IDEA 1:
Project title 1 (Descriptive title that captures your idea1)

Toxic Comment Classifier/Identifier

What you want to do?

To build and deploy a Classifier /Identifier model that can detect harmful comment or tweet. When a text or a few sentences are provided as input, the model must be able to classify the text into the various categories. The text is classified as: hate-speech, offensive language, and neither. We can further explore on more categories of offensive speech racist, sexist, homophobic, or generally offensive. The initial exploratory analysis will involve a study on the words or bag of words that show hateful speech. The models can be supervised, by training using labeled data. We can also try unsupervised learning to identify the fine details.
Some Kaggle datasets available are:
https://www.kaggle.com/mrmorj/hate-speech-and-offensive-language-dataset
https://www.kaggle.com/arkhoshghalb/detecting-hate-tweets
https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/discussion

Why should we care?
Although online communities do not promote hateful reactions from users, it is not preventing these actions. The users do not have any limitation in the words used in comments. But, the threat of abuse and harassment online means that many people stop expressing themselves and give up on seeking different opinions. There have been extreme scenarios which have even led to death of people or artists expressing themselves. My motivation comes from a case when certain Kpop idols (whom I used to follow) committed suicide due to persistent and harsh comments received by anti-fans. I feel that the online communities must have a classifier. These classifiers can be used to detect any harsh comments and hide/delete them immediate. The classifier could also trigger warning notifications to admins of a community or group and even warn the user that such behavior will not be tolerated in the future.

Keywords – To mark topic and domain of the idea.
Hate comments, offensive text, sentiment analysis, abuse, harsh words, NLP, Classifer

IDEA 2:
Project title 2 (Descriptive title that captures your idea2)

Factors for content popularity (Trending content)

What you want to do?

To understand the contributing factors for a content to be successful on an online community, we will study popular content based on their attributes such as number of views, shares, comments and likes. This is mostly a data exploratory project, which may not involve building predictive models.
Related datasets:
https://www.kaggle.com/datasnaek/youtube-new

Why should we care?
Digital Marketing and content advertisement has become a new method to raise awareness of a product or content. Understanding the success factors for trending items can give us insight on product. The trending content have certain parameters which can be identified and used for marketing content through influencers.

Keywords – To mark topic and domain of the idea.
Trending content, popular topics, marketing, influencers, content advisors