Assessment Brief: CW1 Data Science 2024-25

Module title	Data Science
Module code	COMP5122M
Assignment title	Coursework 1: Understanding The Impact of Social Isolation and Loneliness in a Game Environment
Assignment type and description	Assignment
Rationale	To learn how to analyse data and communicate your findings.
Word limit and guidance	The report has a limit of six pages and should be written single-spaced in an 11 pt font (or larger).
Weighting	50%
Submission deadline	23:59, Friday 13th Dec 2024
Submission method	Via Minerva
Feedback provision	Marks and feedback will be returned via Minerva.
Learning outcomes assessed	Understand the work of a data scientist Acquire, link, and investigate the quality of data Do exploratory data analysis Visualize data Model data using Machine Learning approaches Evaluate Machine Learning models Apply problem-solving skills to effectively analyse data and communicate findings for a given application scenario
Module leader	Duygu Sarikaya

1. Assignment guidance

The objective is to assess your ability to analyse data and communicate your findings. You need to do this coursework in the groups that are defined on Minerva. A group comprises a maximum of four people.

Animal Crossing: New Horizons (ACNH) is a typical life-simulation game that hit the market in March 2020 when almost everyone around the globe was forced to stay at home. The game was swiftly sold to more than 13.4 million copies in the first six weeks. When playing ACNH, the players immerse into an idyllic and deserted island with the responsibility of building their island by developing the ecosystem and community. Their daily activities are related to the environment, such as growing flowers, planting fruit, catching fish, snaring bugs, or submitting the fish and bugs to the museum.

You can watch a teaser of the game here: What Is Animal Crossing: New Horizons? A Guide for the Uninitiated

A different and free version of the game is also available as a mobile app: Animal Crossing: Pocket Camp

The Impact of Social Isolation and Loneliness on Gaming

During the COVID-19 pandemic, anxiety and depression symptoms rose significantly. In response, many sought new ways to connect and mitigate the effects of self-isolation. Online gaming provides a virtual space that may meet social needs and offer a sense of competence often lacking in isolation (Cole et al. 2007). A study during the first wave of the pandemic found a positive correlation between social gaming frequency and perceived loneliness, suggesting that social gaming plays a significant role in alleviating loneliness during social distancing (Nebel et al. 2022). The World Health Organization endorsed the gaming industry's #PlayApartTogether

campaign, promoting healthy online gaming practices that foster social interaction. This led to a marked increase in online gaming and activities like e-sports viewing and live streaming. During the Covid-19 pandemic, while many industries are shaken, the video game industry rose rapidly in terms of game players and revenue. As people worldwide were forced to stay at home for public health safety, video games exploded in popularity, with an estimate 2.7 billion game players worldwide in 2020. In the first nine months of 2020, the number of video games sold in the United States surged 23% from the same period of last year to 29.4 billion. The relationship between playing video games with gamers' behavior and perception has been a pivotal topic of many researchers worldwide (Yang et al. 2023).

Exploring the correlation between the in-game behaviors of game players in isolation

Activities such as "enjoying the scenery" or "Participate in Mystery Island Tours" could be attractive to players who are being self-isolated/social distancing as it virtually gives them a chance to travel. Similarly, activities such as "Send it to a friend as a gift (for flowers, bugs etc.)" could be attractive to players who are being self-isolated/social distancing as it gives them a chance to connect with other players. We hypothesize that individuals, who are being self-isolated/social distancing for a longer time, are likely to participate in certain in-game behaviours that give them a chance to virtually travel and connect with other players more than the players who have not been self-isolated/social distancing for long.

Limitations: We are aware that our hypothesis on in-game activities and isolation and social distancing is limited as humans can separate games from reality, regardless of how well games resemble reality.

A multinational dataset of game players' behaviors in a virtual world and environmental perceptions

The dataset (Vuong et al. 2021) offers valuable resources regarding environmental perception and behaviors in the virtual world of 640 Animal Crossing: New Horizons (ACNH) gameplayers from 29 countries around the globe. The dataset consists of six major categories: 1) socio-demographic profile, 2) COVID-19 concern, 3) environmental perception, 4) game-playing habit, 5) in-game behavior, and 6) game-playing feeling. The 12th item in New Ecological Paradigm Scale by Dunlap et al. (Denlap et al. 2000) is used to measure human-centered attitudes of the game players. The players' attitudes toward the idea of human centeredness were measured ordinally on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree).

Dataset

The dataset is publicly available for download at Science Data k

You can find more information about the dataset and its collection here (See Data Availability Statement): A Multinational Data Set of Game Players' Behaviors in a Virtual World and Environmental Perceptions

2. Assessment tasks

Your tasks are:

Detailed analysis:

a) Do exploratory data analysis to show:
   - The distribution of the players' length of being self-isolated/social distancing
   - The distribution of the players' length of being self-isolated/social distancing according to regions
   - The relationship between the players' length of being self-isolated/social distancing and the game-playing frequency of the players.

- A comparison of the frequency of the different lengths of being self-isolated/social distancing and game-playing feeling response "I lost connection with the outside world".

b) Identify the most important in-game behaviours that indicate the player's length of being self-isolated/ social distancing.

c) Develop a machine learning model that can be used to predict a player's length of being self-isolated/ social distancing based on in-game behaviour variables only. To do this, you will need to split the dataset into training, test, and evaluation sets. Evaluate your model's performance using appropriate metrics. Please describe the method you used for your model, and your motivation for choosing the specific methods.

