Will Eye Gaze Show Your Focus: Using Fuzzy Cluster to Predict if People Can Identify Images are Manipulated or Not

Zongge Ren

Research School of Computer Science, Australian National University
u5431907@anu.edu.au

Abstract. To determine if people can be predicted on which images they are looking at by collect their eye gaze tracking data, I use their eye gaze data as input of a fuzzy clustering model and try to predict which images they are looking at, manipulated image or unmanipulated image. Then I compare the prediction with the ground truth. However, I find that the accuracy is just above 60% which can not be regarded as a good result.

Keywords: Prediction; fuzzy clustering model; eye gaze tracking

1 Introduction

Image is a significant way for human to know the world. “A picture is worth a thousand words” is an English adage which shows that an image can include a large amount of information and spread the information of a high efficiency. A person may look through plenty of images per day. Some applications track users’ eye gaze to find what attracts their attention (Stiefelhagen, et al., 1997). Figuring out which images they are looking at can better understand what they are interested in. I use the data of a research paper “Imperfect Understandings: A Grounded Theory And Eye Gaze Investigation Of Human Perceptions Of Manipulated And Unmanipulated Digital Images” (Caldwell, et al., 2015). This research let 80 volunteer participants view 14 manipulated and unmanipulated images (in the dataset, there are just 5 images of them), and capture their eye gaze tracking data, which images they looked at and verbal response on if they can perceive manipulated images. Gosain and Dahiya claims that fuzzy clustering is an important technique which can partition the dataset, and it is used in plenty of fields. Therefore, I will use a fuzzy clustering model and K-means clustering model to learn on eye gaze tracking data and predict which images they looked at and verbal response and finally comparing the prediction with the ground truth to get the accuracy.

2 Method

2.1 Data description

The dataset is from “Imperfect Understandings: A Grounded Theory And Eye Gaze Investigation Of Human Perceptions Of Manipulated And Unmanipulated Digital Images”. (Caldwell, et al., 2015) There are 8 columns in the dataset.

1 Participant: the id number of the participant
2 Num_fixs: the total number of fixations by the participant when looking at the image
3 Fixs_dur: the total amount of time (in seconds) that the participant spent looking at the image
4 Num_man_fixs: the total number of fixations by the participant when looking within the target area
5 Man_fixs_dur: the total amount of time (in seconds) that the participant spent looking within the target area
6 Image: the id number of the image
7 Image_manipulated: this is whether the image the participant views is the manipulated or unmanipulated version 0 = unmanipulated, 1 = manipulated
8 Vote: the verbal opinion of the participant as to whether the image is manipulated or unmanipulated, 0 = voted

unmanipulated, 1 = voted manipulated, 2 = do not know

The initial input of the model is 2,3,4,5 and the output of the model is the prediction of 7 and 8. But later I want to find if 2,3 is important enough for the model, therefore I will get rid of 2 or 3 or both of them to train the model and see if there exist significant changes on the accuracy.

2.2 Data preprocessing

In order to get all attribute weighted evenly, I normalize the input value. The method I use is min-max normalization and the formula is as below:

$X'_i = (X_i - \min)(new\_max-new\_min)/(max-min)+new\_min$

I set new_max = 1 and new_min = 0, the all the input value will range from 0 to 1.

2.3 Fuzzy clustering

I use cluster.cmeans() function from skyfuzz to set up the model. I define 2 clusters because the ground truth contains only 1 and 0. I found that in the dataset, the group of who claims the image is manipulated has obvious high num_man_fixs. Therefore, after clustering, I calculated the average num_man_fixs of the 2 groups and define the group with higher average num_man_fixs as 1-class and the group with lower average num_man_fixs as 0-class. And I will change the m value to see its influence on the prediction. The result will show in part 3.

2.4 K-means clustering

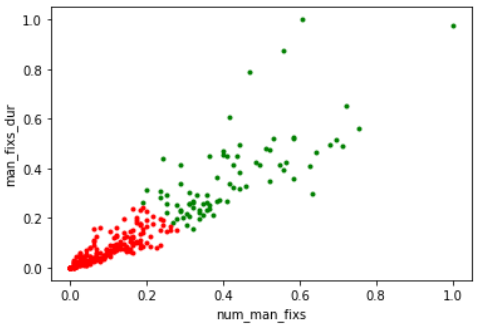
I use KMeans() function from sklearn.cluster to set up the model. I also define 2 clusters and define the group with higher average num_man_fixs as 1-class and the group with lower average num_man_fixs as 0-class. K-means is not the essential part, it is just used to compare with fuzzy clustering. The result will show in part 3.

