

## Module 5: Assessment Assignment

1. In the YouTube video "Machine Learning with Python - Part 2: Decision Tree," why did Wes Doyle not change the min\_samples\_split to 2 on the first tree he built?

Ans: Because a minimum of 100 samples are needed to split one node, Wes Doyle does not adjust the min\_samples\_split setting on the first tree he constructs in "Machine Learning with Python: Part 2: Decision Tree" to 2. A decision tree that is highly particular to the data set that is trained upon is at a very high danger of being overfit because we have one observation on either side of a split, and we would have this dense tree with, in this example, perhaps hundreds of leaves.

Refer to this video at [www.youtube.com/watch?v=XDbi6PxaSf0](https://www.youtube.com/watch?v=XDbi6PxaSf0)

2. In the YouTube video "Cloud Provider Comparisons: AWS vs Azure vs GCP - AI and Machine Learning," Scott identified nine building block services in the video?

Ans: The nine foundational services identified by Scott in the YouTube video "Cloud Provider Comparisons: AWS vs. Azure vs. GCP: AI and Machine Learning" are

Speech-to-text

Text-to-speech

Translation

Chatbots

Speech-to-text

Text Analytics

Document Analysis

Image/Video Analysis

Anomaly Detection

Personalization

Reference: <https://www.youtube.com/watch?v=PNnSHPn×2Xc>

3. After watching the YouTube video "Cloud Provider Comparisons: AWS vs Azure vs GCP - AI and Machine," Which of the three cloud providers would you recommend to your friend who is starting a new customer service company specializing in chatbots and translation services? Explain why you select the cloud provider.

Ans: I would suggest AWS, Microsoft Azure, and Google Cloud Platform to my friend who is beginning a new customer care business that specializes in chatbots and translation services.

- Azure could be incorporated into Responsible ML and the Fairness SDK.

- AWS offers Sage maker Clarify, which can aid with insight into how data elements influence the model generation process and evaluation fairness.
- GCP develops services under AI Explanations and offers Sage maker Clarify.

These cloud service providers offer a variety of services that help businesses grow, cut costs, increase privacy, and concentrate on their main objectives. These advantages have increased the reliance of the contemporary digital environment on cloud service providers.

<https://www.youtube.com/watch?v=PNnSHPNx2Xc> is a good source.

4. In the article, "DoS attack detection method based on improved KNN with the degree of DDoS attack in software-defined networks," what were the two methods to detect the DDoS attack in SDN?

Ans: There are two ways to identify DoS assaults in software-defined networks, according to the article DDoS attack detection method based on improved KNN with the degree of DDoS attack in SD. One method bases the identification of the DDoS attack on the intensity of the attack. The second method finds DDoS attacks using an upgraded K-Nearest Neighbours (KNN) technique built on Machine Learning (ML).

See <https://ieeexplore.ieee.org/abstract/document/8945375> for more information

5. In the article, "DDoS attack detection method based on improved KNN with the degree of DDoS attack in software-defined networks." how did the authors improve the KN algorithm? Do you think this was necessary? Explain why or why not.

Ans: According to the study "DDoS attack detection method based on improved KNN with the degree of DDoS attack in software-defined networks," The weight parameter ( $w$ ) was added by the authors to the KN technique to address situations where conflicting labels among nearest neighbours led to inaccurate findings. They suggested using the distance between the detected item ( $x$ ) and other flows ( $i$ ) to compute the weight as  $1/d(xp, xi)$ . By expressing the weight as  $1/e^t$  where  $t = d(xp, Xi)$ , they significantly enhanced it.

This enhancement, in my opinion, was crucial because the original KNN algorithm produced ambiguous results when numerous neighbours had different tags. The weight parameter has been added to provide a more fair and accurate impact of nearby flows on the final result. The authors sought to enhance the effectiveness of DDoS attack detection using machine learning techniques by resolving this problem.

See <https://ieeexplore.ieee.org/abstract/document/8945375> for more information.