**INFO 7390 – Advances in Data Sciences**

**Research Project Proposal**

**Topic**: Image Caption Generator

**Project Description**

* Objectives  
  Our objective is to build a working model that will generate image caption after reading the image and also to learn concepts of deep learning with CNN & LSTM
* Requirements  
  Good knowledge of Deep Learning, Python, Jupyter notebook, Keras library, Numpy & NLP. Also, below libraries should be installed before start of the project.  
  - Tensorflow  
  - Keras  
  - Pillow  
  - Numpy  
  - tqdm  
  - Jupyterlab

**Algorithms Sources**

* Convolutional Neural Network (CNN) will be used to extract features from images
* Long Short-Term Memory (LSTM) will be used to generate image description

**Data Sources**

<https://www.kaggle.com/ming666/flicker8k-dataset>

**References**

[1] Oriol Vinyals, Alexander Toshev, Samy Bengio, Dumitru Erhan. [Show and Tell: A Neural Image Caption Generator](https://arxiv.org/pdf/1411.4555.pdf)

[2] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard Zemel, Yoshua Bengio. [Show, Attend and Tell: Neural Image Caption Generation with Visual Attention](https://arxiv.org/pdf/1502.03044.pdf)

**Members**

Harshad Jaju – 001439898  
Aman Kapoor – 001448282