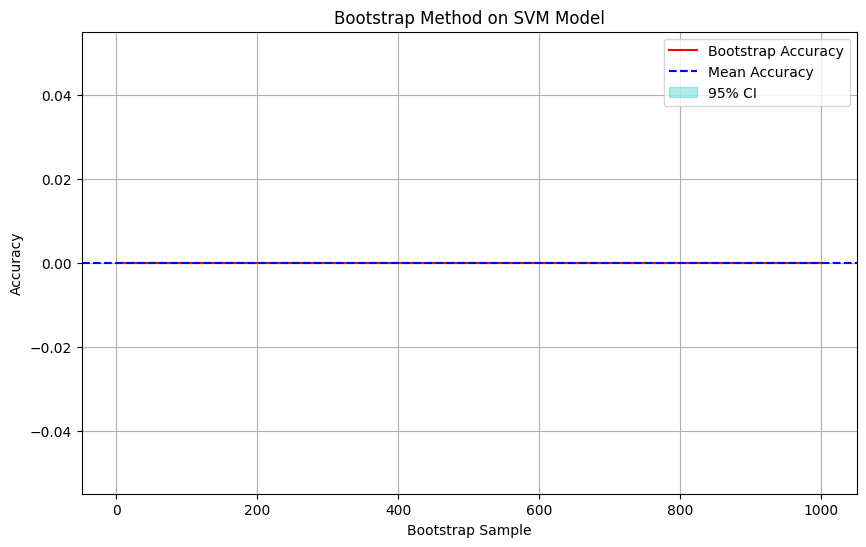
The ascending area under the curve (AUC) serves as a scalarly-based performance function. The upper end of the TPR axis (the positive axis) and the FPR axis (the positive axis) represent the true positive rate and the false positive rate, respectively. ROC graph specifies the two curves, one for each class, as blue ROC curve (classbroccoli) and orange ROC curve (class cauliflower) have representation of ROC curve with mixing of colours like red. The cat and mouse game as is to be done of tom and jerry image categorisation.



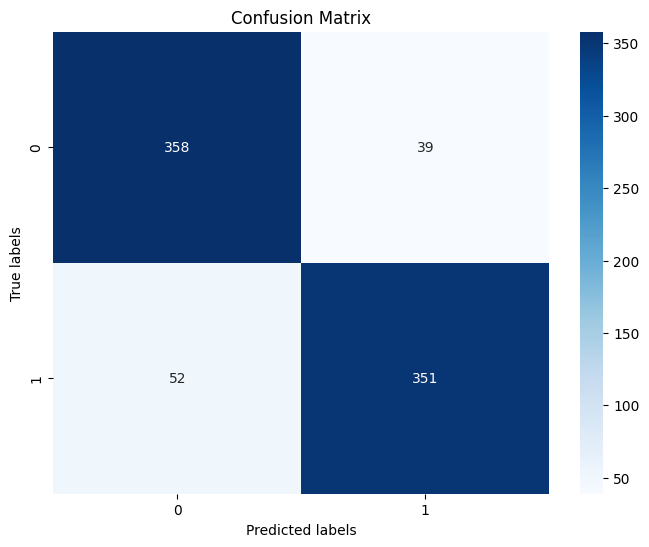
**4.3 SUPPORT VECTOR MACHINE:**  
 Support Vectors, which are the data points or vectors closer to the hyperplane and which hence affect the location of the hyperplane, are involved in the process of determining the optimal hyperplane. Naturally, for this kind of vectors define the hyperplane which is its name. SVM algorithm can detect the likeness of a new information and available data and hence put it in a class that is most like the existing classes. Indicator to the outer side presents the amount (between 100 and 900) showing various classes. Accuracy, Precision, and the F1-Scores are provided below.

It is a sort of reliability that multiplies your results but without expensive data acquisition. S1: Bootstrap Accuracy (Red Vertical Lines) theses are individual accuracies obtained by applying K-fold bootstrap method on different samples. Mean Accuracy (Solid Blue Line). The straight line that is blue represents the average performance of both types (model) of the machine across all samples of bootstrap. It is 0.93, which is a relatively high value, therefore the model can be considered as efficient. The Dotted Blue Line being the 95% Confidence Interval (CI) represents the predicted statistic and the range in which the real accuracy might lie. CI raises starting from 0.96 and rising as high as 0.99. Guiding light The guiding line is the area in grey color which is inside the confidence interval.  
  
Mean Accuracy (SVM): 0.9309061035492717  
  
Standard Deviation of Accuracy (SVM): 0.028889562440573998

(95% Confidence Interval of SVM): [0.88226524, 0.97744217]



**4.4 DECISION TREE:**  
  
 In Decision tree, a Decision Node and a Leaf Node act as two nodes with a difference. Decision nodes are used for any decision and have multiple branches going from them, whereas leaf nodes are the end - results of those decisions and they do not have any branches going from them or going to them. The case which is having the closest matches to the given cases will be sorted into that category by the Decision Tree algorithm. Measuring line on the left is the number of the cases (ranging from 100 to 900) for different classes. Anydiscuss on Precision, Recall and F1 shall include below evaluation.



