- Tiernon Riesenmy, Duc Trinh, Eli Lleshi: Analyzing and Predicting the Popularity of Memes -feedback

Please share this feedback with the other team members as well!

Your proposed project is really interesting and challenging. Predicting the popularity of a meme is a hard task, but it is a great chance to learn about the fields of deep learning and computer vision. I have searched for some links and articles and added a few thoughts to help you start this project:

I recommend you to handle it as a binary classification task: dank or not dank
You can define poplar, if upvote > downvote.
Or another idea is to handle controversial memes, and just define:
popular = upvote > 2*downvote
unpopular = downvote > 2*upvote
controversial = in between
Or you can also drop the controversial group just care about the purely popular or unpopular memes and handle it as a binary classification task.

Perhaps you can make advantage of this webpage using some web scraping techniques: https://knowyourmeme.com/memes
Anyway, did you know that this guy is Hungarian?
https://knowyourmeme.com/memes/hide-the-pain-harold

What makes an image popular?
(https://people.csail.mit.edu/khosla/papers/www2014_khosla.pdf)

- This article has some nice ideas to start with. As far as i understood the reddit dataset (https://www.kaggle.com/kuldeepchoudhary/reddit-meme-analysis) also has upvotes for each image. It can be a really good starting point to extract some features from the images and see if they correlate with the number of upvotes (our popularity measure) or not.
- The article suggests first to extract low-level computer vision features from images. For extracting e.g. color features I recommend you to use opencv:
  - https://realpython.com/python-opencv-color-spaces/;
  - https://www.analyticsvidhya.com/blog/2019/09/feature-engineering-images-introduction-hog-feature-descriptor/?utm_source=blog&utm_medium=3-techniques-extract-features-from-image-data-machine-learning

To get familiar with convolutional neural networks I recommend you to begin it with using pre-trained models we will also cover neural networks in class:

- https://www.learnopencv.com/keras-tutorial-using-pre-trained-imagenet-models/;
- https://medium.com/@ODSC/how-to-leverage-pre-trained-layers-in-image-classification-31fb9b8cdd0;
- https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751

If you don't stick to computer vision, here's an article about meme text generation: https://towardsdatascience.com/meme-text-generation-with-a-deep-convolutional-network-in-keras-tensorflow-a57c6f218e85