**Problem Statement – RNN**

**Introduction :-**

The Social Dilemma, a documentary-drama hybrid explores the dangerous human impact of social networking, with tech experts sounding the alarm on their own creations as the tech experts sound the alarm on the dangerous human impact of social networking. This dataset brings you thetwitter responses made with the #TheSocialDilemma hashtag after watching the eye-opening documentary "The Social Dilemma" released in an OTT platform(Netflix) on September 9th, 2020. You can categorize the tweets into different groups to identify the sentiment of the users regarding the documentary.

**Goal:** You are hired as a deep learning engineer and you are asked to categorize the tweets into three categories.

**Constraints:** You should be using only RNN to generate results and should not be using any LSTMor ML classification models to generate results.

**Data Description:** The dataset was extracted using TwitterAPI, consisting of nearly 20000 tweets from twitter users all over the globe!

**1. Attribute Information:**

**2. user\_name -** The name of the user, as they’ve defined it.

**3. user\_location -** The user-defined location for this account’s profile.

**4. user\_description -** The user-defined UTF-8 string describing their account.

**5. user\_created -** Time and date, when the account was created.

**6. user\_followers -** The number of followers an account currently has.

**7. user\_friends –** The number of friends an account currently has.

**8. user\_favourites -** The number of favorites an account currently has.

**9. user\_verified -** When true, indicates that the user has a verified account.

**10.date -** UTC time and date when the Tweet was created.

**11. hashtags -** All the other hashtags posted in the tweet along with #TheSocialDilemma

**12. source -** Utility used to post the Tweet, Tweets from the Twitter website have a sourcevalue – web

**13. is\_retweet -** Indicates whether this Tweet has been Retweeted by the authenticating user.

**14. clean\_text –** Cleaned text of the tweet.

**15. Sentiment (target) -** Indicates the sentiment of the tweet, consists of three categories: Positive, neutral, and negative.

**Provided Files:**

**Train\_data -** Should be used to feed your model.

**Test\_data -** Should be used only to generate prediction

**Evaluation Criteria:**

The evaluation metric for this problem statement is the validation Accuracy Score.