

XAI Question Bank

Unit 4,5,6

1. Explain different XAI techniques.
2. Explain the mathematical foundation of SHAP and explain how SHAP approximates complex models.
3. Discuss a comparative analysis of the XAI methods covered (DiCE, and LRP, Lime and SHAP).
4. Describe the Shapley Additive exPlanations (SHAP) method.
5. Explain how does LRP work to backpropagate relevance scores.
6. Explain how LIME provide local explanations for complex machine learning models along with its mathematical representation.
7. Describe Layer-wise Relevance Propagation (LRP) with its application.
8. Explain Local Interpretable Model-Agnostic Explanations (LIME) in Explainable AI.
9. Discuss the role of counterfactual explanations in XAI.
10. Explain Diverse Counterfactual Explanations (DiCE) in AI interpretability.
11. Explain the advantages and limitations of LIME and SHAP as model-agnostic explanation methods.
12. Discuss the key metrics to evaluate XAI.
13. Explain the concept of Power Quality Disturbance (PQD) classification.
14. Explain any three methods for measuring human intelligence.

Unit 5

1. Explain the importance of pre-model explainability in the context of unstructured data.
2. Explain the preprocessing steps involved in making textual data interpretable for machine learning models.
3. Describe how LIME is used to explain text data classification.
4. How can techniques like tokenization, stopwords removal, and word embeddings contribute to pre-model explainability? Illustrate with an example.
5. Explain the role of feature extraction and engineering in unstructured data.
6. Discuss the main challenges in applying explainability methods to text data.
7. Illustrate how XRAI improves upon Grad-CAM for visual explanations.
8. Compare LIME and Shape for text classifiers in terms of effectiveness and interpretability.
9. Explain how Layer-wise Relevance Propagation (LRP) works with a deep neural network model for image classification.
10. Compare and contrast explainability approaches for text data versus image data.
11. Explain how LIME and Shape work together for explainability in text classification models.
12. Discuss the use of supervised wrappers for clustering models in the context of unstructured data explainability.
13. Explore the role of feature selection in ensuring explainability for both text and image classifiers.

14. Discuss the explainability techniques applicable to time series models, including Grad-CAM and LRP.
15. Explain the challenges and methods of using foundation models like LLMs for explainability in unstructured data.

Unit 6

1. Discuss two recent trends in Explainable AI that are focused on building trust in models.
2. What are the key factors in making AI models trustworthy in medical diagnosis?
3. Describe how XAI can improve the trustworthiness of sales predictions in real estate.
4. Why is explainability important for physicians when using AI in medical decision-making?
5. Describe how recent trends in XAI, such as counterfactual explanations or model transparency, are being used in the medical domain.
6. How does explainability improve model performance and trustworthiness?
7. Discuss how explainable AI can assist physicians in making critical medical decisions.
8. Evaluate how the integration of explainability in AI models can influence decision-making in both healthcare and real estate.
9. Discuss the recent trends in Explainable AI that are focused on building trustworthy AI models.
10. Discuss the importance of XAI in medical diagnosis.
11. What are the challenges in building trustworthy sales prediction models in real estate?
12. Describe XRAI (eXtended Region-based Activation Importance)?
13. How can explainable AI help build trust among patients when AI is used in medical decision-making?
14. Explain the challenges and solutions for making AI-based sales prediction models more trustworthy.
15. Explore recent trends in Explainable AI (XAI).