

Data Version Control (DVC) in Machine Learning Projects

Nhóm MLOps

Ngày 18 tháng 10 năm 2025

Nội dung về DVC được chia thành 2 phần chính, phần 1 giải thích Logistic Regression đến Linear Regression 1 cách tổng quát và trực quan. Phần 2 tập trung giải thích sâu hơn về việc tại sao lại sử dụng những công thức này và ý nghĩa của chúng.

- **Phần 1: Từ Logistic Regression đến Linear Regression đến Logistic Regression**
- **Phần 2: Tại Sao**

Phần 1: Từ Logistic Regression đến Linear Regression đến Logistic Regression

1.1 Phần 1.1: Tổng quan về Linear Regression

- giải thích thuật ngữ - giải thích công thức qua 1 ví dụ từ đầu đến cuối - giải thích ý nghĩa công thức
- code lại sử dụng ví dụ

1.2 Phần 1.2: Logistic Regression

- Liên kết với Linear Regression và vấn đề Logistic Regression giải quyết - giải thích thuật ngữ - giải thích công thức qua 1 ví dụ từ đầu đến cuối - giải thích ý nghĩa công thức - code lại sử dụng ví dụ

Phần 2: Tại sao

2.1 Tại sao lại sử dụng hàm tăng trưởng Sigmoid ?

2.2 Tại sao lại sử dụng $\ln()$ thay vì $\log()$

2.3 Mối liên hệ giữa $\ln()$ và Sigmoid ?

2.4 Tại sao lại sử dụng loglikelihood

2.5 Convex là gì ?

2.6 Convex trong Logistic Regression cho tính toán ma trận