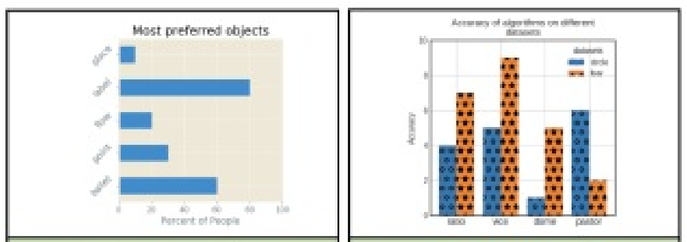
Automated text-translation and data visualization using Generative Adversarial Networks (GANs)

Interns: Samya Bose, Rahul raaghav A. **Mentors**: Aditya Panwar, Andrea Furtado, Kalind Karia.

Abstract (DVQA)

1. Charts and graphs are an effective way to convey numeric information, but exisiting methods fail when faced with even minor variations in appearance.
2. DVQA is the task of **Understanding data visualization via Question answering.**
3. Traditional VQA models perform poorly on DVQA, hence specialized models are of need.



Performance Analysis of MOM (Multi Output Model):

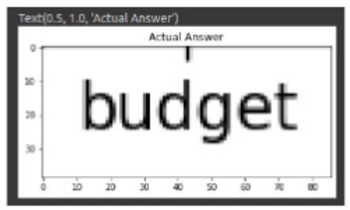


Figure: Actual Bounding-Box

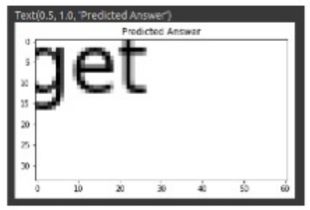


Figure: Predicted Bounding-Box

Performance Analysis of PReFIL:

```
Metrics for val_easy split
        Category: overall:
                Overall accuracy is: 96.9323
        Category: by_qtype:
                Accuracy for reasoning is: 96.5045
                Accuracy for data is: 96.4636
                Accuracy for structure is: 99.7739
        Category: by_imgtype:
                Accuracy for bar is: 96.9323
Metrics for val_hard split
        Category: overall:
                Overall accuracy is: 97.001
        Category: by_qtype:
                Accuracy for reasoning is: 96.5624
                Accuracy for data is: 96.6326
                Accuracy for structure is: 99.6253
        Category: by_imgtype:
                Accuracy for bar is: 97.001
```

Abstract (pix2pix)

1. pix2pix works by training on pairs of images such as building facade labels to building facades and then they attempt to generate a corresponding output image from any given input image.
2. The models are trained on thousands of sample images which are combined in such a way that **the actual image and its ground truth is presented to the model together**, and the model is then trained to predict/convert the images from B (the rough edges) to A (the better representation).