**BACKGROUND**:

The goal of the exercise is to design a state-of-the-art model to solve the problem of entity-based sentiment analysis. Entity-based sentiment analysis is widely used to automatically detect the opinion expressed towards named entities. It is particularly social and news media analytics

**PROBLEM DEFINITION**:

Given a sentence and a named entity contained within a sentence, predict the sentiment towards the named entity

**EXAMPLE**:

Sentence: “Donald Trump and Barack Obama are US presidents but the later is better“.  
Requirement: Find the targeted sentiment for Donald Trump and Barack Obama.

**DATA**:

You are provided with sample training data to build models. The training data is a tab-separated file with the sentence, named-entity and the sentiment label for the named-entity. The train data can be used to also create a development dataset for tuning the model hyperparameters. You can also use other open-source data sets available for training. However, you must describe how external data sets are used in your test submission. Also, there is a test dataset provided with sentences and named entities. You are expected to provide the sentiment predictions for the named entities using your model.

**EXPECTED DELIVERABLES**:

1. Build a machine learning model that can predict the sentiment of an entity
2. Describe the technicalities of your approach and evaluation in a 1-2 page report.
3. Performance statistics of the model such as precision, recall, F-score and accuracy (Use stratified k-fold cross-validation on the training data provided)
4. A working system with access to the code (github or bitbucket) and how to run instructions

**QUESTIONS TO THINK**:

1. How does your model handle contrastive conjunction and negation?
2. How do you deal with sentences that have multiple named entities and opposing sentiment (e.g. The car has great cruise control but dash controls are terrible)
3. How do you handle unseen vocabulary
4. Identify the key limitations of your approach