1. Name of Team and Members
2. The Issue/ Problem We’re Working On
   * Since the 2016 presidential election, one topic dominating political discourse is the issue of “Fake News”. A number of political pundits claim that the rise of  significantly biased and/or untrue news influenced the election, though a study by researchers from Stanford and New York University concluded otherwise. Nonetheless, fake news posts have exploited Facebook users’ feeds to propagate throughout the internet.
   * The data science community has looked to respond to this problem through a Kaggle competition called the “Fake News Challenge”
3. What Makes It Interesting or Hard.  Your approach to the problem and what AI techniques/algorithms you will explore, develop, implement to address the problem
   * Dataset 1 (fake) - <https://www.kaggle.com/mrisdal/fake-news>
     1. Includes: author, title, language, site url, country, domain rank, spam score, type
   * Dataset 2 (real) – allsides.com
     1. Articles are categorized by topic: environment, economics, etc. and by political lean (left, center, right)
     2. Web crawling / scraping to get thousands of articles
   * Implement Naïve Bayes to predictive fake news vs real news.
     1. Import packages such as pandas, numpy, csv.
     2. Read the csv file and display the data set (use python notebooks)
     3. Remove erroneous null and NAN values
     4. Calculate the length of news to check the length of fake and real news
     5. Data Preprocessing – NLTK tokenization
        1. Tokenize and normalize text
        2. Lemmatization and vectorization (count vectorizer vs tfidf vectorizer)
     6. Feature extraction
     7. Training a model to detect fake
     8. Build & evaluate
   * Try other supervised machine learning algorithms such as SVM, Logistic regression and KNN
4. How it can be evaluated
5. How we are dividing responsibilities