Typical exam

- Discuss differences between bagging and boosting
- Discuss how a random forest can be used to estimate the similarity between two instances
- Discuss a hierarchical clustering algorithm
- Discuss the learning process of a binary classification tree
- Discuss the Information Gain splitting criterion
- Discuss how to handle categorical data

Typical exam

- Find the first node of a decision tree predicting the class "PlayTennis" by using the GINI criterion on the following dataset

TRAINING				PlayTennis
Outlook	**Temperature**	**Humidity**	**Wind**	**PlayTennis**
Sunny	Hot	High	Weak	**No**
Sunny	Hot	High	Strong	**No**
Overcast	Hot	High	Weak	**Yes**
Rain	Mild	High	Weak	**Yes**
Rain	Cool	Normal	Weak	**Yes**
Rain	Cool	Normal	Strong	**No**
Overcast	Cool	Normal	Strong	**Yes**
Sunny	Mild	High	Weak	**No**
Sunny	Cool	Normal	Weak	**Yes**
Rain	Mild	Normal	Weak	**Yes**
Sunny	Mild	Normal	Strong	**Yes**
Overcast	Mild	High	Strong	**Yes**