Assignment 3 **ML** Group Project

short line

Group <put\_group\_number\_here>

| Student Last Name | Student First Name | Student ID | Group Allocation |
| --- | --- | --- | --- |
|  |  |  | Student A (regression) |
|  |  |  | Student B (classification) |
|  |  |  | Student C (clustering) |
|  |  |  | Student D (anomaly) |

36106 - Machine Learning Algorithms and Applications

Master of Data Science and Innovation

University of Technology of Sydney

Table of Contents

[**1. Executive Summary 2**](#_xvtyf97iwmd0)

[**2. Regression (<put student name>) 3**](#_1rui1qefuzi8)

[A. Business Understanding 3](#_q6mtcw9n7wwd)

[1. Business Use Cases 3](#_oe8lnln5357a)

[2. Key Objectives 3](#_cubtjsvjtgfp)

[B. Data Understanding 4](#_vwhwplaw6w2h)

[C. Data Preparation 4](#_ucm82spv7pj)

[D. Modeling 5](#_8pvx1ec4yopz)

[E. Evaluation 5](#_3upide8zhmdg)

[1. Evaluation Metrics 5](#_ef95tucbd03i)

[2. Results and Analysis 5](#_nnw596xflel9)

[3. Business Impact and Benefits 5](#_rmdht2iyi7zh)

[4. Data Privacy and Ethical Concerns 6](#_bfmg88x9b3fb)

[**3. Classification (<put student name>) 7**](#_nr2i2nrqk6md)

[A. Business Understanding 7](#_m7coov7xs2f8)

[1. Business Use Cases 7](#_jvel3s1jotwt)

[2. Key Objectives 7](#_206s3wqcpftx)

[B. Data Understanding 8](#_ftjdngnewk72)

[C. Data Preparation 8](#_bt0efj2o5ojk)

[D. Modeling 9](#_y1rf3zz4vili)

[E. Evaluation 9](#_8t46mi137o1c)

[1. Evaluation Metrics 9](#_dn4cu82ov7xz)

[2. Results and Analysis 9](#_plp4b8r1j5on)

[3. Business Impact and Benefits 9](#_58omp279i818)

[4. Data Privacy and Ethical Concerns 10](#_482jdtgyq5ii)

[**4. Clustering (<put student name>) 11**](#_69lgbngoqmr2)

[A. Business Understanding 11](#_88ygae4sz1iz)

[1. Business Use Cases 11](#_1b8waeb2fhez)

[2. Key Objectives 11](#_bhm3cwa32o1j)

[B. Data Understanding 12](#_f2uo6q62m7oh)

[C. Data Preparation 12](#_piogu2ckvygl)

[D. Modeling 13](#_c42p4nv9yffi)

[E. Evaluation 13](#_ndqhm9bdr3v)

[1. Evaluation Metrics 13](#_ov995qr8xxj8)

[2. Results and Analysis 13](#_yqotml9irkbv)

[3. Business Impact and Benefits 13](#_c1yzdej50l8k)

[4. Data Privacy and Ethical Concerns 14](#_kv2w5l9ojw4b)

[**5. Anomaly Detection (<put student name>) 15**](#_m6tactda1vvw)

[A. Business Understanding 15](#_d1az8qfrrq5a)

[1. Business Use Cases 15](#_b7w0mc5uiyve)

[2. Key Objectives 15](#_uyb4rept6c7r)

[B. Data Understanding 16](#_qa2wy5ozjagv)

[C. Data Preparation 16](#_mn2y1kr1ussz)

[D. Modeling 17](#_e25yirgk9gjz)

[E. Evaluation 17](#_kzr2xmr3yged)

[1. Evaluation Metrics 17](#_o8yls7j3ihaf)

[2. Results and Analysis 17](#_25v57zsj5m7g)

[3. Business Impact and Benefits 17](#_xx3recqgybcd)

[4. Data Privacy and Ethical Concerns 18](#_uf1z6gbsejg6)

[**6. Collaboration 19**](#_tx2wigyi41kv)

[A. Individual Contributions 19](#_5fycwkjr7bjp)

[B. Group Dynamic 19](#_4tp33skt3su4)

[C. Ways of Working Together 19](#_t9y1jnesjjm)

[D. Issues Faced 20](#_oeeyog5w6nd)

[**7. Conclusion 21**](#_cbq8x1dbp07g)

[**8. References 22**](#_wrebp7dsc6jv)

# 

# Executive Summary

* Provide an overview of the project, including its objectives and significance.
* Describe the problem statement and the context in which the project was undertaken.
* State the achieved outcomes and results of the project.

