12304 12304 4189

2208 → Dev & Seema

10323 → Acc mai hona, change

14305

Machine Learning:

Unit 1:

Feature Engineering: a process of selecting, manipulating and transforming raw data into features that can be used in supervised learning.

Feature Engineering Process:

1. Feature creatn: finding most useful variables to be used in a predictive model.
2. Transformations: adjusting the predictor variable to improve the accuracy & performance of model.
3. Feature Extraction: generates new variables by extracting them from raw data.
4. Feature Selection: A way of selecting the subset of most relevant features from original features set by removing noisy features.

Feature Engineering Steps:

1. Data Prep: Getting raw data ready for ML by cleaning and organizing it.
2. Exploratory Analysis: Using visualizations to understand data patterns and select imp features.

Benchmark: setting a standard to measure and ~~imp~~ improve model accuracy.

Feature Engineering Techniques:

1. Imputation: Fill missing vals. with column's mean, median or mode.
2. Handling Outliers: Identify and remove outliers using standard deviation or z-score.
3. Log Transform: Make skewed data more normal by applying a log transformation.
4. Binning: Group continous data into bins to reduce noise.
5. Feature split: Divide features into parts to create new features.
6. One-Hot Encoding: Convert categorical data into a numerical format suitable for machine learning.

K-nn K-Nearest Neighbor

K = no. of neighbors used

its a lazy learning algo used for both classification and regression tasks.

Algo Steps:

1. Data Preparation:
   - collect dataset & label instances
   - Normalise dataset

2. Choose K

3. Prediction:
   1. Classification:
      1. compute distⁿ betⁿ new data point and all training data points.
      2. identify 'K'.
      - Assign the class labeled based on majority class among K neighbors.

   2. Regression:
      Same from above 1 and 2.
      Predict value by avg vals of K neighbors.

4. Model Evaluation:
1. Use accuracy | recall | precisio/ f1-score for Classification.
2. Use MSE | RMSE for Regression

Unit 4

Linear Regression
↳ predicts a continous dependent variable based on 1 or more independent variable
finds the best fitting straight line

$$y = mx + c$$

y = dependent var
x = independent var

Adv:
1. Simple and easy
2. Provides a clear relⁿ betⁿ var.
3. Useful for making predictions.

Limitations:
1. Assumes a linear relⁿ betⁿ var
2. Sensitive to outliers.

Multiple Linear Regression | Multivariabe Regression

↳ predicts the value of dependent variable based on the values of multiple independent variables

Equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \cdots + \varepsilon$$

$y \rightarrow$ dependent variable

$x_1, x_2, x_3 \ldots \rightarrow$ independent variable

$\beta_0 \rightarrow$ y - intercept

$\beta_1, \beta_2 \ldots \rightarrow$ coeff

$\varepsilon \rightarrow$ error term

Logistic Regression:

↳ 1. predicts the probability of an outcome that can only have 2 values.

2. used to solve classification problems.

3. o/p val is discrete

Ridge Regression:

↳ A type of linear regression that includes a regularization term to prevent overfitting by shrinking the coeff of features.

Lasso And Ridge Regression

Feature / Aspect	Lasso	Ridge
Regularization type	L1 regularizatn (absolute val. of coeff)	L2 regularizatn (square of coeff)
Coeff adjustment	can force some coeff to be exactly 0.	Shrinks coeff of less significant features towards 0.
feature selection	Performs both regularization and feature selection	Does not perform feature select-ion, only shrinks coeff values.
Suitability	more suitable for high-dimensional datasets	more suitable for datasets with highly correlated predictors.

Regression Metrics:

Mean Absolute Error (MAE): the avg of the absolute differences bet^n the actual values and predicted values.

Lower MAE = Better Accuracy

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

no. of observations — actual value — Predicted value

Mean Squared Error (MSE): avg of the squared differences bet^n the actual values and the predicted values.

Lower MSE = Better Fit

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

R-squared (R2) score: R^2 aka coefficient of determination, indicates how well the independent variables explain the variability of the dependent variables.

Higher R^2 = Better model performance

RMSE: square root of avg of squared differences bet^n the actual values and predicted values.

Lower RMSE = Better fit, easier to interpret.

Markov Model

↳ mathematical system that describes transitions from one state to another in a chain like process.

e2 depends on e1 and so on.

POS with HMM

Unit 5

Bayesian Belief Network.

Reinforcement learning : Completely diff

↳ A type of ML where an agent learns to make decisions by performing actions in an environment to achieve the best possible outcome typically through trial and error.

- Agent : learner
- Environment: world with which agent interacts
- Action: Choices agent can make
- State : current situation of agent within environment.
- Reward: feedback from environment based on action taken.

Working:

1. Agent Starts.
2. Action Selection
3. Environment Response
4. Learning

Policy, Reward function
Value function
Model of the env.

Transfer Learning:

↳ A ML technique where a model developed for one task is reused as the starting point for a model on a different but related task.

↑ Efficiency.
Improved performance.

Unit 3: Imp Only

K-means clustering:
↳ it divides a dataset into K clusters, where each data point belongs to the cluster with the neareast mean (centroid).

Working:
1. Initialization: Choose K inital centroids
2. Assign each data point to nearest centroid
3. Update: cal new centroids by taking mean of all data points in each cluster.
4. Iteration: Repeat the assignment and update until centroids no longer change.

Adv: Simple
efficient

Disadv: Predefined K
Sensitivity to initialization

Hierarchical Clustering:
↳ builds a hierarchy of clusters by either merging smaller clusters (agglomerative approach) or splitting larger clusters (divisive approach)

Agglomerative: merging into one big cluster

Divisive: breaking a single cluster into multiple clusters (small).

BIRCH: