

# MCA Project Document

Group 4

Shahid Nawaz Khan (2017102)

Sudhanshu Mohan (2017110)

Aarushii Agrawal(2017324)

Link to the dataset: <https://drive.google.com/open?id=1i8vTKl6vVmMSTZccq2ptEQKkhJ96YdGf>

The link is to a folder called Data. It has 2 sub folders Images and Tweets. Tweets has 2 types of files, namely "tweet" and "data". The "tweet" files contain the PDIPs generated for the corresponding hashtag.

Link to the GitHub repository: <https://github.com/sudoplox/MCA-Project>

MCA\_Twitter.ipynb is the notebook used for scraping data using Twitter APIs.

DATA\_TRAINING.ipynb is the notebook used for preprocessing and training.