Roll No.												

PRESIDENCY UNIVERSITY

BENGALURU

Mid - Term Examinations – October 2025

Date: 27-10-2025 **Time:** 11.00am to 12.30pm

School: SOE/SOCSE	**Program:** B.Tech [CSE(AIML)]	
Course Code: CAI3405	**Course Name:** Explainable AI	
Semester: VII	**Max Marks**: 50	**Weightage**: 25%

CO - Levels	**CO1**	**CO2**	**CO3**	**CO4**	**CO5**
Marks	**26**	**24**	**-**	**-**	**-**

Instructions:

(i) Read all questions carefully and answer accordingly.

(ii) Do not write anything on the question paper other than roll number.

Part A

Answer ALL the Questions. Each question carries 2marks. **5Q x 2M=10M**

1	What is the main motivation behind Explainable AI (XAI)?	**2 Marks**	**L1**	**CO1**
2	Define interpretability in the context of AI systems.	**2 Marks**	**L1**	**CO1**
3	What is the trade-off between accuracy and interpretability in AI models?	**2 Marks**	**L1**	**CO1**
4	What are transparent models? Give one example.	**2 Marks**	**L1**	**CO2**
5	Define Permutation Feature Importance (PFI).	**2 Marks**	**L1**	**CO2**

Part B

Answer the Questions. **Total Marks 40M**

6.	**a.**	A company uses an AI system to shortlist candidates for interviews. Some candidates question why they were rejected. Explain how trust and accountability in AI can be improved in such a system.	**5 Marks**	**L2**	**CO 1**
	b.	You have a black-box model with high accuracy. To make it understandable, you decide to build a surrogate model. Explain how surrogate models provide global explanations and why they are useful.	**5 Marks**	**L2**	**CO 2**
		Or			
7.	**a.**	A healthcare AI model gives accurate predictions but is too complex for doctors to interpret. Discuss the trade-off between accuracy and interpretability in this case, and suggest how to balance both.	**5 Marks**	**L2**	**CO 1**
	b.	You are training a regression model and want to understand how each input feature affects predictions globally. Explain how Partial Dependence Plots (PDP) can help visualize these relationships.	**5 Marks**	**L2**	**CO 2**

8.	**a.**	During model development, the data scientist decides to include explainability checkpoints at each phase. Explain how explainability can be integrated throughout the model development lifecycle.	**7 Marks**	**L2**	**CO 1**
	b.	Given a Linear Regression model: $y=-2+0.4\,X1+0.03\,X2$ If a student studied 5 hours and had 80% attendance, calculate the predicted score y. Explain how this simple transparent model is interpretable compared to a neural network.	**8 Marks**	**L3**	**CO 2**
		Or			
9.	**a.**	Imagine a smart traffic management system powered by AI. Discuss why transparency and interpretability are critical for public trust in such a system.	**8 Marks**	**L2**	**CO 1**
	b.	An e-commerce platform uses SHAP to understand its recommendation model. Explain how SHAP values are used to assign importance to features and why they enhance explainability.	**7 Marks**	**L3**	**CO 2**

10.	**a.**	Different stakeholders (developers, regulators, end-users) expect different levels of explanation from an AI model. Explain how audience segmentation affects the design of	**8 Marks**	**L2**	**CO 1**

		explanations in XAI.			
	b.	You are working on a complex financial prediction model. Explain how a RuleFit model can be used to extract human-understandable decision rules while maintaining model accuracy.	**7 Marks**	**L2**	**CO 2**
		Or			
11.	**a.**	An autonomous vehicle company wants to build public trust before deployment. Describe how XAI principles (transparency, interpretability, trust, accountability) can be applied to achieve this goal.	**7 Marks**	**L2**	**CO 1**
	b.	A healthcare model uses multiple explainability methods—PDP, LIME, and SHAP. Compare these three techniques in terms of *scope of explanation (local vs global)* and *ease of interpretation*.	**8 Marks**	**L2**	**CO 2**