Instructions: In this section, provide a brief introduction to the project, including its goals, relevance, and achieved outcomes. Add a concise summary of the problem statement and the overall context of the project.

short dash

# Regression (<put student name>)

## Business Understanding

### Business Use Cases

* Describe the specific business use cases or scenarios where the project is applied.
* Discuss the challenges or opportunities that motivated the project.

Instructions: Describe the business use cases or scenarios where the project is applied. Discuss the challenges or opportunities that motivated the project, explaining why machine learning algorithms are relevant in this context.

### Key Objectives

* Specify the key objectives or goals of the project.
* Identify the stakeholders and their requirements.
* Explain how the project aims to address these requirements.

Instructions: Specify the key objectives or goals of the project, highlighting the desired outcomes. Identify the stakeholders involved and their specific requirements. Explain how the project aims to address these requirements through the use of machine learning algorithms.

short dash

## Data Understanding

* Provide insights into the dataset used for the project.
* Describe the data sources, data collection methods, and any data limitations.
* Discuss the variables/features present in the dataset and their significance.

Instructions: Describe the dataset used for the project, including its sources and any limitations. Discuss the variables or features present in the dataset and their relevance to the project. Include any exploratory data analysis conducted to understand the data better.

short dash

## Data Preparation

* Describe the steps taken to prepare the data for modelling.
* Discuss the data cleaning, preprocessing, and feature engineering techniques applied.
* Explain the rationale for data splitting strategy used.
* Document any handling of missing values, outliers, or imbalanced data.

Instructions: Describe the data preparation steps taken before modelling. Include details about data cleaning, preprocessing, and feature engineering techniques applied. Explain how missing values, outliers, or imbalanced data were handled and any transformations performed on the dataset. Provide clear explanations on data splitting strategy.

short dash

## Modeling

* Describe the machine learning algorithms used for modeling.
* Discuss the rationale behind selecting these algorithms.
* Explain the parameter tuning and model selection process.

Instructions: Describe the machine learning algorithms used for modeling, providing a rationale for their selection based on the project goals. Explain the process of parameter tuning and model selection. Include details about the algorithms' implementation and any considerations made during the modeling phase.

## Evaluation

### Evaluation Metrics

* Describe the evaluation metrics used to assess the models' performance.
* Explain why these metrics were chosen and how they relate to the project goals.

Instructions: Describe the evaluation metrics used to assess the models' performance, including the specific metrics chosen and their relevance to the project goals.

## Results and Analysis

* Present the results of the model evaluation, including accuracy, precision, recall, F1-score, etc.
* Analyse and compare the performance of each model.
* Discuss the key insights gained during the experimentation phases.

Instructions: Present the results of the model evaluation, including accuracy, precision, recall, F1-score, or any other relevant metrics. Analyze and compare the performance of each model, highlighting the key insights gained during the experimentation phases. Discuss the implications of these insights on the project's goals and potential areas for further improvement.

### Business Impact and Benefits

* Assess the impact and benefits of the final model on the business use cases.
* Discuss how the model contributes to solving the identified challenges or exploiting opportunities.
* Quantify the improvements achieved and the potential value generated.

Instructions: Assess and discuss the impact and benefits of the final model on the identified business use cases. Explain how the model contributes to solving the identified challenges or exploiting opportunities. Quantify the improvements achieved and discuss the potential value generated by the model.

### Data Privacy and Ethical Concerns

* Assess the data privacy implications of the project.
* Discuss any ethical concerns related to data collection, usage, or model deployment.
* Address steps taken to ensure data privacy and ethical considerations.
* Assess potential negative impacts and risks for Indigenous people

Instructions: Assess the data privacy implications of the project, considering any sensitive information or privacy concerns related to data collection, usage, or model deployment. Discuss any ethical concerns and considerations. Address the steps taken to ensure data privacy and mitigate ethical concerns.

short dash

# 

# Classification (<put student name>)

## Business Understanding

### Business Use Cases

* Describe the specific business use cases or scenarios where the project is applied.
* Discuss the challenges or opportunities that motivated the project.

Instructions: Describe the business use cases or scenarios where the project is applied. Discuss the challenges or opportunities that motivated the project, explaining why machine learning algorithms are relevant in this context.

### Key Objectives

* Specify the key objectives or goals of the project.
* Identify the stakeholders and their requirements.
* Explain how the project aims to address these requirements.

Instructions: Specify the key objectives or goals of the project, highlighting the desired outcomes. Identify the stakeholders involved and their specific requirements. Explain how the project aims to address these requirements through the use of machine learning algorithms.

short dash

## Data Understanding

* Provide insights into the dataset used for the project.
* Describe the data sources, data collection methods, and any data limitations.
* Discuss the variables/features present in the dataset and their significance.

Instructions: Describe the dataset used for the project, including its sources and any limitations. Discuss the variables or features present in the dataset and their relevance to the project. Include any exploratory data analysis conducted to understand the data better.

short dash

## Data Preparation

* Describe the steps taken to prepare the data for modelling.
* Discuss the data cleaning, preprocessing, and feature engineering techniques applied.
* Explain the rationale for data splitting strategy used.
* Document any handling of missing values, outliers, or imbalanced data.

Instructions: Describe the data preparation steps taken before modelling. Include details about data cleaning, preprocessing, and feature engineering techniques applied. Explain how missing values, outliers, or imbalanced data were handled and any transformations performed on the dataset. Provide clear explanations on data splitting strategy.

short dash

## Modeling

* Describe the machine learning algorithms used for modeling.
* Discuss the rationale behind selecting these algorithms.
* Explain the parameter tuning and model selection process.

Instructions: Describe the machine learning algorithms used for modeling, providing a rationale for their selection based on the project goals. Explain the process of parameter tuning and model selection. Include details about the algorithms' implementation and any considerations made during the modeling phase.

## Evaluation

### Evaluation Metrics

* Describe the evaluation metrics used to assess the models' performance.
* Explain why these metrics were chosen and how they relate to the project goals.

Instructions: Describe the evaluation metrics used to assess the models' performance, including the specific metrics chosen and their relevance to the project goals.

## Results and Analysis

* Present the results of the model evaluation, including accuracy, precision, recall, F1-score, etc.
* Analyse and compare the performance of each model.
* Discuss the key insights gained during the experimentation phases.

Instructions: Present the results of the model evaluation, including accuracy, precision, recall, F1-score, or any other relevant metrics. Analyze and compare the performance of each model, highlighting the key insights gained during the experimentation phases. Discuss the implications of these insights on the project's goals and potential areas for further improvement.

### Business Impact and Benefits

* Assess the impact and benefits of the final model on the business use cases.
* Discuss how the model contributes to solving the identified challenges or exploiting opportunities.
* Quantify the improvements achieved and the potential value generated.

Instructions: Assess and discuss the impact and benefits of the final model on the identified business use cases. Explain how the model contributes to solving the identified challenges or exploiting opportunities. Quantify the improvements achieved and discuss the potential value generated by the model.

### Data Privacy and Ethical Concerns

* Assess the data privacy implications of the project.
* Discuss any ethical concerns related to data collection, usage, or model deployment.
* Address steps taken to ensure data privacy and ethical considerations.
* Assess potential negative impacts and risks for Indigenous people

Instructions: Assess the data privacy implications of the project, considering any sensitive information or privacy concerns related to data collection, usage, or model deployment. Discuss any ethical concerns and considerations. Address the steps taken to ensure data privacy and mitigate ethical concerns.

short dash

# 

# Clustering (<put student name>)

## Business Understanding

### Business Use Cases

* Describe the specific business use cases or scenarios where the project is applied.
* Discuss the challenges or opportunities that motivated the project.

Instructions: Describe the business use cases or scenarios where the project is applied. Discuss the challenges or opportunities that motivated the project, explaining why machine learning algorithms are relevant in this context.

### Key Objectives

* Specify the key objectives or goals of the project.
* Identify the stakeholders and their requirements.
* Explain how the project aims to address these requirements.

Instructions: Specify the key objectives or goals of the project, highlighting the desired outcomes. Identify the stakeholders involved and their specific requirements. Explain how the project aims to address these requirements through the use of machine learning algorithms.

short dash

## Data Understanding

* Provide insights into the dataset used for the project.
* Describe the data sources, data collection methods, and any data limitations.
* Discuss the variables/features present in the dataset and their significance.

Instructions: Describe the dataset used for the project, including its sources and any limitations. Discuss the variables or features present in the dataset and their relevance to the project. Include any exploratory data analysis conducted to understand the data better.

short dash

## Data Preparation

* Describe the steps taken to prepare the data for modelling.
* Discuss the data cleaning, preprocessing, and feature engineering techniques applied.
* Explain the rationale for data splitting strategy used.
* Document any handling of missing values, outliers, or imbalanced data.

Instructions: Describe the data preparation steps taken before modelling. Include details about data cleaning, preprocessing, and feature engineering techniques applied. Explain how missing values, outliers, or imbalanced data were handled and any transformations performed on the dataset. Provide clear explanations on data splitting strategy.

short dash

## Modeling

* Describe the machine learning algorithms used for modeling.
* Discuss the rationale behind selecting these algorithms.
* Explain the parameter tuning and model selection process.

Instructions: Describe the machine learning algorithms used for modeling, providing a rationale for their selection based on the project goals. Explain the process of parameter tuning and model selection. Include details about the algorithms' implementation and any considerations made during the modeling phase.

## Evaluation

### Evaluation Metrics

* Describe the evaluation metrics used to assess the models' performance.
* Explain why these metrics were chosen and how they relate to the project goals.

Instructions: Describe the evaluation metrics used to assess the models' performance, including the specific metrics chosen and their relevance to the project goals.

## Results and Analysis

* Present the results of the model evaluation, including accuracy, precision, recall, F1-score, etc.
* Analyse and compare the performance of each model.
* Discuss the key insights gained during the experimentation phases.

Instructions: Present the results of the model evaluation, including accuracy, precision, recall, F1-score, or any other relevant metrics. Analyze and compare the performance of each model, highlighting the key insights gained during the experimentation phases. Discuss the implications of these insights on the project's goals and potential areas for further improvement.

### Business Impact and Benefits

* Assess the impact and benefits of the final model on the business use cases.
* Discuss how the model contributes to solving the identified challenges or exploiting opportunities.
* Quantify the improvements achieved and the potential value generated.

Instructions: Assess and discuss the impact and benefits of the final model on the identified business use cases. Explain how the model contributes to solving the identified challenges or exploiting opportunities. Quantify the improvements achieved and discuss the potential value generated by the model.

### Data Privacy and Ethical Concerns

* Assess the data privacy implications of the project.
* Discuss any ethical concerns related to data collection, usage, or model deployment.
* Address steps taken to ensure data privacy and ethical considerations.
* Assess potential negative impacts and risks for Indigenous people

Instructions: Assess the data privacy implications of the project, considering any sensitive information or privacy concerns related to data collection, usage, or model deployment. Discuss any ethical concerns and considerations. Address the steps taken to ensure data privacy and mitigate ethical concerns.

short dash

# 

# Anomaly Detection (<put student name>)

## Business Understanding

### Business Use Cases

* Describe the specific business use cases or scenarios where the project is applied.
* Discuss the challenges or opportunities that motivated the project.

Instructions: Describe the business use cases or scenarios where the project is applied. Discuss the challenges or opportunities that motivated the project, explaining why machine learning algorithms are relevant in this context.

### Key Objectives

* Specify the key objectives or goals of the project.
* Identify the stakeholders and their requirements.
* Explain how the project aims to address these requirements.

Instructions: Specify the key objectives or goals of the project, highlighting the desired outcomes. Identify the stakeholders involved and their specific requirements. Explain how the project aims to address these requirements through the use of machine learning algorithms.

short dash

## Data Understanding

* Provide insights into the dataset used for the project.
* Describe the data sources, data collection methods, and any data limitations.
* Discuss the variables/features present in the dataset and their significance.

Instructions: Describe the dataset used for the project, including its sources and any limitations. Discuss the variables or features present in the dataset and their relevance to the project. Include any exploratory data analysis conducted to understand the data better.

short dash

## Data Preparation

* Describe the steps taken to prepare the data for modelling.
* Discuss the data cleaning, preprocessing, and feature engineering techniques applied.
* Explain the rationale for data splitting strategy used.
* Document any handling of missing values, outliers, or imbalanced data.

Instructions: Describe the data preparation steps taken before modelling. Include details about data cleaning, preprocessing, and feature engineering techniques applied. Explain how missing values, outliers, or imbalanced data were handled and any transformations performed on the dataset. Provide clear explanations on data splitting strategy.

short dash

## Modeling

* Describe the machine learning algorithms used for modeling.
* Discuss the rationale behind selecting these algorithms.
* Explain the parameter tuning and model selection process.

Instructions: Describe the machine learning algorithms used for modeling, providing a rationale for their selection based on the project goals. Explain the process of parameter tuning and model selection. Include details about the algorithms' implementation and any considerations made during the modeling phase.

