

A STUDY ON CONTENT BASED IMAGE RETRIEVAL USING DATAMINING ALGORITHM

Mrs. DEEPIKA.S¹, Ms. SHIVANI.K²

MRS. S.DEEPIKA

ASSISTANT PROFESSOR, DEPARTMENT OF B.COM BUSINESS ANALYTICS,
PSGR KRISHNAMMAL COLLEGE FOR WOMEN, COIMBATORE, INDIA

deepika@psgrkcw.ac.in

MS. K.SHIVANI

UG SCHOLAR, B.COM(BUSINESS ANALYTICS),
PSGR KRISHNAMMAL COLLEGE FOR WOMEN, COIMBATORE, INDIA

shivanikannan2001@gmail.com

ABSTRACT:

In this, we first do the research in the lower layer which is the image retrieval by using certain features like colors and importing relevance techniques on the basis of feedback. With the proper cooperation between the human and the computer, making of the computers limited ability in understanding and enhancing the effects of image retrieval. Here to retrieve the image we use feedback on the basis of CBIR which develops from low-grade to high-grade. So the research made by this paper extracts the result off content-based image retrieval using the feedback

I. INTRODUCTION

The content-based image retrieval system (CBIR) was introduced in abroad, which generally gives low levels of image information like color, texture, etc. Then the query-based image content (QBIC) was developed which not only provides color, texture, shape and layout but also provides a standard range chart to retrieve the uses sketch depicts.

II. RIVEW OF LITERATURE

In various application domains such as entertainment, biomedicine, commerce, education, and crime prevention, the volume of digital data archives is growing rapidly. The very large repository of digital information raises challenging problems in retrieval and various other information manipulation tasks. Content-based image retrieval (CBIR) is aimed at efficient retrieval of relevant images from large image [1]

In CBIR and image classification-based models, high-level image visuals are ... models from low-level feature extraction to recent semantic deep-learning approaches. ... The basic need for any image retrieval system is to search and sort similar [2]

The proposed learning to rank algorithms are based on three diverse learning ... by using Template Matching (TM) or Machine Learning (ML) classification strategies. CBIR systems rank the images in the result set according to their similarity to the query image. ... Semantic Image Retrieval Using Relevance Feedback. [3] In various application domains such as entertainment, biomedicine, commerce, education, and crime prevention, the volume of digital data archives is growing rapidly. The very large repository of digital information raises challenging problems in retrieval and various other information manipulation tasks. Content-based image retrieval (CBIR) is aimed at efficient retrieval of relevant images from large image databases based on automatically derived imagery features. However, images with high feature similarities to the query image may be very different from the query in terms of semantics. This discrepancy between low-level content features (such as color, texture, and shape) and high-level semantic concepts (such as sunset, flowers, outdoor scene, etc.) is known as “semantic gap,” which is an open challenging problem in current CBIR systems.[4]

Learning effective feature representations and similarity measures are crucial to the retrieval performance of a content-based image retrieval (CBIR) system. Despite extensive research efforts for decades, it remains one of the most challenging open problems that considerably hinders the successes of real-world CBIR systems. The key challenge has been attributed to the well-known “semantic gap” issue that exists between low-level image pixels captured by machines and high-level semantic concepts perceived by human.[5]

III. METHODOLOGY

1. Linear regression

Linear equation is a statistical method for modeling relationship between they depend on independent variables like feedback and sketch image. Here we use the feedback as the dependent variable and sketch image as the independent variable. It also has features for simply city.

