**CHALLENGES FACED AND SOLUTIONS**

First of all thank you for giving me this assignment. It was interesting, challenging and a lot of fun.Below are the challenges I faced :

* Dealing with JSON files. I ended up learning about the JSON module from Google.
* The problem with captions and tags : pictures that didn’t contain pav-bhaji also had the word #pavbhaji mentioned. It was difficult for me initially to notice the difference between “ Just had pav-bhaji #food, #blessed” and “Love Mumbai #rain, #pavbhaji, #mumbaikar”. This was solved by distinguishing between those that words that had “#”before them and those that didn’t.

**OTHER CHALLENGES**

* The JSON data does not provide any significant insight about the objective. It was very difficult to build any model ( let alone a decent one) with this data.
* The accuracy of the model is around 65/66%. (bad I know!) This is mainly due to the above point. Some things that can be done to improve the model are further hyperparameter tuning and using GridSearchCV etc,