3. General guidance and study support

See skills@library

4. Assessment criteria and marking process

The report has a limit of six pages and should be written single-spaced in an 11 pt font (or larger). The report should have the following section headings and content:

1. **Group members**: Name & student ID
2. **Introduction** [3 marks]. State the aim of each aspect of the coursework (specific detailed analysis objectives) and describe the data file that was used. One paragraph is sufficient.
3. **Detailed analysis** [44 marks]:
   a) [16 marks] For the items mentioned in the data exploratory analysis, please use appropriate visualizations and report findings.
   b) [8 marks] Please describe the method you used for identifying the most important in-game behaviours that indicate the player's length of being self-isolated/social distancing, and your motivation to choose the method.
   c) [20 marks] Please describe the method you used to develop your machine learning model, and your motivation to choose the specific methods. Give details on how you split your dataset, preprocess data, how you define your task, how you evaluate the performance of your model, the metrics you used to evaluate your model, and the motivation for choosing these metrics. Please report the performance of your model.
4. **Conclusions** [3 marks]: Restate your main findings. One paragraph is sufficient.

To gain full marks your report needs to adhere to the specified structure, be well presented and easy to understand. Moreover, you should show that you are following the data science road map, that you have put in thought and effort, and that you have done research to pick proper visualizations, best machine learning approaches, preprocessing methods, proper metrics etc. for the problem you are working on.

Do not include references in your report. They can go in the appendix (appendix limit: 1 page).

5. Presentation and referencing

The quality of written English will be assessed in this work. As a minimum, you must ensure:

- Paragraphs are used

- There are links between and within paragraphs although these may be ineffective at times
- Word choice and grammar do not seriously undermine the meaning and comprehensibility of the argument
- Word choice and grammar are generally appropriate to an academic text

6. Submission requirements

One person should make the submission on behalf of the group.

Submit a PDF file that contains your report.

Optionally: Before the first page of the report, insert a signed copy of the COMP5122M Data Science *Group Work Form* for your group. You only need to do that if the group members did not contribute equally to the coursework.

7. Academic misconduct and plagiarism

Leeds students are part of an academic community that shares ideas and develops new ones.

You need to learn how to work with others, how to interpret and present other people's ideas, and how to produce your own independent academic work. It is essential that you can distinguish between other people's work and your own, and correctly acknowledge other people's work.

All students new to the University are expected to complete an online Academic Integrity tutorial and test, and all Leeds students should ensure that they are aware of the principles of Academic integrity.

When you submit work for assessment it is expected that it will meet the University's academic integrity standards.

If you do not understand what these standards are, or how they apply to your work, then please ask the module teaching staff for further guidance.

By submitting this assignment, you are confirming that the work is a true expression of your own work and ideas and that you have given credit to others where their work has contributed to yours.

8. Group Conflicts & Grading Individual Contributions

I expect all members of a group to contribute to the solution and attend group meetings. The groups will be graded as a whole, and every member of the group will receive the same grade.

If you think that a member of your group is not contributing to the solution or/and to the report, please contact me as soon as possible as a first action. Informing me about these issues at the last minute, before submission, will not be helpful.

If you prefer to be marked individually, please submit the Conflict Resolution form on Minerva by 13th of December 23:59 explaining your case. Only one submission is needed for the group project and every member is invited to submit this form and support their individual contributions.

I reserve the right to investigate the contributions of a student to the project solution and the report and adjust individual's marks according to his/her contribution. The claims in the form below would only be used to guide me and I will make my own judgement based on concrete evidence available. It's, therefore, important to demonstrate engagement with the project by frequently documenting your contributions with timestamps.

9. Assessment/ marking criteria grid

For details, see §4.

10. Statement about using deep learning

While you are encouraged to go beyond the algorithms we have learned about in class, deep learning-based methods are out of scope for this module and coursework.

11. Use of Gen AI (Generative Artificial Intelligence) instructions:

This assessment is red category. AI tools cannot be used in submitted work.

References:

H. Cole, M.D. Griffiths, Social interactions in massively multiplayer online role-playing gamers, Cyber Psychol. Behav., 2007

S. Nebel, M. Ninaus, Short research report: does playing apart really bring us together? Investigating the link between perceived loneliness and the use of video games during the COVID-19 pandemic, Frontiers in Psychology, 2022

Jin Yang, Ruoxu Wang, Amy Cook, Rhema Fuller, Gaming during the COVID-19 pandemic: Examining its effect on loneliness & motivation, playing and gratification differences between competitive and recreational gamers, Telematics and Informatics Reports, 2023.

Dunlap, R., K.V. Liere, A. Mertig, and R.E. Jones. 2000. Measuring endorsement of the new ecological paradigm: A revised NEP scale. Journal of Social Issues 56: 425–442. https://doi.org/10.1111/0022-4537.00176.

Vuong QH., Ho MT., La VP., Le TT., Nguyen THT., Nguyen MH.; A Multinational Data Set of Game Players' Behaviors in a Virtual World and Environmental Perceptions. Data Intelligence 2021; 3 (4): 606–630. doi: https://doi.org/10.1162/dint_a_00111