Deep Learning + AI Exam

[Time: 5 hrs.]

[Total Marks: 50]

Part I: CNN   [Total Marks - 25]

E-commerce has rapidly grown and their business strategies are completely based on user actions and user experiences. Although it is completely based on users, we should also not forget to say that there is a technology bridge in between users and growth in business. It may be Machine Learning or Deep Learning. Companies apply many image classification techniques on data to improve their catalog and give best suggestions to the users. They need accurate product classification on their platforms for better user experience. But when you talk about products, there exists a huge variety and classifying within varieties is really challenging. As a Deep Learning engineer, you should always try cracking these kinds of challenges by classifying things within a product itself.

Goal: Given the images of a product with multiple categories, train a model which can classify the type of a product.

Data Description: Data is all about images of shoes with multiple categories and data is collected from a popular Ecommerce site. Data set consists of two folders train and test.

Provided Files:
Train: train set consists of images belonging to 3 different categories of shoes in 3 differentfolders: Boots, Sandals and Slippers.
Test: test set consists of images belonging to all 3 categories of shoes into a single folder.

Instructions:
1. Train set should be used to feed the model.
2. Test set should be used to predict labels for test data.

Evaluation Criteria: The evaluation metric for this problem statement is the Accuracy scorewhere each shoe category is matched with the actual shoe label.

Part II: RNN   [Total Marks - 25]

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 predictions.

Evaluation Criteria:

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