**General theme:**

Sentiment analysis AI (and other Game Theory AIs, if time is ample)

**Participant requirements:**

Proficient in Python (or other coding language)

**Platforms:**

* Google Colab
* (Potentially) Wolfram Machine Learning Platform (more research on this): <https://reference.wolfram.com/language/guide/MachineLearning.html>

**Methodology:**

* Introduce some basic knowledge and algorithms of machine learning
* Familiarize the participants with Google Colab
* Demonstrate a sample sentiment analysis that focuses on 1 sentiment only. For example, this program only returns whether or not this statement is happy, so it only provides a true or false answer.
* The participants are asked to familiarize themselves with the sample code, and create a project that expands on it.
* The hackathon will be divided into two parts:
  + The introductory project: the participants use the sample as their basis and expand on it. This would be more participant friendly:
  + The long-term project: the participants are asked to use neutral network to program a more complex sentiment analysis AI
* **Introductory Project (no neural network)**: the participants can choose the following projects.
  + (Introductory) AI that returns a range (from 0 to 1) for the degree of one sentiment from a sentence
  + (Basic) AI that determines the prominent sentiment of a sentence from a set of possible sentiments (e.g. happy from the set of: sad, happy, angry).
  + (Intermediate) AI that shows the degree of association for each sentiment category (e.g. 30% happy, 20% sad, 50% angry)
  + (Advance) On top of the intermediate requirement, this AI judges the level of authenticity of the emotions in the sentence!
* **Long term project documentation using neural network**:
  + Provide resources on the neural network.
  + <https://www.tensorflow.org/tutorials/text/text_classification_rnn>
* The participants can do these projects progressively, and are encouraged to save their work on the previous project before advancing to the next one.

**Data:**

* Pandas, matplotlib
* Data is provided
* When using the sample code Dev sent, here are the areas of customization
  + Data manipulation:
    - Vectorization
    - Counting
    - Combination of the above
  + SK-learn model
    - Bayes Theorem (Naive & Multinomial)
    - Decision Trees
    - SVM
    - Random Forest Customization
    - Regression (linear & logistic)
  + Each model have their own parameters

**Awards:**

* Best Accuracy
  + By percentile
* Most efficient (fastest algorithm exceeding benchmark accuracy)
* Best explanation / Best commenting

**Judging:**

* The participants save their trained model
* At the end of the competition perion, participants (or groups) can only submit one project for judging.
* The submissions are judged based on accuracy with Kaggle’s database: <https://www.kaggle.com/>. The awards will be given out to the most accurate participants.
* In addition, there will be an award for creativity, as long as the project relates to sentiment analysis or AI! So be creative as there are no limits for AI!