**XGBoost (eXtreme Gradient Boosting)** là một thuật toán học máy thuộc họ Gradient Boosting, được phát triển để cải thiện và mở rộng các thuật toán truyền thống của Gradient Boosting. Dưới đây là một giải thích chi tiết về XGBoost:

1. **Gradient Boosting:**
   * Gradient Boosting là một phương pháp tập hợp cây quyết định, trong đó mỗi cây được xây dựng để sửa các lỗi của cây trước đó.
   * Các cây được xây dựng tuần tự, mỗi cây hướng dẫn dự đoán sự chênh lệch giữa giá trị thực tế và giá trị dự đoán của các cây trước đó.
2. **Regularization trong XGBoost:**
   * XGBoost thêm các thành phần regularization để kiểm soát sự phức tạp của mô hình và ngăn chặn overfitting.
   * Hàm mất mát của XGBoost bao gồm ba thành phần: mất mát dự đoán, mất mát tổng quát và hàm chống chệch.
3. **Boosting và Học Bậc Cao:**
   * XGBoost có khả năng hỗ trợ việc học bậc cao (higher-order learning), cho phép mô hình học các mối quan hệ phức tạp giữa các biến đầu vào.
4. **Cách Xây Dựng Cây:**
   * Các cây trong XGBoost được xây dựng bằng cách sử dụng thuật toán tối ưu hóa để chọn nút chia tốt nhất và tối ưu hóa hàm mất mát.
5. **Thực Hiện Phương Pháp "Cây Thụt Lề" (Tree Pruning):**
   * Khi xây dựng mỗi cây, XGBoost thực hiện cây thụt lề để loại bỏ các nhánh không đóng góp nhiều vào việc cải thiện mất mát.
6. **Ứng Dụng Rộng Rãi:**
   * XGBoost có thể được sử dụng cho cả bài toán phân loại và hồi quy và thường được chọn cho các cuộc thi và dự án thực tế với hiệu suất cao.

**Ưu Điểm của XGBoost:**

* **Hiệu Suất Cao:** XGBoost thường có hiệu suất cao và là một trong những thuật toán hàng đầu trong nhiều cuộc thi máy học.
* **Chống Chệch và Regularization:** Các thành phần chống chệch và regularization giúp kiểm soát overfitting.
* **Xử lý Hiệu Quả Dữ Liệu Lớn:** XGBoost có thể xử lý hiệu quả dữ liệu lớn và có các tối ưu hóa tăng tốc để cải thiện tốc độ huấn luyện.

**Nhược Điểm của XGBoost:**

* **Cần Điều Chỉnh Tham Số:** XGBoost có nhiều tham số cần được điều chỉnh, và việc chọn các giá trị thích hợp có thể đòi hỏi kiến thức và thời gian.