INF80051 Artificial Intelligence and Insights - Assignment 3 Rubric

Bank Marketing Campaign Analytics (Total: 35 marks)

Machine Learning Pipeline for Term Deposit Prediction

Criterion	Marks	High Distinction	Distinction	Credit	Pass	Fail
1. Problem Definition *ML task formulation*	**5**	Excellent formulation of binary classification problem. Clear definition of target variable and prediction task. Well-articulated objectives for model performance. Understanding of supervised learning task demonstrated. Discussion of why this is a classification problem and what constitutes success.	Good problem formulation with clear classification objectives. Target variable properly identified. Model performance goals stated. Understanding of ML task type shown.	Adequate problem definition. Classification task identified. Basic objectives stated. Target variable mentioned.	Basic problem definition. Unclear understanding of classification task. Minimal objectives. Target variable identified but poorly explained.	No clear problem formulation. Task type not understood. No objectives. Target variable not properly identified.
2. Data Exploration *EDA and visualization*	**6**	Exceptional data exploration with 10+ publication-quality ggplot visualizations (faceted plots, correlation matrices, custom themes). Professional dashboard showing distributions, patterns, outliers, correlations. Deep analysis of feature relationships. Class imbalance (11.3%) clearly identified and quantified. Comprehensive summary statistics with insights.	Thorough data exploration with 6-8 well-designed ggplot visualizations (multivariate plots, faceted charts). Informative dashboard with multiple panels. Feature distributions analyzed. Class imbalance identified and measured. Good statistical summaries.	Good exploration with 4-5 ggplot visualizations showing key patterns (histograms, box plots, correlations). Basic dashboard created. Class imbalance noted. Standard summary statistics shown.	Basic exploration with 2-3 simple ggplot visualizations. Basic summary statistics. Class imbalance may be mentioned. Limited analysis.	No meaningful EDA. No or minimal visualizations. Critical issues like 11.3% class imbalance completely missed. No understanding of data.
3. Data Preprocessing *Data cleaning pipeline*	**7**	Comprehensive preprocessing pipeline: systematic handling of missing values with comparison of imputation methods (mean, median, mode) and justified selection. Outliers identified and handled with clear rationale. Duration variable removed with explanation. Feature scaling applied with justification. Categorical encoding properly implemented. Feature engineering with logic explained. Full reproducible pipeline documented.	Thorough preprocessing: missing values analyzed and imputation method justified. Outliers detected and handled with rationale. Duration variable removed with explanation. Feature scaling applied. Categorical encoding correct. Some feature engineering. Clear pipeline steps.	Good preprocessing: missing values handled with basic justification. Outliers identified. Duration variable removed with brief explanation. Feature scaling applied. Categorical variables encoded. Basic pipeline shown.	Basic preprocessing: simple missing value handling. Duration variable removed. Basic categorical encoding. Minimal justification. Incomplete pipeline.	No preprocessing or major errors. Duration variable kept (critical error). Raw data used. No pipeline. Critical preprocessing steps missing.

