Jerome Ignacio 301406777 Assignment 4

Estimating Human labour Hours

The dataset that I will be utilizing for my project is a medical dataset containing x-rays of the chest. This dataset consists of roughly 6200~ X-ray images If we consider these factors, assuming that we are working from the base level with no xrays at all and they cannot be sourced from a hospital, we can assume that for 6200 ~ xrays, collecting an x-ray could take roughly 9 minutes per patient. Although unrealistic, if we imagine all these 6200 patients one by one, 9 minutes is a good approximation for getting the correct position of the patient, taking the x-ray, and having basic verification. After data collection, we would then have radiologists and trained professionals annotate these X-rays to ensure the accuracy of the labels. Because of the required review, we can assume that each X-ray would need to be verified for 5-10 minutes to identify the disease based on symptoms. After labeling, we would then ensure that these images are anonymous and apply preprocessing techniques to these images such as denoising, these simple preprocessing should take a few seconds per if we assume 10~ seconds are required for every image.

$$X\ Ray\ collection:\ 6200\ \times\ 9\ min\ =\ 55800\ min\ \Rightarrow\ 930\ hrs$$
$$Professional\ annotation:\ 6200\ \times\ 10\ min\ =\ 62000\ min\ \Rightarrow\ 1033\ hrs$$
$$Processing\ image\ time:\ 6200\ \times\ 10\ secs\ =\ 62000\ secs\ \Rightarrow\ 1033\ min\ =\ 17\ hrs$$

$$Total\ Labour\ Hour\ =\ 930\ +\ 1033\ +\ 17\ =\ 1980\ labour\ hours$$

Cost of Dataset commission:

Considering the above approximations for labour hours, we can find an estimation of what it would cost to create a new dataset from scratch, if we consider that we need an X-ray technician, a radiologist to annotate each X-ray, and a data engineer to process these images we can come up with the following costs the following average salary costs are assumptions found using those job titles in Indeed[2] in the BC area.

$$XRay\ Technician:\ \frac{\$46}{hr}Avg\ Salary\ \times\ 930\ hr\ =\ \$42780$$
$$Radiologist:\ \frac{\$196}{hr}Avg\ Salary\ \times\ 1033\ hr\ =\ \$202468$$
$$Data\ Engineer:\ \frac{\$63}{hr}Avg\ Salary\ \times\ 17\ hr\ =\ \$1071$$

$$Total\ Cost\ =\ \$\,246\,319$$

Potential Value

The potential value of this could mainly come down to the amount of time a radiologist can save verifying a chest/lung disease. With the assumption in CAN that roughly 20 million X-rays are performed every year[1], if we assume 5% of those are for X-rays about lung disease, and with the assumed values for time required to verify the disease (5 minutes -10 minutes) and a typical

radiologist salary (196$/hr), we can come up potential values that we could save. If we assume that we can reduce the verification time to a flat rate of 2.5 minutes based on the accuracy of the model. That would mean that we would save approximately 2.5 - 7.5 minutes of a radiologist's time per X-ray. Considering roughly 20 million chest X-rays are performed every year [1], we could assume that based on the time we save for a radiologist, we can generate a potential value of the following:

$$\frac{\$\ Saved}{hour} : Cost \times \frac{2.5\ minutes}{60\ minutes} = \$196 \times (0.0416) = \$8.16$$

$$\$8.16 \times (20\ million)(5\%) = \$8,160,000$$

$$\frac{\$\ Saved}{hour} : Cost \times \frac{7.5\ minutes}{60\ minutes} = \$196 \times (0.125) = \$24.5$$

$$\$24.5 \times (20\ million)(5\%) = \$24,500,000$$

Therefore we can potentially save **$8,160,000** to **$24,500,000** in value by the time that we save for a radiologist verifying our data.

[1]https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt_formats/hecs-sesc/pdf/pubs/radiation/quality-assurance_art-qualite/assurance-eng.pdf
[2]https://ca.indeed.com/