Assignment 2

For this second assessment of this element I will be focusing the design of my research on this research question ‘is there a correlation between social media users who self-harm based on their activity on an app and can AI detect it.

I plan to use computer vision and Natural language processing for this experiment

o detect weather a teen is self -harming. In the last few years there has been a number of cases where teenagers have sadly taken their life through self-harm, investigation have followed following these events and it is regularly found that the victims were exposed to a large amount of content online which may have fuelled these events.

My research question would see if there a correlation between social media users who self-harm based on their activity on an app and can AI detect it.

For this task I will employ several different individuals all in their teens. They will be in three different groups which will be:

1. Participants who have had a history of self-harm but recovered.
2. Participants who have never experienced this.
3. Participants who are currently experiencing self-harm

I will test how much content which can deduced as self-harming content each participant interacted with. Did they comment, like save the post?

I will also assess their usage over a 30 day period and asses what they have posted, is their any text that relates to self-harming, have any pictures been posted to show self-harm.

I will use text classification models to measure the interaction of the participants with other pages via comment sections as well as their captions on their posts and stories. I will use computer vision models to asses’ images they have posted on social media. I will how good it is at detecting physical bruising through self-harm

The tool will use Natural Language processing to assess publicly available content like comments, likes and use data from images that have been posted which show signs of self-harm. The tool will use this information to mark this account and then alter its algorithm to start promoting content regarding seeking solace and getting medical assistance.

The obvious ethical issues would be dealing with those who have a serious medical issue, without any medical qualification would it be wise to undertake a project with individuals with this issue. Another issue would also be data. The aim of the project is to collate social media activity, likes, interactions which may seem intrusive. Another issue is data handling, would lhe data be safe. Its personal data about a serious medical issue. If it happened to fall into the wrong hands that could be ever so detrimental for the participants.

Potential errors may occur if there is a start k contrast as to when is online. As we have not established when harmful content is circualated. If one participant is online at night would this mean they are more exposed to harful content?

Even though the participants will all fall into three groups there will be some confounding variables to consider. How often do the individual post on social media, whether occupation are they studying, does being less busy increases the chances of being exposed to self-harming content?