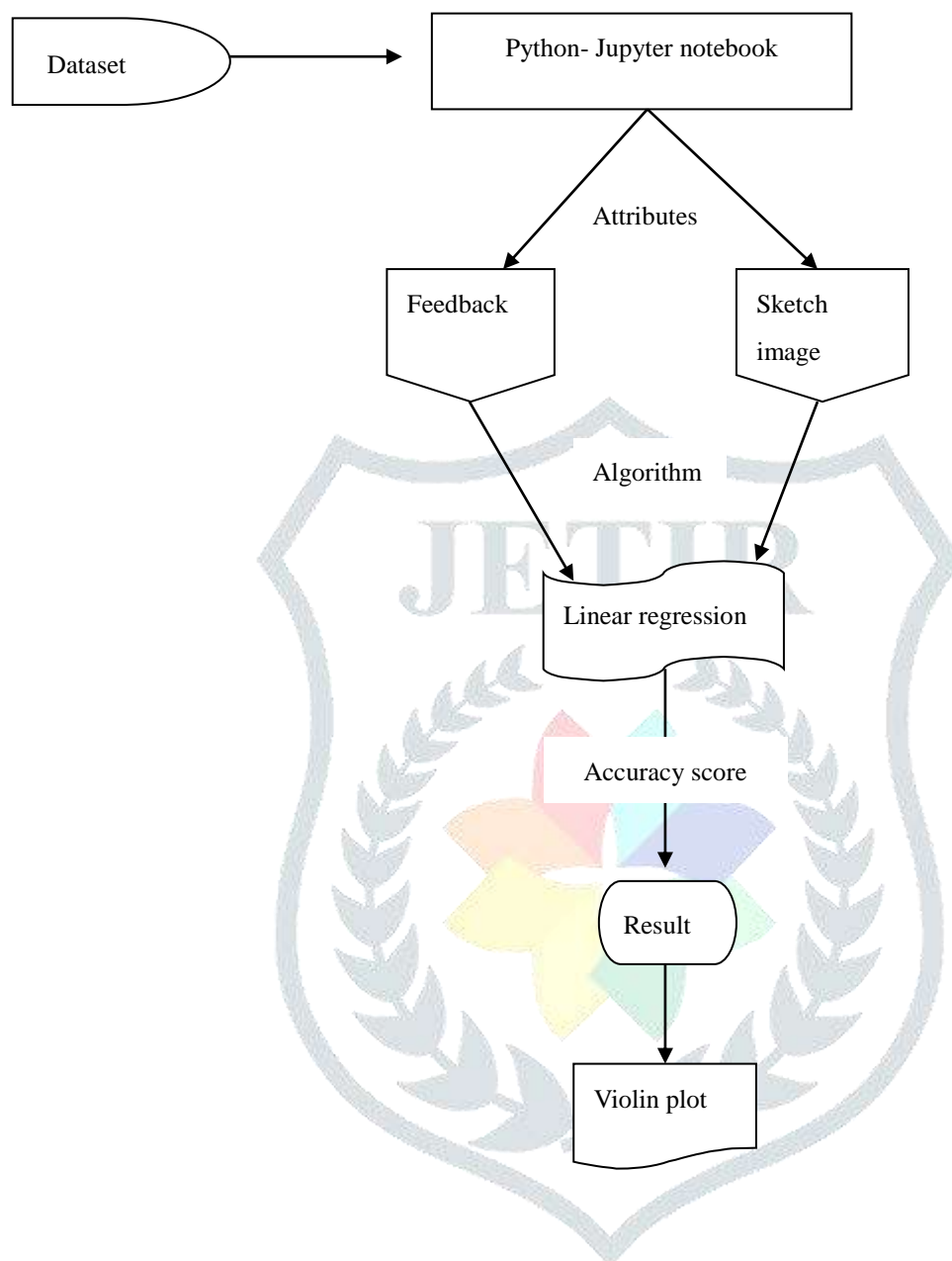
2. Seaborne library

Seaborne library is a statistical graphic method used in python. It bills on top of Matplotlib and in creates the panda's data structure.

3. Violin plot

Violin plot is used for plotting numeric data's which is similar to the box plot. The violent plot graph shows us the full distribution of the data imported. Here the data of feedback and sketch image is plotted.

