VISUAL QUESTION ANSWERING ON STATISTICAL PLOTS

Team Members: Sneha Jayaraman, Sooryanath I T, Himanshu Jain
Guide: Dr. Mamatha H.R
Department of Computer Science and Engineering, PES University

Problem Statement

To build a **Visual Question Answering system** which accepts **statistical plots** along with **questions** on the plot (with respect to the elements of the plot) to provide answers to the questions posed.

Benefit: Helps data analysts question and understand plots on a large scale, **and automate the decision-making capabilities.**

Scope: **Plots = Dot, Line, Bar (Vertical, Horizontal, Grouped), Questions = Open-ended, Boolean (Support/ Refute)**

Background

Statistical plots consist of **Visual Elements** (Bars, lines, dots) and **Textual Elements** (x-label, y-label, title, x-tick, y-tick). We need to extract information out of visual elements and textual elements to create a semi-structured table that can then be used to answer to questions.

Dataset and Features

The dataset used is **PlotQA**. It consists of images with corresponding annotations (bounding boxes of elements), and question-answer pairs.

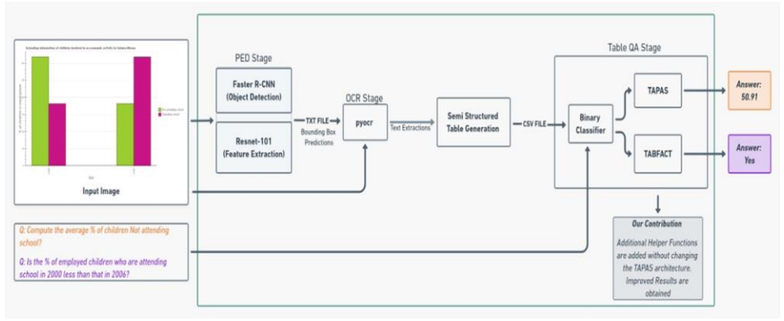
Requirements: Image of statistical plot and a question

Product Features: The product can answer **structural questions, data-retrieval, reasoning type and questions requiring a Yes/No answer** on any dot, line or bar plot.

Design Approach/ Methods

Approach: The design consists of 4 stages

1) Plot Element Detection Stage
2) Optical Character Recognition Stage
3) Semi-Structured Table Generation Stage
4) Table Question Answering Stage – Uses **TAPAS** model to answer open-ended questions, **TABFACT** to answer Yes/No questions



Results

Evaluation: Exact match for textual answers, 5% buffer for numeric answers.

Testing was done on a total of 8000 images and corresponding 1.86 Lakh question-answer pairs.

Plot Type	Number of Images Tested	Total Number of Questions	Number of Correct Answers	Average Accuracy (in %)
Dot	2000	53970	25104	**46.965499**
Vertical	2000	47940	19898	**41.474200**
Horizontal	2000	49241	20128	**40.990114**
Line	2000	35353	14077	**36.669402**

Summary of Project Outcome

Our model is able to answer questions posed on different types of plots significantly better using the TAPAS model with an **overall accuracy of 41.52%.**

It can answer questions requiring textual as well as numeric answers on different types of plots (hand-drawn and computer generated plots) including grouped plots.The product comes with an interactive user interface.

Conclusions and Future Work

This work is a step towards **machine reasoning capabilities**.

Future Work:

1) Extension to other types of plots (ex: pie etc)
2) Expansion to many different types of questions
3) Improving accuracy to match human accuracy.

References

Methani, Nitesh, et al. "Plotqa: Reasoning over scientific plots." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2020.

Kim, Dae Hyun, Enamul Hoque, and Maneesh Agrawala. "Answering questions about charts and generating visual explanations." Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020.


Sneha Jayaraman


Sooryanath I T


Himanshu Jain


Dr. Mamatha H.R