## Evaluation

### Evaluation Metrics

* Describe the evaluation metrics used to assess the models' performance.
* Explain why these metrics were chosen and how they relate to the project goals.

Instructions: Describe the evaluation metrics used to assess the models' performance, including the specific metrics chosen and their relevance to the project goals.

## Results and Analysis

* Present the results of the model evaluation, including accuracy, precision, recall, F1-score, etc.
* Analyse and compare the performance of each model.
* Discuss the key insights gained during the experimentation phases.

Instructions: Present the results of the model evaluation, including accuracy, precision, recall, F1-score, or any other relevant metrics. Analyze and compare the performance of each model, highlighting the key insights gained during the experimentation phases. Discuss the implications of these insights on the project's goals and potential areas for further improvement.

### Business Impact and Benefits

* Assess the impact and benefits of the final model on the business use cases.
* Discuss how the model contributes to solving the identified challenges or exploiting opportunities.
* Quantify the improvements achieved and the potential value generated.

Instructions: Assess and discuss the impact and benefits of the final model on the identified business use cases. Explain how the model contributes to solving the identified challenges or exploiting opportunities. Quantify the improvements achieved and discuss the potential value generated by the model.

### Data Privacy and Ethical Concerns

* Assess the data privacy implications of the project.
* Discuss any ethical concerns related to data collection, usage, or model deployment.
* Address steps taken to ensure data privacy and ethical considerations.
* Assess potential negative impacts and risks for Indigenous people

Instructions: Assess the data privacy implications of the project, considering any sensitive information or privacy concerns related to data collection, usage, or model deployment. Discuss any ethical concerns and considerations. Address the steps taken to ensure data privacy and mitigate ethical concerns.

short dash

# Collaboration

## Individual Contributions

* Describe the individual contributions of each team member.
* Explain the specific tasks, responsibilities, or areas of expertise assigned to each team member.
* Highlight any significant contributions or achievements made by individual team members.

Instructions: In this section, each team member should provide a summary of their individual contributions to the project. Explain the specific tasks or responsibilities assigned to each team member and describe their contributions in terms of data preparation, modeling, evaluation, or other project-related activities. Highlight any notable achievements or contributions made by individual team members.

## Group Dynamic

* Describe the overall group dynamic and collaboration within the team.
* Discuss how team members communicated, coordinated, and shared responsibilities.
* Highlight any strategies or practices adopted to ensure effective teamwork.

Instructions: Reflect on the group dynamic and collaboration within the team. Describe how team members interacted, communicated, and coordinated their efforts throughout the project. Discuss any strategies, tools, or practices that were adopted to ensure effective teamwork and smooth collaboration.

## Ways of Working Together

* Explain the methodologies or frameworks used to manage the project within the team.
* Discuss the frequency and format of team meetings, progress tracking, and decision-making processes.
* Highlight any tools or technologies utilized to facilitate collaboration and project management.

Instructions: Explain the methodologies or frameworks employed to manage the project within the team. Describe the frequency and format of team meetings, the approach to progress tracking, and the decision-making processes followed. Highlight any specific tools or technologies that were utilized to facilitate collaboration and project management.

## Issues Faced

* Identify any challenges, issues, or obstacles encountered during the project.
* Describe how these challenges were addressed and resolved within the team.
* Discuss any lessons learned or recommendations for future group collaborations.

Instructions: Identify and discuss any challenges, issues, or obstacles that the team encountered during the project. Describe how these challenges were addressed and resolved within the team, including any strategies or actions taken to overcome them. Reflect on the lessons learned and provide recommendations for improving future group collaborations.

short dash

# Conclusion

* Summarize the key findings, insights, and outcomes of the project.
* Reflect on the project's success in achieving its goals and meeting stakeholders' requirements.
* Discuss any future work, recommendations, or next steps based on the project's outcomes.

Instructions: Summarize the key findings, insights, and outcomes of the project. Reflect on the project's success in achieving its goals and meeting stakeholders' requirements. Discuss any future work, recommendations, or next steps based on the project's outcomes.

short dash

# References

* Include a list of references used throughout the project report.

Instructions: Include a list of references used throughout the project report, following the appropriate citation style.

short dash

Note: The CRISP-DM steps (Cross-Industry Standard Process for Data Mining) provide a framework for structuring the project report, but feel free to adapt the template to match the specific requirements and guidelines of your project or organization.