Big Data Analytics and Application

Project Submission 1 (Due 9/28/2020 11:59 pm CST)

Team number: **Group 5**

Team Members: **Jonathan Wolfe, Mehmet Acikgoz, Irfan Ahmed**

Kaggle challenge link: <https://www.kaggle.com/rmisra/news-category-dataset>

Kaggle Data link : <https://www.kaggle.com/rmisra/news-category-dataset>

“Assignment 1 Life” should accurately and passionately describe real-world problems facing the targeted users or audience:

Answer the following question:

1. Define your scope or domain where the use case is relevant or prevalent?
   1. Students, journalist, data mining.
2. What is your main story?
   1. I am interested in news articles based on the topics that are popular.
3. Who are the characters or people in the main story?
   1. I am a college student who is interested in news articles based on the topics of interest. My interest now is focused on police shootings.
4. What problem happened to them?
   1. I am looking for articles that relate to politics and police shootings.
5. Where did the problem take place?
   1. Where means two things: 1) The environment and settings that the people or the community is living in, and 2) the place/location where the problem take place.
      1. I am only concerned about articles that relate to the country I currently reside in.
6. Why?
   1. Why means the possible causes and/or origin of the problem?
      1. Mitigate bias in different reporting’s.
7. How?
   1. If you would like, you can add a dimension of how. How did it happen? Sometimes, the answer to how can be covered by what, when and where.
      1. It becomes tedious having to read every article just to compare few differences.
8. What new aspect your team is adding to the Kaggle challenge (please be very specific and clearly explain what you are trying to add to the Kaggle challenge)