

Lab 1 Theory

Rasmus Tengstedt

May 2024

1 Task 1.1



English



Spanish



Swedish



Arabic

Chatgpt 4o seems to only generate white people if it's not specified what color the people are, this was also test with other stuff like vikings and a banana farmer. Also it seems that less data or its worse and translating Arabic because the people look more unrealistic when the prompt was asked. Also its much more likely to generate males in most languages except for Spanish which only much more often generated women.

This happens probably because in a lot of online places and old media white is the default and the same goes for male. So when it learned it has learned a lot more male in the fields so it generates them more often. This went for a lot of testing but not child caregivers, parents and nurses which both were more likely to be women. This is probably cause media portrays women more often in these jobs and the image generation reflects that.

2 Task 1.2

Accuracy is the ratio of correctly predicted instances (both true positives and true negatives) to the total instances. This is very a useful metric in many cases if it is a balanced set of stuff happening say 50/50 chance and it is not that high risk with it being wrong. But sometimes it will not give a complete picture especially when the dataset is imbalanced. Here it could say 95% accuracy when it is just guessing the same thing over and over if that is 95% of the case. An example of this would be cancer detection. Here recall would be better.

2.1 First Matrix:

- True Negatives (TN) = 990

- False Positives (FP) = 10
- False Negatives (FN) = 20
- True Positives (TP) = 30

Precision

$$\text{Precision} = 30 / 40 = 0.75$$

Recall

$$\text{Recall} = 30 / 50 = 0.6$$

Accuracy

$$\text{Accuracy} = 1020 / 1050 \approx 0.97$$

F1 Score

$$\text{F1 Score} = 2 \times 0.3333 \approx 0.67$$

2.2 Second Matrix:

- True Negatives (TN) = 9000
- False Positives (FP) = 50
- False Negatives (FN) = 100
- True Positives (TP) = 850

Precision

$$\text{Precision} = 850 / 900 \approx 0.94$$

Recall

$$\text{Recall} = 850 / 950 \approx 0.89$$

Accuracy

$$\text{Accuracy} = 9850 / 10000 = 0.98$$

F1 Score

$$\text{F1 Score} = 2 \times 0.8445 \approx 0.92$$