MODULE CODE: IB98D0

MSOR/MSBAC Advanced Data Analysis Module 2020/2021

A.D.A. Assessment

This group assignment counts for **25%** of the assessment for the Advanced Data Analysis module.

Down load the file SkillCraft_Dataset.csv from the ADA module website on my.wbs. This is a '.csv' format data file containing video game telemetric data from a cognitive study of a Complex Skill Learning project. A random sample has been chosen for this assignment: 697 Real-time Strategy game players at 7 different levels of expertise, ranging from novices to full-time professionals. The data set information[1] has been provided on page 2, further details of this project and the full data set can be found:

http://archive.ics.uci.edu/ml/datasets/SkillCraft1+Master+Table+Dataset

Assignment question:

Question 1: What affects the expert level? How can the players improve their level?

Question 2: Are there any identifiable sets or "groups" of players within the data? If so, what are their distinguishing features?

Question 3: Investigate the association of the different variables. How many hidden factors are related to these characteristics? Can you interpret these factors?

Use appropriate techniques to answer the above questions and give justification to your answers.

Format of response:

Write a technical report with a brief executive summary. The whole report should not be more than 12 pages (excluding appendices) in 12 point Times New Roman font (single spaced). Marks will be deducted for exceeding the page limit. It is essential that the report contains relevant numbered tables and figures to assist discussion while other output tables and figures can be put in appendices.

The main focus should be on the application of the techniques. Detailed descriptions of the theories are not necessary. A brief outline of their purpose & assumptions is enough. For each analysis discuss what you are doing and why. For each analytical approach explain its purpose, state/discuss any assumptions made, check on the suitability of the data for the analyses. Explain and discuss interpretations and implications of results: What do they mean? How could they be used? Outline any potential problems. Remember to be critical! Feel free to clothe the assignment questions in an appropriate context if this helps your discussion of relevant issues.

You will be working on this project in groups. The groups have been formed and posted on-line by the Programme Office.

The deadline for this assessment is **12noon Monday 22 March**. Clearly label the report with the ID numbers of the students in the group. Submit the report electronically.

[1] In RTS games, players develop game pieces called units with the ultimate goal to destroy their opponent's headquarters. First, the games have an economic component such as player's strategic decisions are related to balancing spending on military and economic strength. Second, the game board, called a map, is much larger than what that player can see at any one time. The resulting uncertainty about the game state leads to a variety of information gathering strategies, and requires vigilance and highly developed attentional processes. Third, in

RTS games players do not have to wait for their opponent to play their turn. Players that can execute strategic goals more efficiently have an enormous advantage. Consequently, motor skills with a keyboard and mouse are an integral component of the game. Each game produces lots of behavioral data.

Data set information:

-- The screen movements have been aggregated into screen-fixations using a Salvucci & Goldberg (2000) dispersion-threshold algorithm, and defined Perception Action Cycles (PACs) as fixations with at least one action.
-- Time is recorded in terms of timestamps in the StarCraft 2 replay file. When the game is played on 'faster', 1 real-time second is equivalent to roughly 88.5 timestamps.
-- List of possible game actions is discussed in Thompson, Blair, Chen, & Henrey (2013)

Attribute Information: below is a brief description of each variable in the dataset.

1. GameID: Unique ID number for each game player (integer)
2. LeagueIndex (Expert level): Bronze, Silver, Gold, Platinum, Diamond, Master, and Professional leagues coded 1-7 (Ordinal)
3. Age: Age of each player (integer)
4. HoursPerWeek: Reported hours spent playing per week (integer)
5. TotalHours: Reported total hours spent playing (integer)
6. APM: Action per minute (continuous)
7. SelectByHotkeys: Number of unit or building selections made using hotkeys per timestamp (continuous)
8. AssignToHotkeys: Number of units or buildings assigned to hotkeys per timestamp (continuous)
9. UniqueHotkeys: Number of unique hotkeys used per timestamp (continuous)
10. MinimapAttacks: Number of attack actions on minimap per timestamp (continuous)
11. MinimapRightClicks: number of right-clicks on minimap per timestamp (continuous)
12. NumberOfPACs: Number of PACs per timestamp (continuous)
13. GapBetweenPACs: Mean duration in milliseconds between PACs (continuous)
14. ActionLatency: Mean latency from the onset of a PACs to their first action in milliseconds (continuous)
15. ActionsInPAC: Mean number of actions within each PAC (continuous)
16. TotalMapExplored: The number of 24x24 game coordinate grids viewed by the player per timestamp (continuous)
17. WorkersMade: Number of SCVs, drones, and probes trained per timestamp (continuous)
18. UniqueUnitsMade: Unique unites made per timestamp (continuous)
19. ComplexUnitsMade: Number of ghosts, infestors, and high templars trained per timestamp (continuous)
20. ComplexAbilitiesUsed: Abilities requiring specific targeting instructions used per timestamp (continuous)

Perception-Action-Cycle (PAC) variables. Each variable pertains to a period of time where players are fixating and acting at a particular location. Many of these variables will therefore reflect both attentional processes (because Perception-Action-Cycles have consequences for what players are able to attend to), perceptual processes (because shifts of screen imply new stimuli) and cognitive-motor speed (in the sense that actions must not only be fast but meaningful and useful).

SCV: Space Construction Vehicle, is a staple in intra-Colonial construction and engineering due to its ability to perform a multitude of tasks, including the construction of new buildings and the transportation of raw resource materials. It is this versatility, and an unmatched reliability, that make the SCV an invaluable tool in rapidly establishing Marine encampments and strike bases on any terrain.

Actions per minute. This variable is often used as a predictor of expertise in the StarCraft community and is automatically calculated by the game. It is a measure of cognitive motor speed.