Hacettepe University Department of Computer Engineering

BBM 409: Introduction to Machine Learning Lab.

Name: Denizkaan Araci

Number: 21426606

Subject: Classification of Bird Species from Images and Attributes

Advisors: Aykut Erdem, Levent Karacan, Tuğba Erdoğan

1.Introduction

In this project, we implement k-nearest neighbour algorithm to predict and classify the given bird data. We have 5033 image with 200 different classes, bird attributes, segmentations and bounding box coordinates.

2.About k-Nearest Neighbour

When implementing this algorithm, i used normal k-NN and weighted k-NN.

I created test and train data with k-fold cross validation method. I picked k = 10 then data set tested ten times with different test and train data.

In k-NN i picked first 21 neighbour. Because larger we make k, the accuracy of prediction is increasing. Larger values of k reduce the effect of noise on the classification.

3.Data Analysis

After making test and train data for image color histogram, image attributes, we can get %40-45 of correct prediction rate.

First of all, using k-NN and weighted k-NN with k-fold cross validation in bird attributes, each prediction has between %40-%60 correction rate.

Mean accuracy of bird attributes with k-fold cross validation is 47.395

Even though masking and boxing operation in image, we couldn’t get get high predicton rate. Image color histogram prediction has between %5-%10 correction rate.