Project n^{o} [1/2/3]: [Your Project Title]
Advanced Machine Learning (MDS)

Student Name 1 Student Name 2

Date:

Abstract

[Your abstract here]

1 Introduction

[Your introduction here. Start with the problem and its motivation. Then describe your approach and main findings.]

2 Problem statement

3 Related Work

Algorithm 1 Algorithm name

Require: Input data X, parameters θ_0
Ensure: Optimized parameters θ^*
1: Initialize $\theta \leftarrow \theta_0$
2: **while** not converged **do**
3: Compute gradient $g \leftarrow \nabla_\theta L(\theta)$
4: Update $\theta \leftarrow \theta - \alpha g$
5: **end while**
6: **return** θ

4 Data and Preprocessing

4.1 Data description

[Your data description here]

Table 1: Dataset characteristics: add/remove rows if needed; replace the '?'s by the corresponding values

Property	**Variable**	**Value**
Number of observations	n	?
Number of features	d	NA
Feature types	continuous, categorical	?/?/ ...
Number of classes	C (for classification)	?
Class distribution	balanced/imbalanced	?/?/..
Missing values	percentage	?

4.2 Exploratory data analysis

4.3 Preprocessing steps

5 Methodology

5.1 Experimental protocol

[Your experimental protocol here]

5.2 Method 1: [Method name]

5.2.1 Model formulation

For example, for a GLM, present the likelihood and regularized objective.

The model minimizes the regularized empirical risk:

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell(y_i, f(x_i; \theta)) + \lambda R(\theta) \right\} \tag{1}$$

where ℓ is the loss function, $f(\cdot; \theta)$ is the prediction function, $R(\theta)$ is the regularization term, and $\lambda > 0$ controls regularization strength. Recommended level of description:

one paragraph for MLAs seen in class

half page for MLAs not seen in class but mentioned

up to you for MLAs not seen in class or own work

[Your mathematical formulation here]

5.2.2 Theoretical properties and justification

[Your theoretical discussion here]

5.2.3 Implementation details

[Your implementation details here]

Table 2: Hyperparameters for Method 1

Hyperparameter	**Search space**	**Best value**
Regularization λ	$\{10^{-4}, 10^{-3}, \ldots, 10^{2}\}$	10^{-2}
Learning rate	$\{0.001, 0.01, 0.1\}$	0.01

5.3 Method 2: [Method name]

[Your description of Method 2 following the same structure as Method 1]

5.4 Comparison framework

[Your comparison framework here]

6 Discussion of results

6.1 Overall performance

[Your overall results here]

Table 3: Performance comparison of methods on test set. Results shown as mean ± standard deviation over 5-fold cross-validation. Bold indicates best performance.

Method	Accuracy	F1-score	AUC	Training time (s)
Method 1	0.85 ± 0.02	0.83 ± 0.03	0.90 ± 0.02	15.3
Method 2	$\mathbf{0.88 \pm 0.02}$	$\mathbf{0.86 \pm 0.02}$	$\mathbf{0.92 \pm 0.01}$	127.8

6.2 Detailed analysis

[Your detailed analysis here with insightful interpretation]

6.3 Model selection and generalization error

[Your final model selection and generalization error estimate here]

7 Conclusions

References

A Additional results

[Your additional results here if needed]

B Mathematical derivations

Lemma B.1. *State your lemma here with precise mathematical notation.*

Proof. Provide detailed proof here. □

[Your derivations here if needed]

C Implementation Details

[Your implementation details here if needed]