Most Asked Questions

Lecture date: 18-08-2025

Top 10 Technical Queries

Query 1: What is the difference between TensorFlow and Scikit-learn?
Answer: TensorFlow is mainly for deep learning and neural networks, while Scikit-learn is best for traditional ML algorithms (linear regression, SVM, clustering, etc.).

Query 2: What does it mean to fit data points to a curve?
Answer: Fitting data means finding a mathematical function that best represents the relationship between input and output. If the curve is linear, it's linear regression; otherwise, it's non-linear regression.

Query 3: What are parameters in a model (like intercept)?
Answer: Parameters are values learned by the model from data. For linear regression, these include slope(s) (coefficients) and the intercept.

Query 4: Are most real-world data patterns linear or non-linear?
Answer: Most real-world data is non-linear. Linear models are a starting point but often too limited. ML algorithms like decision trees, random forests, and neural networks capture non-linear patterns.

Query 5: Can some models only capture linear patterns?
Answer: Yes. Linear regression, Logistic regression, Ridge/Lasso, and Linear SVM assume linear relationships. To capture non-linear patterns, feature transformations or different models are needed.

Query 6: How do outliers affect machine learning models?
Answer: Outliers can heavily influence models like linear regression, pulling the line or boundary and making predictions inaccurate. Robust methods or outlier treatment are needed.

Query 7: If I remove outliers, can new ones appear?
Answer: Yes. Handling outliers is iterative — once some are removed, new ones may appear after retraining. Robust models reduce this problem.

Query 8: What is `random_state` in ML code?
Answer: It's a seed value for randomness (used in train-test split, shuffling, weight initialization). Setting it ensures reproducible results.

Query 9: What does `drop_first=True` mean in `pandas.get_dummies()`?
Answer: It drops the first category to avoid multicollinearity in models like regression. Remaining dummy columns still represent all categories.

Query 10: Does linear regression always have only 2 parameters?
Answer: No. The number of parameters depends on the number of input features. For n features, there are n coefficients + 1 intercept.

Top 10 Non-Technical Queries

Query 1: Where can I raise doubts if not resolved in class?
Answer: You can raise a support ticket on LMS. If you have any further doubts or need clarification, we encourage you to attend the upcoming tutorial session with your Teaching Assistant for additional support, and the team will get back to you.

Query 2: Will class notes/recordings be shared?
Answer: Yes. Notes and recordings are shared after class. If delayed, they'll be provided before the next session.

Query 3: Can we get the theory + practice material for the libraries used?
Answer: Yes. Materials will be shared, and official documentation links are recommended (e.g., MDN, Scikit-learn docs).

Query 4: What if I miss a class?
Answer: You can refer to the recording and shared notes.

Query 5: Do attendance or assignments affect CGPA?
Answer: No, Attendance and assignment submissions are not included in the calculation of the overall score for grading purposes. Only evaluation scores count toward CGPA

Query 6: Will we get the program curriculum?
Answer: Yes. The module-wise curriculum is shared via email and available on the official website.

Query 7: How are evaluations scheduled?
Answer: Mid-semester after Module 2 (25% weightage) and after Module 4 (25% weightage), with final evaluation carrying 50% (per FAQ).