Assignment-7

1. Use following sentences to find the tf*idf vectorized data.

   1. The postman delivered the package to the wrong address.

   2. I wrapped a beautiful present for my friend's birthday.

   3. The delivery truck arrived late due to heavy traffic.

   4. We need to check the shipping address before sending the order.

   5. Online shopping offers a wide variety of products with fast delivery.

2. Download the file spam_messages.csv from the link below:

   https://mitu.co.in/dataset

   This file contains the collection of spam and not spam (ham) messages received in the mail box.

   Perform the following operations on it:

   1. Read the file

   2. Seperate input and output data

   3. Explore the output data

   4. Prepare and vectorize the data

   5. Cross validate the data

   6. Build the Decision Tree model for classification

   7. Evaluate the performance of algorithm

   8. Download the file sample.csv which contains the unknown mails. Predict the category of these mails.