

Huấn luyện mô hình Tree Model quyết định CPU Frequency

1. Dữ liệu huấn luyện:

Tham khảo từ bài báo "*An Interpretable Machine Learning Model Enhanced Integrated CPU-GPU DVFS Governor*", các đặc trưng của hệ thống được xem xét để đánh giá tần số của CPU bao gồm:

- CPU Utilization:

Mức sử dụng của từng core hoặc trung bình từng hệ thống. Lấy từ /proc/stat - file thống kê toàn bộ về hoạt động của hệ thống bao gồm CPU, tiến trình, ngắt v.v...

Ở đây ta sẽ quan tâm các thống kê về CPU. Trường dữ liệu của CPU bao gồm, thời gian user, nice, system, idle, iowait, irq, softirq, steal, guest, guest_nice. Để tính utilization, ta quan tâm đến các thời gian idle và non-idle; trong đó, các công thức được biểu diễn như sau

$$total_time = user + nice + system + idle + iowait + irq + softirq + steal$$

$$idle = idle + iowait$$

$$non_idle = user + nice + system + irq + softirq + steal$$

Do các thông số này được chồng chất trong hệ thống do đó mỗi lần Sample cần trả vào dữ liệu ở 2 thời điểm khác nhau và tính dựa vào chênh lệch.

$$CPU\ Utilization = \frac{delta_total}{(delta_total - delta_idle)}$$

- Load Average:

Tải trung bình trong 1/5/15 phút. Được hiển thị trong /proc/loadavg.

- CPU Temperature (mili Celcius)

Lấy từ /sys/class/thermal/thermal_zone0/temp.

- Tần số hiện tại của CPU,

Lấy trong /sys/devices/system/cpu/cpu0/cpufreq/scaling_cur_freq.

- Số lượng Run Queue Length.

Số lượng process đang chờ CPU. Lấy từ /proc/loadavg.

2. Kịch bản thu thập dữ liệu:

3. Cách lấy Features dưới Kernel Space: