

Most Asked Questions

Lecture date: 13-08-2025

Top 10 Technical Queries

1. Query: Why is the yellow line showing in the import function?

Answer: The yellow underline in code editors (like Jupyter or VS Code) usually indicates a warning, such as an unused import or a style suggestion. It does not stop the code from running.

2. Query: Why are we using `sns.load_dataset` instead of pandas?

Answer: `sns.load_dataset` is a quick way to load example datasets built into Seaborn, without needing to download CSV files manually. Pandas is used when loading datasets from your local system or an external URL.

3. Query: Can we use `pd` of pandas instead of `sns`?

Answer: Yes, you can use pandas (`pd.read_csv`) to load datasets instead of Seaborn. The choice depends on where your dataset is coming from.

4. Query: What does `.copy()` mean?

Answer: `.copy()` creates a duplicate of a DataFrame or Series to avoid modifying the original data unintentionally.

5. Query: What is `axis=1` in pandas?

Answer: In pandas, `axis=1` means the operation is applied across columns (horizontal), while `axis=0` applies across rows (vertical).

6. Query: Why are we splitting data into X and y?

Answer: In ML, `X` contains the features (inputs) and `y` contains the target variable (output we want to predict).

7. Query: Why did we drop some features?

Answer: Features are dropped when they are irrelevant, have too many missing values, or may cause overfitting. This improves model performance.

8. Query: What do 1 and 0 mean in the `survived` column?

Answer: 1 indicates the passenger survived; 0 indicates they did not.

9. Query: Can we change the value of `random_state` other than 42?

Answer: Yes, any integer can be used. Changing it affects how the data is split into train/test sets, which can slightly change results.

10. Query: What happens if we use 90% data for training and 10% for testing?

Answer: The model will have more data to learn from, but less data to test on. Accuracy might vary; sometimes it increases, but you risk less reliable test evaluation.

Top 10 Non-Technical Queries

1.

Query: Can we get the link of the CSV file?

Answer: The Titanic CSV file is available on LMS or can be loaded via Seaborn's

`sns.load_dataset('titanic')`.

2.

Query: Can we get the dataset?

Answer: Yes, datasets are provided in the LMS resources section or as part of the code examples.

3. Query: Can we get the theory and some practice materials for all of these libraries?

Answer: Yes, materials are available in the LMS and GitHub repository shared by the instructor.

4. Query: For sklearn, do we need to install any library first?

Answer: Yes, install it using `pip install scikit-learn`.

5. Query: What can I do to rectify the error "No module named 'sklearn'"?

Answer: Install the library with `pip install scikit-learn` in your environment.