2026 Volume: 02 Issue: 05 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-03

L1 REGULARIZATION USING GRADIENT DESCENT ON CALIFORNIA HOUSING DATASET

*Sumanaa B. S., Samartha B. S.

Department of Computer Science and Engineering, Jyothy Institute of Technology, Bengaluru, India.

Article Received: 04 April 2026
Article Revised: 24 April 2026
Published on: 14 May 2026

***Corresponding Author: Sumanaa B. S.**
Department of Computer Science and Engineering, Jyothy Institute of Technology, Bengaluru, India.

DOI: https://doi-doi.org/101555/ijrpa.3829

ABSTRACT

Machine learning regression models often face the challenge of overfitting when trained on real-world datasets. Regularization techniques are widely used to improve generalization performance. Among these, L1 regularization (Lasso) is particularly effective as it promotes sparsity by shrinking less important feature coefficients to zero, thereby performing implicit feature selection. However, the optimization of L1 regularization is challenging due to its non-differentiability at zero. This survey paper analyzes the use of gradient descent and its variants for training L1-regularized models on the California Housing dataset. It reviews existing research on optimization techniques such as sub-gradient descent, coordinate descent, and stochastic gradient descent. The study highlights the effectiveness of L1 regularization in reducing overfitting and improving model interpretability.

1. INTRODUCTION

Regression models are widely used in machine learning to predict continuous values. However, when models become complex or contain many features, they tend to overfit the training data, resulting in poor generalization on unseen data. L1 regularization, also known as Lasso, adds a penalty proportional to the absolute value of model parameters. This results in sparse models where irrelevant features are eliminated. Gradient descent is a widely used optimization technique in machine learning. Applying it directly to L1 regularization requires modifications such as sub-gradient methods. This paper explores how gradient descent can be adapted for L1 regularization and evaluates its performance using the California Housing dataset.

2. METHODOLOGY

This survey is based on a comparative analysis of L1 regularization techniques and gradient-based optimization methods. • Reviewing research papers on L1 regularization and sparse learning.

- Studying gradient descent and its variants for non-smooth optimization.
- Analyzing the behavior of model coefficients under L1 penalty.
- Evaluating performance based on convergence speed, sparsity, and prediction accuracy.

Methods discussed include L1 Regularization (Lasso), Gradient Descent, Sub-Gradient Method, Stochastic Gradient Descent (SGD), and Coordinate Descent.

3. SYSTEM ARCHITECTURE AND DATA FLOW

The system architecture for L1-regularized regression consists of the following components:

1. Input Layer – Receives dataset features from the California Housing dataset.
2. Processing Layer – Applies linear regression and computes predictions.
3. Regularization Component – Adds L1 penalty to the loss function.
4. Optimization Layer – Uses gradient descent or its variants for parameter updates.
5. Output Layer – Produces predicted housing prices.

Data Flow Process:

- Input data is normalized and fed into the model.
- Predictions are computed using current parameters.
- Loss is calculated with L1 penalty.
- Gradients are computed and parameters are updated.
- The process repeats until convergence is achieved.

4. RESULT AND DISCUSSION

Effect of L1 Regularization:

- Reduces overfitting.
- Eliminates irrelevant features.
- Produces sparse models.

Gradient Descent Behavior:

- Converges slower compared to L2 regularization.
- Sensitive to learning rate.
- Requires careful tuning.

Key Findings:

- L1 regularization effectively performs feature selection.
- Sub-gradient methods enable optimization of non-smooth functions.
- SGD improves scalability for large datasets.
- Coordinate descent provides faster convergence for L1 problems.

5. CONCLUSION

L1 regularization plays a crucial role in improving regression models by reducing overfitting and enabling feature selection. Although it introduces optimization challenges due to non-differentiability, gradient descent-based methods such as sub-gradient descent and coordinate descent provide effective solutions. The study concludes that combining L1 regularization with gradient-based optimization techniques leads to efficient, sparse, and interpretable models. Future work can focus on advanced optimization techniques such as proximal gradient descent and Elastic Net.

6. ACKNOWLEDGMENT

I would like to express my sincere gratitude to my faculty guide for valuable guidance and support throughout this study. I also thank my institution for providing the necessary resources and environment to carry out this work.

7. REFERENCES

1. R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” 1996.
2. J. Duchi et al., “Efficient Projections for L1-Regularization,” 2008.
3. S. Shalev-Shwartz, “Stochastic Gradient Descent for L1-Regularized Models,” 2011.
4. Friedman et al., “Regularization Paths for Generalized Linear Models,” 2010.
5. Scikit-learn Documentation – California Housing Dataset.
6. Goodfellow, Bengio, Courville – Deep Learning, MIT Press.
7. Research papers on Coordinate Gradient Descent.