Các bước xác định được hệ số α:

Bước 1: Tạo cây đầy đủ (full tree) trên tập train (tập đầy đủ)

Bước 2: Từ cây đầy đủ (α = 0), ta lần lượt chặt từng nhánh, từ thấp tới cao, mỗi lần chặt sẽ ứng với mỗi model cần xét (ví dụ ta chặt nhánh 4 lần)

Bước 3: ta có công thức tính tree score:

Tree score = SSR + α.T

SSR: Sum of error

T: Số none lá

Từ công thức tính trên, ứng với mỗi model ta có SSR và T. Việc còn lại của ta là chon α sao cho ứng với mỗi model thì Tree score đó sẽ nhỏ nhất. Suy ra ta sẽ có 5 α (tính cả α=0)

Bước 4: Chia tập data train thành K phần. sau đó lần lượt ta sử dụng K – 1 phần thành tập data\_train, 1 phần còn lại thành tập data\_test. Thường ta sẽ sử dụng là K = 5

Sau đó, ứng với mỗi tập datatrain, ta sẽ xây dựng cây quyết định lần lượt với 5 hệ số α ta đã có lúc nãy. Mỗi model ứng với mỗi α ta SSR đối với tập data\_test.

Suy ra, mỗi α sẽ có K model với K SSR khác nhau. Ta tính trung bình SSR của mỗi α, lấy α nào có trung bình SSR thấp thất (đồng thời chọn cây quyết định tạo ra α để làm model chuẩn).