ONE PAGE SUMMARY

IT & Software › Machine Learning › Feature Engineering

Introduction to Feature Engineering

Unlocking Data's Potential Through Feature Engineering

DEFINITION

> Feature engineering is the process of selecting, modifying, or creating features to improve model performance. It helps machine learning algorithms understand data better.

KEY CONCEPTS

Feature Selection

Choosing the most relevant features from the dataset to improve model accuracy.

Feature Creation

Generating new features from existing data to capture additional information.

Normalization

Scaling features to a common range to ensure equal contribution to model training.

Encoding Categorical Variables

Transforming categorical data into numerical format for better model interpretation.

Handling Missing Values

Applying techniques to manage gaps in data, ensuring models can learn effectively.

EXAMPLES

→ Converting 'date' into 'day', 'month', and 'year' features.

→ Using one-hot encoding for categorical variables like 'color'.

→ Scaling income data to a range of 0 to 1.

MEMORY TIPS

★ Remember 'F' for Feature and 'S' for Selection - FS helps focus on what matters.

★ Think of features as puzzle pieces; only the right pieces fit the picture.

★ Use the acronym 'CEN' for Categorical Encoding and Normalization.

COMMON MISTAKES

✗ Neglecting to preprocess data before model training.

✗ Overfitting by creating too many features without validation.

✗ Ignoring the impact of irrelevant features on model performance.

QUICK RECAP

Feature engineering is crucial for enhancing model performance by refining data inputs. Focus on selecting, creating, and processing features effectively to ensure your model learns accurately.