Final Project Report

Akshay Iyengar and Graham Trogdon

November 21, 2020

https://github.com/iyengarakshay00/COMP562

1 Abstract

As the importance of data visualization–knowing how to read graphics intended to convey information about data, truly understanding what they mean, and how certain choices in visualization affect the audience's understanding–increases, it is necessary to address what this situations means for those with disabilities and their relationship with the typical paradigm of visual consumption of data. Whether one has a visual impairment, or a learning disability, many of the new demands on the modern worker have the potential to pose as an obstacle to integration into the labor force or, in some cases, normal life (i.e. the data visualization heavy nature of political elections). In this study, we explore the effectiveness of standard machine learning techniques in classifying different types of graphs for the purposes of laying the foundation for deeper analysis on individual chart types so that accessibility technology (such as screen readers) might evolve to consume charts and graphs in the future.

2 Introduction

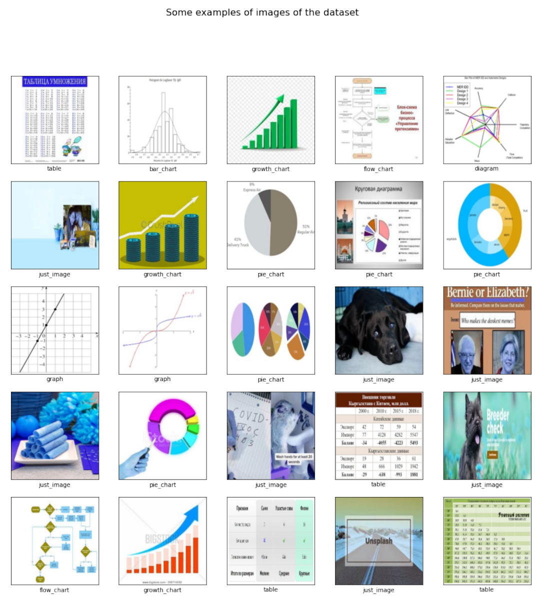
People with disabilities have a history of being disadvantaged with regards to the labor market, and for that reason struggle when it comes to obtaining and maintaining employment. Within recent years, employement rates reached a low in 2011 at 36.3% and peaked in 2017 at 44.2%[1]. As technology continues to embed itself in human production to a greater degree, we reach a crossroads where we can either exclude all except a few people with disabilities from this technology, or we can, with some effort, work towards adapting many of these new technologies to people with different needs, thereby making the gains in production that is a result of technology accessible to all.

3 Data Source

We selected a dataset comprised over over 15,000 images, where each image could be of one of eight classes. Each image could be one of the following; a simple, non-data related image, a bar chart, a diagram, a flow chart, a (line) graph, a growth chart, a pie chart, a table

The following is a subset of the images and their labels as an example:

[1] McDonnall, Michele C., and Zhen Sui. "Employment and Unemployment Rates of People Who Are Blind or Visually Impaired: Estimates from Multiple Sources." Journal of Visual Impairment & Blindness 113, no. 6 (November 2019): 481–92. https://doi.org/10.1177/0145482X19887620.



4 Methods

To begin, we needed to preprocess and label the dataset. We read each image using OpenCV and labeled each image with a number between 0 to 7, where 0 represented a standard image and 1-7 represented one of the 7 types of graphs. A resize transform (courtesy of albumentations) was also applied to each image such that each image became the same size (224x224).

A random 70/30 split was performed on the dataset to partition the training set and the testing set. Then, for each of the sets, the images and labels were split into two respective sets.

We used 7 models throughout our investigation:

1. A logistic regression to perform a binary classification between simple images and information graphics (label 0 vs. 1-7).

2. A naive Bayes model that also performed a binary classification between simple images and information graphics (lael 0 vs. 1-7)

3. A naive Bayes model that performed multiclass classification and tried to discern which chart type out of the 8 categories an image was.

