MSc. in Computing
Practicum Approval Form

Section 1: Student Details

Project Title:	Data-Driven Review System
Student ID:	20210415 21263045
Student name:	Kashish Kishinchandani Aditi Bhat
Student email	kashish.kishinchandani2@mail.dcu.ie aditi.bhat2@mail.dcu.ie
Chosen major:	Data Analytics
Supervisor	Prof. Manoj Kesavulu
Date of Submission	28/01/2022

Section 2: About your Practicum

Please answer all questions below. Please pay special attention to the word counts in all cases.

What is the topic of your proposed practicum? (100 words)

The focus of this practicum is to create a data-driven review system to help E-game developers reduce the time invested in understanding how the changes made to the game affect the users.

Please provide details of the papers you have read on this topic (details of 5 papers expected).

1. Tan, Jie Ying, Andy Sai Kit Chow, and Chi Wee Tan. "SENTIMENT ANALYSIS ON GAME REVIEWS: A COMPARATIVE STUDY OF MACHINE LEARNING APPROACHES." *International Conference on Digital Transformation and Applications (ICDXA)*. Vol. 25. 2021.

2. Hu, Minqing, and Bing Liu. "Mining and summarizing customer reviews." *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2004.

3. Yauris, Kevin, and Masayu Leylia Khodra. "Aspect-based summarization for game review using double propagation." *2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA)*. IEEE, 2017.

4. Aljuhani, S.A. and Alghamdi, N.S., 2019. A comparison of sentiment analysis methods on Amazon reviews of Mobile Phones. *Int. J. Adv. Comput. Sci. Appl*, *10*(6), pp.608-617.

5. Viggiato, Markos, et al. "What Causes Wrong Sentiment Classifications of Game Reviews." *IEEE Transactions on Games* (2021).

How does your proposal relate to existing work on this topic described in these papers? (200 words)

The proposal focuses on gathering the data from the user reviews to reduce the time spent by game developers in identifying the areas in which users (gamers) have a concern. In [1], the researchers focus on classifying the reviews using text classification to gather the improvements that can be made in the game. Using the research done [2][3][4], the motive of this practicum is to reduce the time invested by game developers to comprehend the sentiments of the user reviews. Our research focuses on the patch notes and updates released. Using sentiment analysis on the reviews of one specific e-game, we analyze how users are accepting the updates. This review system can then be extended to other platforms altogether.

What are the research questions that you will attempt to answer? (200 words)

- How will a data-driven approach of a review system affect the time invested in review analysis?

How will you explore these questions? (Please address the following points. Note that three or four sentences on each will suffice.)

- **What software and programming environment will you use?**
  - Programming Language: Python
  - Environment: Jupyter Notebooks, Jupyter Labs, AWS/GCP

- What coding/development will you do?
  - Using WEB APIs to gather reviews from gaming review Websites such as https://www.metacritic.com/.
  - Data collection via web scraping from various sources such as social media pages of E-games, E-gaming forums and Reddit, in addition to some Kaggle datasets.
  - Sentimental Analysis on the reviews gathered to identify how changes in the game affect the users.

- What data will be used for your investigations?
  - The dataset will be created from different sources like:
    - Official E-gaming websites for patch notes and update releases
    - E-gaming social media pages
    - Kaggle datasets
    - Official E-gaming forum
    - Other Community forums (Reddit)

- Is this data currently available, it not, where will it come from?
  - Using WEB APIs to gather reviews from gaming review Websites such as https://www.metacritic.com/. Using Reddit API to gather more such user reviews from Reddit

- The Kaggle datasets are readily available.

- What experiments do you expect to run?

- We will perform Sentimental Analysis on the reviews gathered via web scraping to gather insights on the users' feedback regarding the game's updates, patches, and bug fixes.
- We will classify the reviews into useful/valid, spam and hateful by performing Sentimental Analysis. Using these valid reviews, we will categorize based on patch notes with the help of text summarization or feature extraction techniques.
- Classify the valid reviews into positive or negative sentiments based on the category of the patches. Create a train-test split on the dataset and predict the positive/negative sentiments using different classification models and calculate accuracy of each.

- What output do you expect to gather?

- The practicum's output is to build a robust model that uses a data-driven approach to perform Sentimental Analysis on the reviews gathered to reduce the time invested by the game developers.