Criterion	Marks	High Distinction	Distinction	Credit	Pass	Fail
4. Class Imbalance Handling *SMOTE implementation*	**5**	SMOTE optimally implemented with parameters justified. Detailed before/after analysis showing impact on class distribution. Applied correctly after train-test split (no data leakage). Stratified sampling used. Clear visualization of balanced dataset. Discussion of why SMOTE is needed for 11.3% minority class.	SMOTE properly implemented with parameters explained. Good analysis of impact on class distribution. No data leakage. Stratified split used. Visualization of results shown.	SMOTE applied with basic parameter settings. Some analysis of impact. Proper train-test split maintained. Basic understanding demonstrated.	SMOTE attempted but may have minor implementation issues. Limited analysis of impact. Basic validation approach used.	No imbalance handling despite severe 11.3% minority class. Data leakage present. Complete misunderstanding of the problem.
5. Model Development and Tuning *Model building and optimization*	**6**	3-4 diverse models implemented (logistic regression, decision trees, neural networks with 5+ architecture variations, ensemble methods). Comprehensive hyperparameter tuning with systematic grid/random search. All choices justified with technical reasoning. Cross-validation properly implemented. Excellent code documentation with clear pipeline.	2-3 models implemented with multiple variations tested (e.g., neural networks with different architectures - varying layers/neurons). Good hyperparameter tuning performed. Technical justifications provided. Cross-validation used. Well-documented code.	2 models implemented (e.g., logistic regression + decision tree or neural network). Basic hyperparameter tuning attempted. Some technical justification. Standard train-test validation. Adequate code documentation.	1 basic model implemented (logistic regression or decision tree). Minimal hyperparameter tuning. Poor justification. No cross-validation. Limited code comments.	No models properly implemented or code fails. No hyperparameter tuning. No justification. Models produce errors or meaningless results.
6. Model Evaluation and Analysis *Performance assessment and justification*	**6**	Comprehensive evaluation using appropriate metrics for imbalanced data: confusion matrices, precision, recall, F1-score, ROC curves, AUC for all models. Detailed comparison with strengths/weaknesses analysis. Critical discussion of why models perform well or poorly. Variable importance analyzed. Threshold optimization explored. Clear recommendation with technical justification. Understanding of metric trade-offs demonstrated.	Thorough evaluation: confusion matrices, precision, recall, F1-score, ROC curves, AUC reported for all models. Good comparison of model performance. Analysis of why models succeed or fail. Variable importance examined. Clear conclusions with justification.	Good evaluation: confusion matrices, precision, recall, F1-score for each model. Basic model comparison performed. Some analysis of performance differences. Adequate conclusions drawn.	Basic evaluation: accuracy and 1-2 other metrics. Simple confusion matrix. Limited model comparison. Weak analysis of results. Minimal justification of conclusions.	No proper evaluation. Only accuracy reported (wrong for imbalanced data). No confusion matrices. No model comparison. No understanding of results.
7. Report Quality *Documentation and presentation*	**0**	Professional report with clear ML pipeline structure. Modular R code with functions. All steps thoroughly documented. Perfect word count (4000 $\pm$2%). Publication-quality formatting.	Well-structured report with logical flow. All code documented with clear comments. Good explanations of technical choices. Within word count (4000 $\pm$5%).	Clear report structure. Code has section comments. Technical steps explained. Within word count (4000 $\pm$10%).	Basic report structure. R code included with minimal comments. Within word count (4000 $\pm$20%).	No report or missing code. Word count ¡2000 or ¿5000. No clear structure.
TOTAL: 35 marks		5 + 6 + 7 + 5 + 6 + 6 + 0 = 35 marks				

Important Technical Requirements

What This Assignment Tests

This is a **machine learning pipeline assignment**. You must demonstrate your ability to:

- Clean and preprocess data systematically
- Handle class imbalance appropriately
- Build and tune multiple predictive models
- Evaluate model performance with appropriate metrics
- Provide technical justifications for all decisions
- Critically analyze what works and what doesn't

Mandatory Requirements

- **Word Count:** 4000 words ±10% (main body only, excluding references, code appendices, and figures)
- **Duration Variable:** MUST be excluded from all models (it's only known after the call - using it is a critical error)
- **Class Imbalance:** Dataset has only 11.3% positive cases - SMOTE is required
- **Code:** All R code must be in appendix with clear comments explaining each step
- **Visualizations:** Use ggplot2 for all visualizations
- **Metrics:** Must report precision, recall, F1-score, AUC - not just accuracy
- **Data Leakage:** Apply SMOTE only to training data AFTER train-test split

Techniques You Should Use

These are the techniques covered in INF80051:

- **Data Preprocessing:** Missing value imputation (mean, median, mode), outlier detection, feature scaling, categorical encoding
- **Visualizations:** ggplot2 plots (histograms, box plots, scatter plots, correlation matrices, faceted plots)
- **Imbalance Handling:** SMOTE only
- **Models:** Logistic Regression, Decision Trees, Neural Networks (vary architectures), Ensemble Methods
- **Evaluation:** Confusion matrices, ROC curves, AUC, precision, recall, F1-score, cross-validation

How Grades Differ

Higher grades = deeper application of the SAME techniques:

- **Pass:** Apply techniques correctly, basic analysis
- **Credit:** More thorough application, better justifications
- **Distinction:** Systematic comparison of options, detailed analysis of results, strong technical justifications
- **High Distinction:** Comprehensive experimentation, critical evaluation, deep understanding shown through analysis of why things work or don't work