4. A Support Vector Machine (SVM) with the polynomial kernel running as a binary classifier

5. A Support Vector Machine (SVM) with the polynomial kernel running as a multiclass classifier

6. A Support Vector Machine (SVM) with the rbf kernel running as a binary classifier

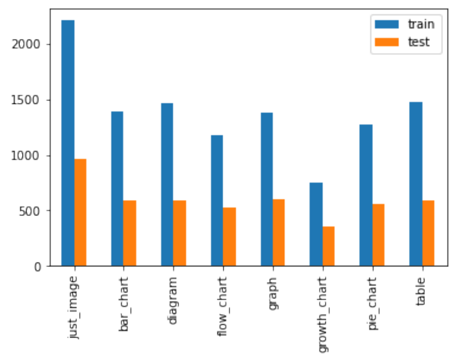
7. A Support Vector Machine (SVM) with the rbf kernel running as a multiclass classifier

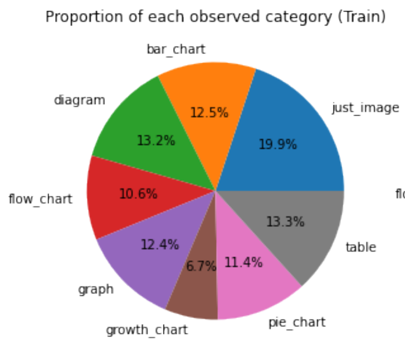
Our images came in a 4D array of shape (no. of images, image height, image width, 3), so in order to train and test on the above models, the data was modified before training such that the RGB values of each datapoint were averaged (to produce a greyscale effect), and the width and height were flattened into a single array so that the training data was a 2D array of the shape (no. images, image height * image width).

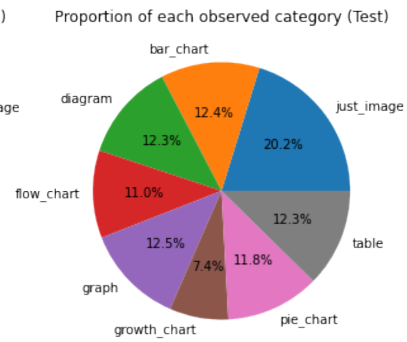
5 Results

5.1 Data Exploration

The following is a visual breakdown of training vs testing set quantities as well as of how much of each type of image comprised the training set and the testing set.:





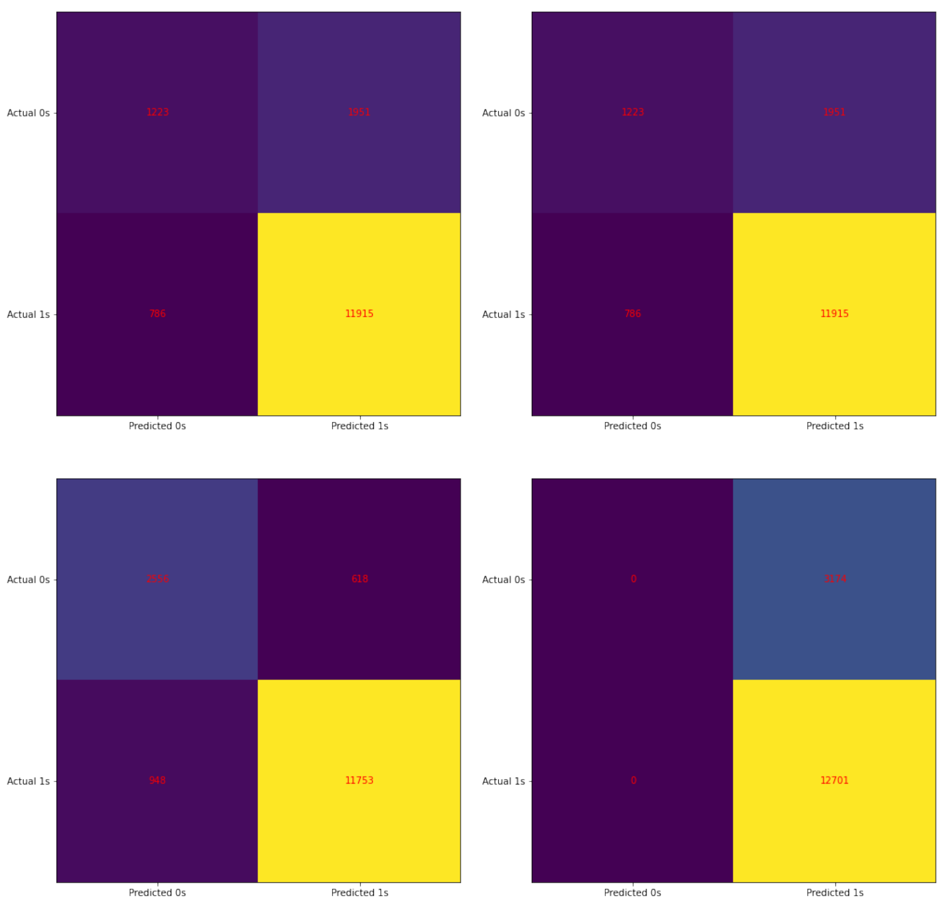


5.2 Scores

Model	Success
Logistic Regression	0.814
Naive Bayes (Binary)	0.896
Naive Bayes (Multiclass)	0.264
SVM (poly) (Binary)	0.904
SVM (poly) (Multiclass)	0.522
SVM (rbf) (Binary)	0.778
SVM (rbf) (Multiclass)	0.32

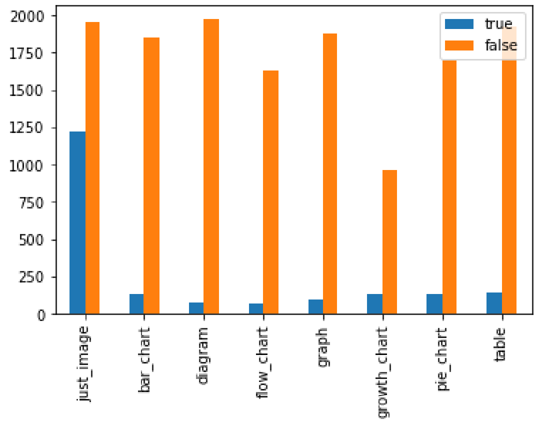
5.3 Confusion Matrices

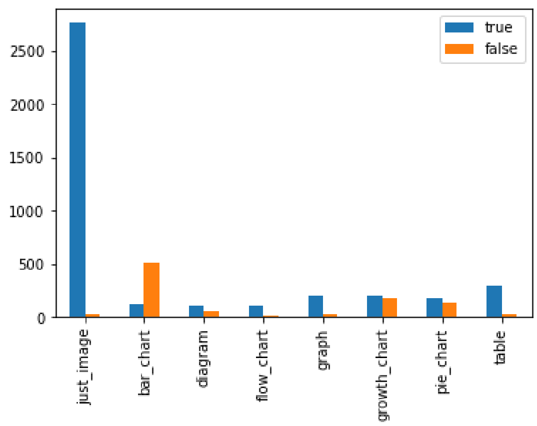
Confusion Matrices: Logistic Regression (top left), Naive Bayes (top right), SVM polynomial (bottom left), SVM rbf (bottom right)

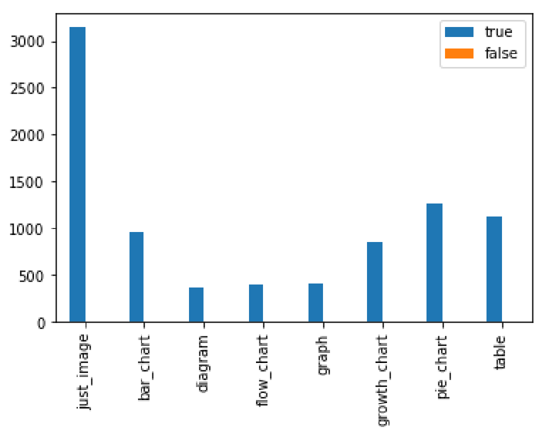


5.4 ConfusionConfusion Matrices: Logistic Regression (top left), Naive Bayes (top right), SVM polynomial (bottom left), SVM rbf (bottom right)

Confusion Matrices represented as bar charts to compare which were truthfully labeled and which were falsely labeled as each category. Top is Naive Bayes multiclass, Center is SVM Poly multiclass, and the last is SVM RBF multiclass.







6 Conclusion

Overall, our models performed fairly well, considering the approach was limited in its sophistication. Getting over 50% on the multiclass classification from an SVM with a polynomial kernel was exciting, because it indicates that machine learning and computer vision are most definitely useful tools in the pursuit of accessiblity in technology.

7 References

McDonnall, Michele C., and Zhen Sui. “Employment and Unemployment Rates of People Who Are Blind or Visually Impaired: Estimates from Multiple Sources.” Journal of Visual Impairment & Blindness 113, no. 6 (November 2019): 481–92. https://doi.org/10.1177/0145482X19887620.