# CSE422 Viva Questions

- How many classes are you detecting in your project?

- How many features are there in your project?

- How many data points are in your dataset?

- What type of dataset are you using - balanced or imbalanced?

- Explain why it is balanced or imbalanced.

- What is LabelEncoder?

- What is the purpose of using LabelEncoder?

- Why do we convert categorical values to numerical values?

- What is correlation, and why is it important?

- Is high correlation good or bad? Why?

- What is preprocessing in machine learning?

- How do you handle rows with null values?

- Why is scaling necessary?

- Which scaling technique did you use, and how does it work?

- What type of data splitting technique did you use? Why?

- If you used stratified splitting, how does it work?

- If you used random splitting, how does it work?

- What is random state, and why do we use it?

- Which models did you use, and how do they work?

- What are precision and recall, and how do they differ?

- What is a confusion matrix?

- What are False Positive, False Negative, True Positive, and True Negative?

- Show False Positive, False Negative, True Positive, and True Negative?