Câu 1: Học có giám sát (Supervised Learning) là phương pháp học máy trong đó mô hình được huấn luyện với tập dữ liệu có nhãn. Điều này có nghĩa là mỗi đầu vào (input) sẽ có một đầu ra mong muốn (output) đã biết trước. Ví dụ: phân loại email thành spam hoặc không spam.  
  
  
Câu 2: Confusion Matrix là một bảng gồm 4 ô dùng để đánh giá hiệu suất của mô hình phân loại:  
True Positive (TP): Dự đoán đúng mẫu thuộc lớp dương.  
False Positive (FP): Dự đoán sai mẫu không thuộc lớp dương thành lớp dương.  
True Negative (TN): Dự đoán đúng mẫu thuộc lớp âm.  
  
  
Câu 3: Overfitting là hiện tượng mô hình học máy học quá kỹ trên tập huấn luyện đến mức nó không thể tổng quát hóa tốt trên dữ liệu mới.