3 Results and Discussion

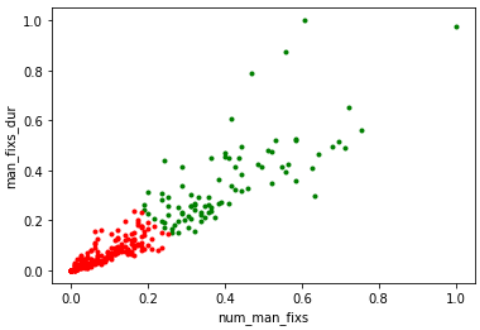
3.1 Result of fuzzy clustering

input	m	average accuracy
4,5	2	0.650
3,4,5	2	0.653
2,4,5	2	0.658
2,3,4,5	2	0.653
2,3,4,5	3	0.656
2,3,4,5	4	0.656
2,3,4,5	5	0.653

Input=4,5 m=2



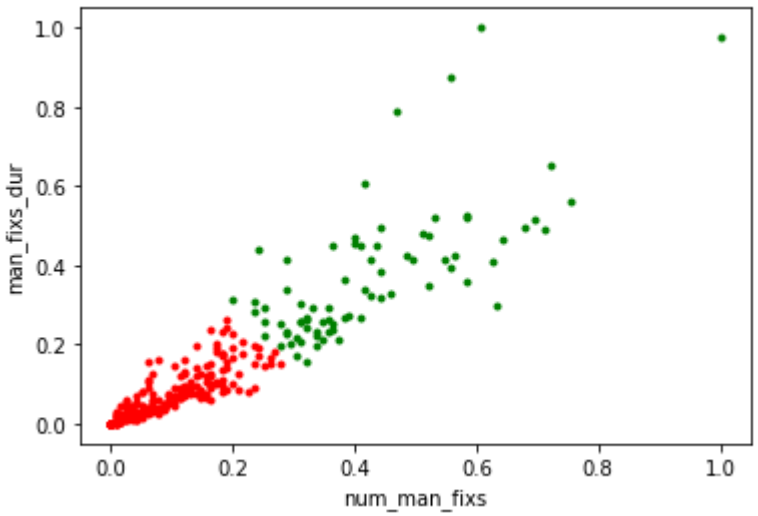
Input=4,5 m=3



3.2 Result of K-means clustering

input	average accuracy
4,5	0.645
3,4,5	0.642
2,3,4,5	0.653

Input=4,5



3.3 Discussion

It can be observed that my research has just around 65% prediction accuracy. These are the reasons which might have influence on the accuracy.

a) Using unsupervised machine learning algorithm to predict data with labels is not appropriate. The number of clusters has been defined in the beginning which will be a limitation. We cannot choose the best number of clusters.

Number of clusters	Fuzzy partition coefficient
2	0.748
3	0.651
4	0.581
5	0.521

Fortunately, in this case, 2 clusters are the best.

b) The definition on which class is 1 and which class is 0 depends on the average value of num_man_fixs, because I found that group of who claims the image is manipulated has obvious high num_man_fixs. This is true in the dataset. However, this might be a coincidence and makes no sense. Moreover, in the prediction, the group has higher average num_man_fixs might not be 1-class, because the group is not exactly the same as the group in the ground truth.

4 Conclusion and Future Work

4.1 Conclusion

The following analysis results are just relative to this research.

a) Fuzzy clustering model performs better than K-means clustering model, but no significant difference.
b) The result shows that changing the parameters will influence the accuracy. I change one parameter per training process to see how it influences on the accuracy. Some changes make the accuracy better and some make it worth. However, the combination of several positive changes might not lead to a better accuracy.
c) Num_fixs and Fixs_dur leave few influences on the prediction. Accuracy might not reduce a lot or even get better when I get rid of these two inputs. I assume that they might lead to overfitting.
d) It is not suitable to use unsupervised algorithm to predict labeled data.
e) The prediction does change a lot when m value become larger.

4.2 Future Work

The accuracy of these two models (fuzzy clustering and k-means) are not high. I do not a find an exact same research to compare the result. The major failure of this research is that I used unsupervised algorithms to predict labeled data. I cannot give a reasonable definition of each cluster and the number of clusters is limited. In the future, another supervised method will be applied on the dataset and we will see if it will perform better.

5 Reference

Caldwell, S., Gedeon, T., Jones, R. & Copeland, L., 2015. Imperfect Understandings: A Grounded Theory And Eye Gaze Investigation Of Human Perceptions Of Manipulated And Unmanipulated Digital Images. *Proceedings of the World Congress on Electrical Engineering and Computer Systems and Science*, 13 7.

Gosain, A. & Dahiya, S., 2016. Performance Analysis of Various Fuzzy Clustering Algorithms: A Review. *Procedia*, 14 3.

Stiefelhagen, R., Yang, J. & Waibel, A., 1997. *Tracking Eyes and Monitoring Eye Gaze,* s.l.: University of Karlsruhe.