

Logistic Regression

Introduction

Logistic Regression is a popular supervised learning algorithm used for binary classification tasks. Despite its name, logistic regression is used for classification rather than regression. It models the probability that a given input belongs to a specific class.

Model Representation

- 1. Sigmoid Function:** In logistic regression, the output of the algorithm is passed through a sigmoid function, which maps any real value into the range $[0, 1]$. The sigmoid function is defined as: $\sigma(z) = \frac{1}{1 + e^{-z}}$.
- 2. Decision Boundary:** The decision boundary in logistic regression is a threshold value (usually 0.5) applied to the output of the sigmoid function. If the output is above the threshold, the input is classified as belonging to one class; otherwise, it is classified as belonging to the other class.

Training a Logistic Regression Model

1. **Objective Function:** The training objective of logistic regression is to maximize the likelihood of the observed data given the model parameters. This is typically done using the maximum likelihood estimation or gradient descent optimization.
2. **Cost Function:** The cost function used in logistic regression is the log loss or binary cross-entropy loss, which penalizes the model for making incorrect predictions. The cost function is minimized during training to improve the model's performance.

Evaluation Metrics

1. **Accuracy:** Accuracy measures the proportion of correctly classified instances out of the total instances.
2. **Precision and Recall:** Precision measures the proportion of true positive predictions among all positive predictions, while recall measures the proportion of true positive predictions among all actual positive instances.

Regularization Techniques

Similar to linear regression, logistic regression models can also benefit from regularization techniques such as **L1 (Lasso)** and **L2 (Ridge)** regularization to prevent overfitting and improve generalization to unseen data.

Applications of Logistic Regression

Logistic regression is widely used in various fields such as healthcare (for disease diagnosis), marketing (customer segmentation), and finance (credit scoring) due to its simplicity, efficiency, and interpretability in modeling binary classification problems.

Conclusion

Logistic Regression is a powerful algorithm for binary classification tasks, providing a probabilistic interpretation of the predicted class labels. By modeling the relationship between input features and the likelihood of belonging to a specific class, logistic regression offers valuable insights for decision-making in various domains.

Producer: Elham Jafari

Computer Engineering