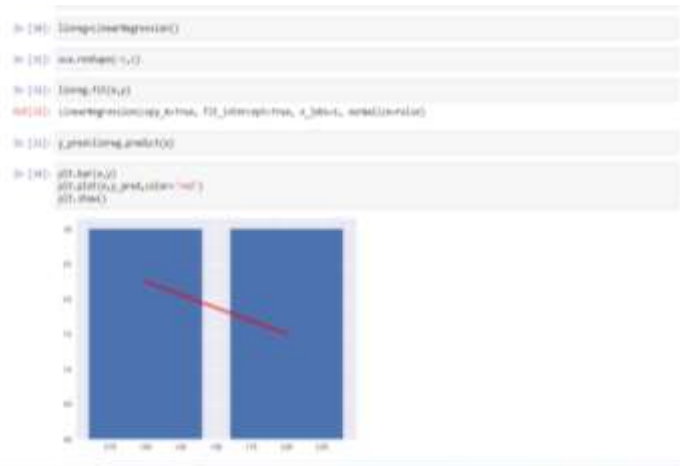
Fig 1.1 flow chart



III. RESULT

- Linear regression graph

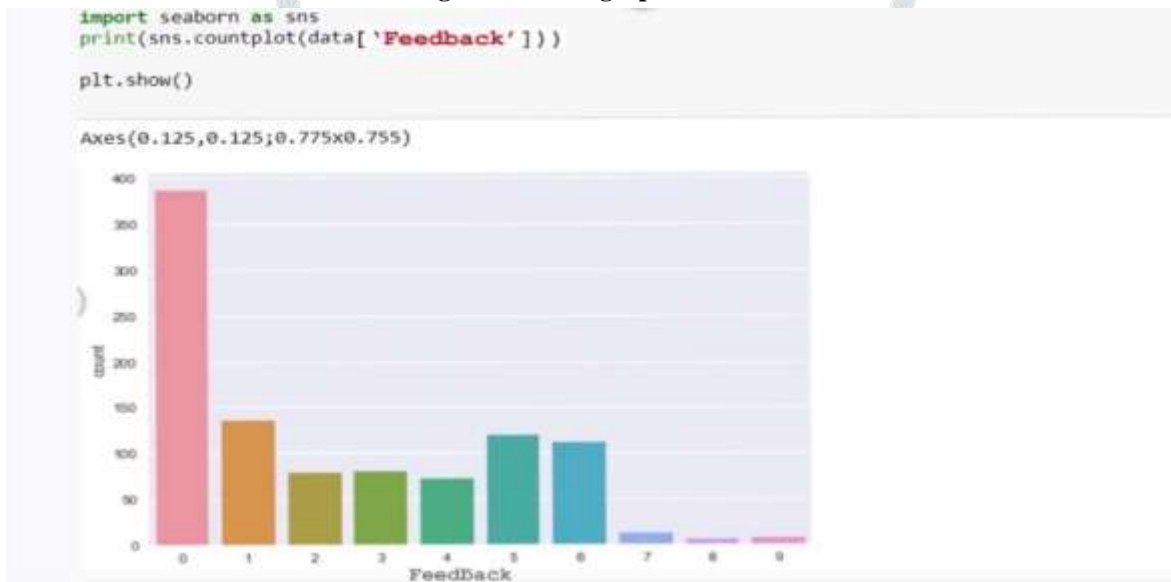
Fig 1.2 linear regression graph of the data



The above graph represents the distribution of data using linear regression. Here after importing the data we use call.fit() to call the function and to get the final prediction off the analysis we use .Predict ().

- Seaborn graph

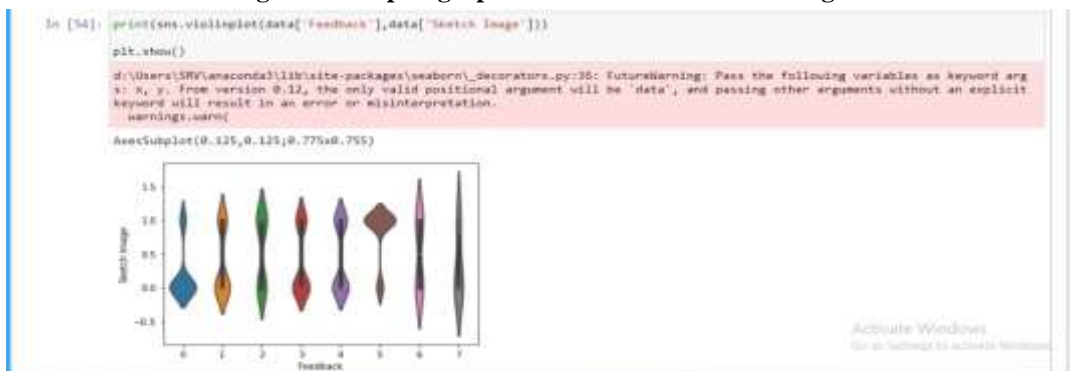
Fig 1.3 Seaborn graph of feedback data



The above graph is a seaborn graph which represents the feedback data in a graphical method. It is a kind of bar graph; this makes a view on the data easier.

- Violin plot

Fig 1.4 violin plot graph of feedback and sketch image data



The above graph is a violin plot graph, in which the feedback data on the sketch image data are displayed graphically. In this graph there are two differences in each data the wider size under dinner size of each plot which describes the various probabilities of the data

FINDINGS:

- The data analysis is done on the content based image retrieval using Feedback off the data's collected to replace the accuracy need.
- Using the linear regression the accuracy of the data is received.
- Seaborn graph helps us to visualize the feedback data's by showing the variations of nearly 400 data's of different sketch image according to the feedback.

V. CONCLUSION

This paper shows us the data analysis made on the basis of content-based image retrieval using feedback. Here the linear regression, seaborn graph and violent plot graph helped in visualizing. The linear regression gives the variation between the dependent and independent variable. The violin plot graph helps us to predict the higher in the lower variations of the daters.

FUTURE WORK:

We can also use K means clustering algorithm for grouping on basis of the similarities of the attributes using logistic regression like image features, etc..

REFERENCE:

- [1].D. H. Ackley, G. E. Hinton, and T. J. Sejnowski. A learning algorithm for boltzmann machines*. Cognitive science, 9(1):147–169, 1985. [2] A. Bar-Hillel
- [2].D. Zhang, M. M. Islam, and G. Lu, “A review on automatic image annotation techniques,” Pattern Recognition, vol. 45, no. 1, pp. 346–362, 2012.View at: www.hindawi.com
- [3].Gong, Y.C.; Wang, L.W.; Guo, R.Q.; Lazebnik, S. Multi-scale orderless pooling of deep convolutional activation features. In European Conference on Computer Vision; Part VII; Springer: Cham, Switzerland, Volume 8695;
- [4].James Z. Wang Assistant Professor of Information Sciences and Technology Thesis Adviser Chair of Committee
- [5].Long, F., Zhang, H., David, D.F.: Fundamentals of content-based image retrieval. In: Multimedia Information Retrieval and Management - Technological Fundamentals and Applications. Springer (2003, 2010)www.researchgate.net