Assignment 3 – Maze Ball Player Modeling

1. The Maze-Ball dataset is a publicly available game experience corpus (released by Hector P. Martinez) that contains two modalities of data obtained from Maze-Ball players: their gameplay attributes and three physiological signals: blood volume pulse, heart rate and skin conductance.
   The dataset also contains demographical data of the players and self-reported annotations of experience in two forms: ratings and ranks.
   a. Download the files ***mazeball.json*** and ***WID3009_Tutorial_11.ipynb*** from Spectrum.
   b. Using Jupyter-Notebook or Google Colab, launch ***WID3009_Tutorial_11.ipynb*** to view the format of the dataset.
   c. Choose one or more of the affective or cognitive states that are available to model (the **output** of the model).
      For example: under "Preferences" you can choose to model "Anxiety", "Challenge", "Excitement", etc.
   d. Extract information that might be useful for player modeling:
      i. Gameplay information;
      ii. Keystroke information;
      iii. Physiological information;
      iv. Demographic information
      All four types of information have different data sizes and resolution; how do you represent them as one feature vector?
   e. Use at least two different **supervised learning algorithms** to build player experience models that predict both the ratings and the rank labels of experience.
      The ***player_experience.ipynb*** notebook gives a simple example for one type of affective information ("Fun") and one type of data (physiological signals).
      Do the same for the other affective information (i.e. "Anxiety", "Challenge", "Excitement") and for other types of player data (gameplay, keystroke).
   f. Save your codes and results in the same Jupyter Notebook and upload to Spectrum as your assignment submission.
      Instead of having to write a report, all your work must be shown in the Jupyter Notebook file. If you use two different supervised learning algorithms, they must be shown in different cells along with their results.
      If you have more than one Notebook file, zip them and upload the zip file.

**Hint**: the Python package ***scikit-learn*** has many algorithms for predictive modeling. See the following for examples: https://scikit-learn.org/stable/auto_examples/index.html

**Hint**: perform feature extraction on the physiological signals for dimensionality reduction. For a list of useful features that can be extracted from physiological signals, you can refer to the research paper "Towards Affective Camera Control in Games" by Yannakakis et al. Try to achieve better performance than the paper by testing different feature extraction and selection methods, or using different supervised learning models.

| Marking scheme | (10%) |
|---|---|
| Compared at least two supervised learning methods | 2% |
| Modeled at least two affect (i.e. "Challenge", "Fun") | 2% |
| How many modalities (physio / gameplay / keystroke) are used | 3% |
| Presented results (i.e. graphs, accuracy) | 1% |
| How initial results are improved (i.e. with feature selection, or ensemble methods, or other methods). | 2% |