Q1. Ensemble techniques in machine learning involve combining multiple individual models to improve predictive performance or robustness.

Q2. Ensemble techniques are used in machine learning for several reasons:

- They can reduce overfitting by averaging out biases from individual models.
- They can increase predictive accuracy by leveraging the strengths of different models.
- They provide better generalization to new data.
- They are more robust to noisy data.
-

Q3. Bagging, or Bootstrap Aggregating, is a technique in ensemble learning where multiple models are trained on different subsets of the training data, sampled with replacement. The final prediction is typically the average of predictions made by individual models.

Q4. Boosting is another ensemble technique where models are trained sequentially, with each model focusing on the instances that previous models have misclassified. Boosting aims to reduce bias and variance by iteratively improving the model's performance.

Q5. The benefits of using ensemble techniques include:

- Improved predictive performance.
- Reduced overfitting.
- Increased robustness to noise.
- Better generalization to unseen data.
- Ability to handle complex relationships in data.
-

Q6. Ensemble techniques are not always better than individual models. Their effectiveness depends on factors such as the quality and diversity of the base models, the nature of the data, and the ensemble method used.

Q7. The confidence interval using bootstrap is calculated by resampling the dataset with replacement, generating multiple bootstrap samples, computing the statistic of interest (e.g., mean, median) for each sample, and then calculating the confidence interval based on the distribution of these statistics.

Q8. Bootstrap is a resampling technique used for estimating the sampling distribution of a statistic by repeatedly resampling from the observed data. The steps involved in bootstrap are:

1. Randomly draw a sample (with replacement) from the observed data.
2. Compute the statistic of interest (e.g., mean, median) for the resampled data.
3. Repeat steps 1 and 2 a large number of times (e.g., thousands of times).
4. Construct the bootstrap sampling distribution and estimate confidence intervals or other properties of interest.

Q9. To estimate the 95% confidence interval for the population mean height using bootstrap:

- Randomly sample 50 heights (with replacement) from the observed sample of heights.
- Calculate the mean height for each bootstrap sample.
- Repeat this process a large number of times (e.g., 10,000 times).
- Compute the 2.5th and 97.5th percentiles of the bootstrap sample means.
- This interval (from the 2.5th to the 97.5th percentile) will provide the 95% confidence interval for the population mean height.