Capstone Check-in #1

Proposal 1: YouTube Comment Summarizer

Proposal 1: YouTube Comment Summarizer and Sentiment Analyzer

1. Generate advanced analytics on a given YouTube video based on comments, including:
    a. Sentiment analysis
    b. Text summarization of top comments (What are the people saying on your video?)
2. YouTubers
    a. Comments can help inform decisions on what future content to make
    b. Manually wading through the comment section can be a mentally exhausting task.
        i. Let an app do it automatically!
3. Data can be obtained using YouTube API or by manually scraping
4. Some possible limitations on manual scraping or number of results returnable by API
    a. Manual scraping can take a while to complete
    b. API might be difficult to use

Proposal 2: Facial Expression Recognition

Proposal 2: Facial Expression Recognition

1. What is your problem statement? What will you actually be doing?
    a. Build a Flask app that identifies emotion based on an uploaded picture
2. Who is your audience? Why will they care?
    a. Product testers/designers
    b. During focus testing, this tool can be used to detect non-verbal response to the product
3. What is your success metric? How will you know if you are actually solving the problem in a useful way?
    a. Success will depend on the accuracy of my classification model
4. What is your data source? What format is your data in? How much cleaning and munging will be required?
    a. Model will be trained on Google Facial Expressions Dataset
        i. https://research.google/tools/datasets/google-facial-expression/
5. What are potential challenges or obstacles and how will you mitigate them?

Proposal 3: Handwritten Text Recognition System

Proposal #3: Handwritten Text Recognition System

1. What is your problem statement? What will you actually be doing?
    a. Build a Flask app that identifies identifies text from handwriting
    b. Based on the following TDS article:
        i. https://towardsdatascience.com/build-a-handwritten-text-recognition-system-using-tensorflow-2326a3487cd5
2. Who is your audience? Why will they care?
    a. Anyone who needs to digitize handwritten forms
3. What is your success metric? How will you know if you are actually solving the problem in a useful way?
    a. Success will depend on the accuracy of my classification model
4. What is your data source? What format is your data in? How much cleaning and munging will be required?
    a. Model to be trained on IAM handwriting database
        i. https://fki.tic.heia-fr.ch/databases/iam-on-line-handwriting-database
5. What are potential challenges or obstacles and how will you mitigate them?