OPEN INNOVATIONS IN DATA SCIENCE (OIDS) COURSEWORK REPORT

Do not write your name on your work unless your lecturer has explicitly told you to do so.

Student ID number	Title of degree studying	Level/Year
2200918	Bachelor of Science (Honours) Data Science and Analytics	6

Short unit name:	M32366 - OIDS		Due date: 14 March 24	Deadline: 14 Mar
Full unit name:	OPEN INNOVATIONS IN DATA SCIENCE			
Unit lecturer name:	Mr. Jackson Maximillian Yap		Group: *(if applicable)*	
Additional items e.g. CD/disk/USB:	Yes	No ✔	Details:	

All additional items should be clearly labelled with ID number and unit name and securely attached to your work.

Candidates are reminded that the following are defined as Assessment Offences and will be dealt with in accordance with the University's Code of Student Discipline:

a)   Any attempt to complete an assessment by means considered to be <u>unfair.</u>

b)   <u>Plagiarism,</u> which the University defines as the incorporation by a student in work for assessment of material which is not their own, in the sense that all or a substantial part of the work has been copied without any adequate attempt at attribution or has been incorporated as if it were the student's own work when in fact it is wholly or substantially the work of another person or **persons.**

Please note: Group **coursework** will be filed under the first Student ID number at the top of the list. Ensure you know <u>all</u> group member's ID numbers.

NB: **Coursework not collected** will be disposed of six months **after** the hand-in date.

<u>FOR OFFICIAL USE ONLY</u>

Date received/Office <u>stamp</u>	Provisional mark % / Comments

Administration Office   Academic Staff Member

Team Members:

UP2200910, UP2200923, UP2200918, UP2200933, UP2200919, UP2200934

Recording Link

Table of Contents

Evolution of your mini project (How we deploy Crisp DM)

(597 words)

Leveraging the CRISP-DM framework, our project **aimed at reversing declining viewership** trends through a sophisticated Movie Recommender System. This system, envisioned **to enhance user engagement** by tailoring movie suggestions to individual preferences, is a testament to the power of prescriptive analytics and data mining. The CRISP-DM methodology, with its structured phases, provided a rigorous approach to tackling the project, each phase building upon the last towards achieving our objective.
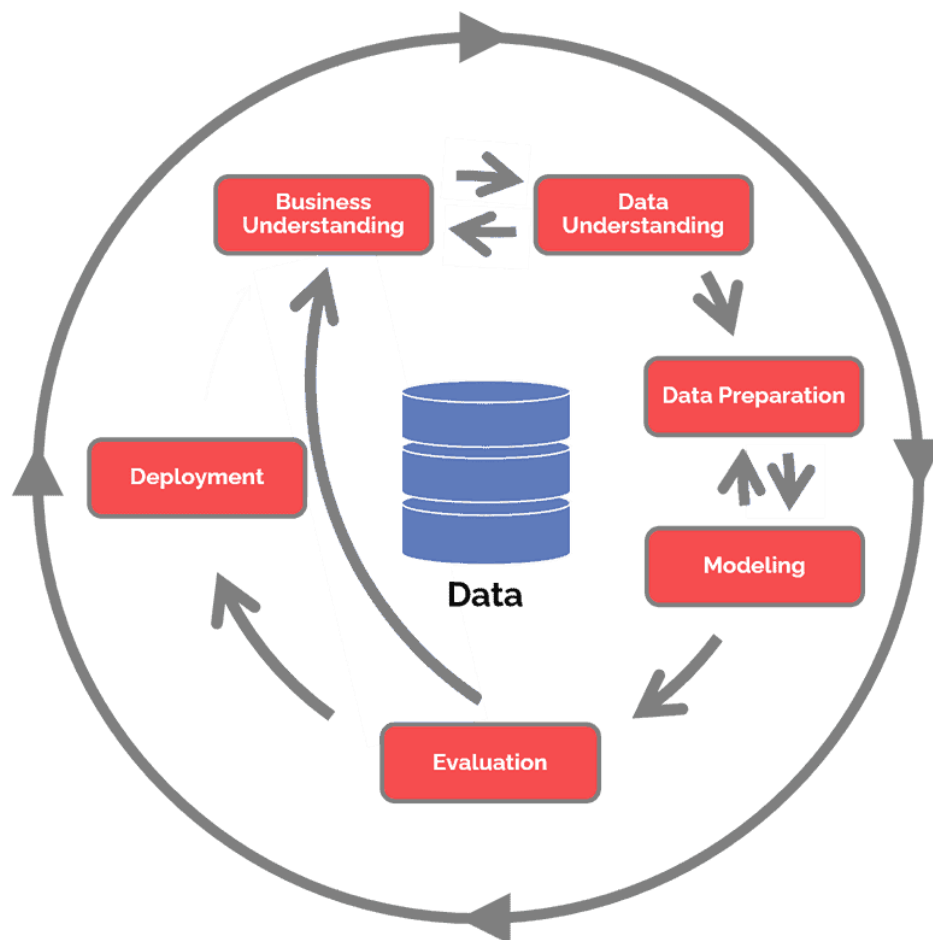
v

Figure 1: CRISP-DM (Source: Hotz, N., January 19, 2023, Data Science Process Alliance)

Below is a detailed account of how each phase was meticulously applied in the context of our movie recommendation system project.

1. Business Understanding

Our project's foundation was the articulation of clear business objectives aimed at **boosting user engagement** to **counteract falling viewership** numbers. Key performance indicators were **set increased viewer engagement time, higher retention rates, and positive user feedback** on the recommendations. This stage was pivotal in ensuring that our project's goals were in line with overarching corporate strategies.

2. Data Understanding

The project delved into a comprehensive examination of available data, encompassing user profiles, historical viewing patterns, and detailed movie metadata. This crucial phase helped **identify the scope and quality of data essential** for the recommendation system's development, laying a solid foundation based on accurate and relevant data insights.

3. Data Preparation

The preparation of data involved meticulous tasks such as **data cleaning, integration, and transformation**. New features were engineered to **enhance model performance**, and datasets were combined and reformatted to ensure compatibility with our chosen algorithms. This phase ensured the data was primed for modelling, laying the groundwork for effective recommendation algorithms.

4. Modelling

The exploration and implementation of modelling techniques formed the core of this phase. We experimented with collaborative and content-based filtering, each offering unique benefits in movie recommendation. Collaborative filtering **harnessed user similarity metric**s, whereas content-based filtering **focused on movie attributes**. A hybrid model, **combining these two approaches**, was ultimately chosen for its **enhanced ability to deliver personalized and diverse suggestions.**

5. Evaluation

This phase was dedicated to rigorously testing the hybrid model against our defined success criteria. By closely **monitoring user interaction** and **gathering feedback**, we **assessed the model's effectiveness in delivering relevant and satisfactory** movie recommendations. The

evaluation ensured the system's readiness for deployment, confirming its potential to meet our business objectives.

6. Deployment

The culmination of the project involved the strategic integration of the recommendation system into the existing platform. This phase addressed key technical infrastructure considerations, established a comprehensive monitoring plan for **performance tracking**, and outlined procedures **for regular maintenance and updates**. The deployment strategy was meticulously crafted to introduce the new system to users smoothly, prioritizing efficiency and minimal disruption.

7. Project Success and Conclusion

The project's impact was clearly demonstrated through the achievement of our objectives: **notable improvements in user engagement metrics, an increase in average viewing times**, and overwhelmingly **positive feedback on the recommendation quality.** The successful deployment of the recommendation system represented a crucial stride towards **reversing the trend of declining viewership,** showcasing the effectiveness of a disciplined problem-solving approach and the transformative potential of prescriptive analytics in elevating user experience.

In wrapping up, the CRISP-DM framework's application significantly contributed to the successful development of the Movie Recommender System. By carefully progressing through each of the framework's phases, our team not only met the project's immediate goals but also laid down a strategic blueprint for future endeavours in analytics. This project highlights the indispensable role of structured data mining and prescriptive analytics in driving forward business objectives **through informed decision-making and customized user engagements, setting a new standard for industry practices.**

Lessons learnt (158 words):

New Technology: The project was an excellent opportunity to deepen my understanding of data science technologies, particularly in the areas of collaborative and content-based filtering techniques. Utilizing the CRISP-DM framework allowed for a structured approach to the complex problem of personalized recommendations. This experience highlighted the importance of choosing the right algorithm based on the specific needs of the project and the data available, enhancing my technical proficiency in data science tools and methodologies.

Personal Development: On a personal level, the project was instrumental in developing my problem-solving and project management skills. It taught me the value of critical thinking when faced with incomplete and contradictory data, pushing me to find creative solutions to complex problems.

Working autonomously within agreed guidelines honed my ability to manage tasks efficiently and make informed decisions with minimal supervision. Moreover, navigating the challenges of this project improved my resilience and adaptability, skills that are crucial in the ever-evolving field of data science.

In your opinion, have you met the learning outcomes of this module? (99 words)

Through the Movie Recommender System project, I've met the module's learning outcomes by critically examining complex data to address real-world challenges, notably in handling and analysing diverse and incomplete datasets. I applied data science knowledge innovatively, using a mix of collaborative and content-based filtering techniques to devise a solution that enhances user engagement.

Working autonomously, I navigated the project's demands with minimal supervision, showcasing the ability to tackle new domain problems using data science skills effectively. This project not only demonstrated a thorough understanding of data science principles but also their practical application in developing a tangible, user-centric solution.

References

1. Hotz, N. (2023, January 19). What is CRISP DM? - Data Science Process Alliance. Data Science Process Alliance. https://www.datascience-pm.com/crisp-dm-2/

2. Smart Vision Europe. (2020, June 17). *Crisp DM methodology - Smart Vision Europe*. https://www.sv-europe.com/crisp-dm-methodology/

3. Yennhi95zz. (2023, March 13). #6. The Deployment Phase of the CRISP-DM Methodology: A Detailed Discussion. *Medium*. https://medium.com/@yennhi95zz/6-the-deployment-phase-of-the-crisp-dm-methodology-a-detailed-discussion-f802a7cb9a0f#:~:text=The%20Deployment%20phase%20of%20the%20CRISP%2DDM%20methodology%20refers%20to,used%20to%20drive%20business%20value.