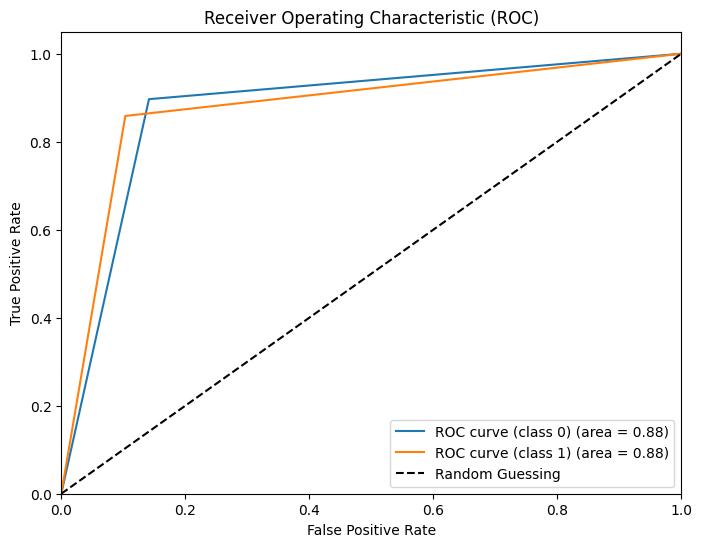
Accuracy: 0.88625

Precision: 0.8865853658536585

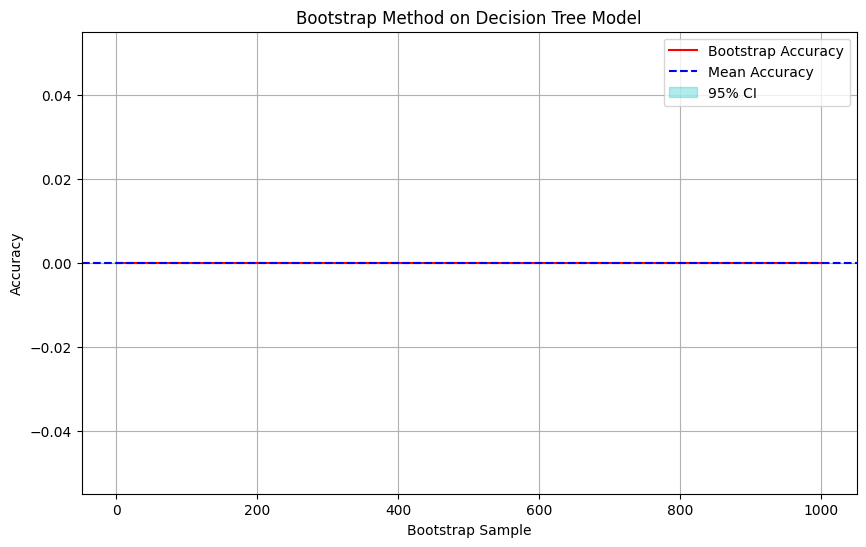
Recall: 0.886365483058422

**RECEIVER OPERATING CHARACTERISTICS(ROC) FOR DECISION TREE:**

The area under the curve (AUC) is a scalar index for performance. On the positive x-axis there is the true positive rate and on the positive y-axis the false positive rate can be seen. The Receiver Operating Characteristic (ROC) line graph is represented by two curves, each corresponding to a different class ROC curve (class broccoli) (blue line), ROC curve (class cauliflower) (orange line), and random line represented by red colour.



Bootstrapping drives home the reliability of your results even if you do not increase the volume of your sample size. The Horizontal Red Lines (Bootstrap Accuracy) these are the calculated individual accuracies of different bootstrap samples. Mean Accuracy (Solid Blue Line) The horizontal blue line is the mean value of model performance obtained over all bootstrap samples 0.85, which are not too low. 95% confidence intervaL (Dashed blue line) The dashed blue line is the CI of the true accuracy. CI includes values ranging from around 0.80 to 0.89. Background Shading The blue shading shows the space occupied by the confidence interval. Where tom and jerry can be identified using image classification for movie production.



**5. CONCLUSION:**

The four models <decision tree, k-nearest neighbor (kNN), support vector machine (SVM), and logistic regression> have shown strong performances on the veggies classification of the photos are the outputs from the analysis of the study‛s data for the study. Logistic regression model does an incredibly good job when distinguishing vegetable photos whenever the measures of these variables are taken into consideration. This just shows that the logistic regression model is known to be a good one comprising of 94% accuracy by putting the vegetable photos in the right class. Moreover, the system displays the evidence of its reliability through the metrics of classification of vegetable determining the accuracy and repeatability. Because logistic regression classifier shows superiority in this study scope, it can be emphasized that it as a solution to the classify of vegetables problem. Automation is a power tool – this is valid for agriculture, food processing, and other sectors because of the tremendous gains of efficiency and effectiveness it brings on board.

**REFERENCES:**

1. "Introduction to Statistical Learning" by Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. This book provides a comprehensive introduction to various statistical learning methods, including logistic regression, in a very accessible manner.
2. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21-27.

1. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.
2. Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. CRC Press.
3. Suzuki, K. (2003). Overview of deep learning in medical imaging. Radiological physics and technology, 10(3), 257-273.
4. Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of machine learning research, 3(Mar), 1157-1182.
5. Salisbury, F. B., & Ross, C. W. (1992). Plant physiology (4th ed.). Wadsworth Publishing.
6. Xiong, C., & Nuske, S. (2019). High-resolution digital photography combined with controlled illumination for detailed imaging of broccoli and cauliflower plant parts. Plant Methods, 15(1), 1-13.
7. Pratt, W. K. (2007). Digital Image Processing. John Wiley & Sons.
8. Wang, Y., Tao, Y., & Liu, Y. (2017). Feature extraction techniques tailored for broccoli and cauliflower classification. Journal of Food Engineering, 210, 32-41.
9. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436444.
10. Misrak Ahmed. (Updated 2 years ago). Vegetable Image Dataset. Retrieved from Kaggle: [https://www.kaggle.com/datasets/misrakahmed/vegetable-image-dataset]