

IMPLEMENTAÇÃO DE ESCALONADORES TUTORES PARA APRENDIZADO SUPERVISIONADO¹

Guilherme Diel¹², Guilherme Piêgas Koslovski¹³.

¹ Vinculado ao projeto “Mecanismos para Alocação de Infraestruturas Virtuais baseados em Aprendizado de Máquina e Acelerados por GPUs”

² Acadêmico do Curso de Ciência da Computação – CCT – Bolsista PIBIC/CNPq

³ Orientador, Departamento de Ciência da Computação – CCT – guilherme.koslovski@udesc.br

É fato que o *High Performance Computing* (HPC) é muito utilizado em diversas áreas de pesquisa, desde previsões do tempo até descobertas de novos medicamentos. Desse modo, urge a necessidade de aprimorar cada vez mais os métodos de escalonamento e, para obter esse resultado, foi pensado em adicionar Inteligência Artificial (IA) para fazer a árdua tarefa de otimizar métodos de escalonamento que já são utilizados a muito tempo, visto que essas conseguem tomar decisões baseadas em diversos fatores e dados relevantes, levando a um resultado mais eficiente. Nesse contexto, esta pesquisa está relacionada à implementação de políticas de escalonamento para serem utilizadas como tutores de aprendizado de máquina supervisionadas. Dessa forma, o aprendizado supervisionado é um paradigma de aprendizado de máquina em que algoritmos são treinados para aprender a relação entre entradas e saídas. Assim, neste estudo, as entradas foram os resultados de escalonamentos realizados com 7 políticas de escalonamento reconhecidas pela comunidade, com o intuito de fazer com que a máquina consiga tomar decisões de escalonamento que sejam, no mínimo, tão boas, ou até melhores que os métodos de escalonamento originais.

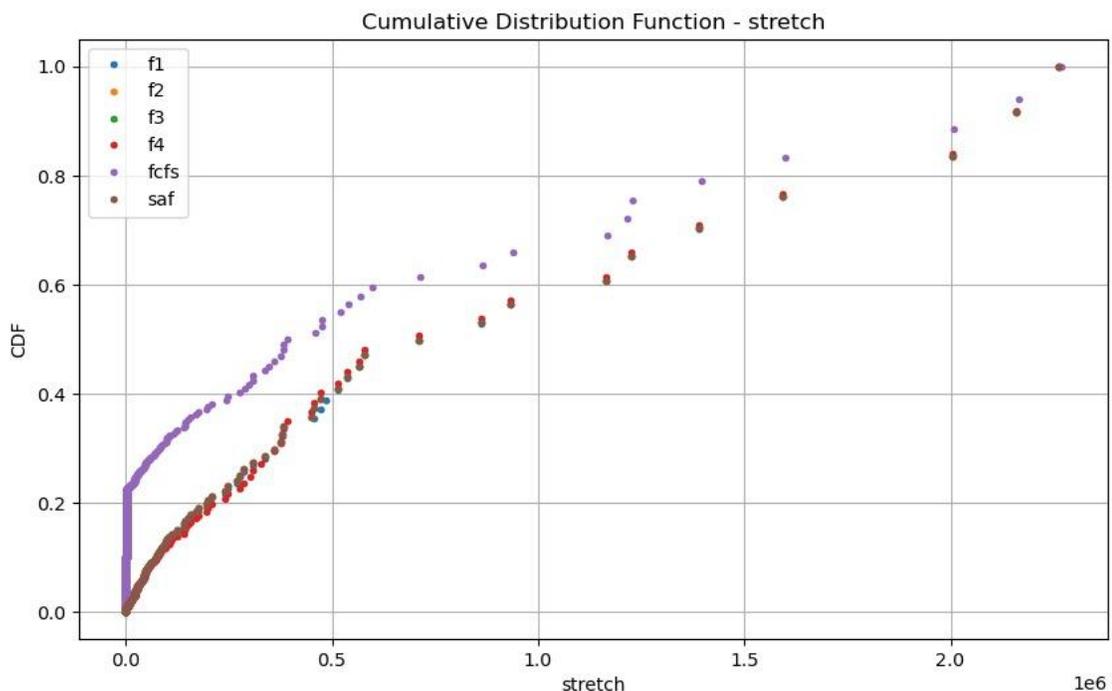
Dessa maneira, o primeiro passo foi escolher políticas de escalonamento que fossem bons tutores para a IA, de modo que sejam reconhecidos pela comunidade e sejam eficientes. Assim sendo, foram selecionados seis métodos: F1, F2, F3, F4, SAF, FCFS, sendo que os métodos F1, F2, F3 e F4 são baseados em fórmulas que levam em conta diversos atributos dos processos (*jobs*) a serem escalonados; o SAF dá prioridade ao *job* que possui a menor área - que é o produto entre a quantidade de recursos utilizados pelo processo, multiplicado pela quantidade de tempo estimado para completar a tarefa -, no FCFS as tarefas são postas na fila de processamento em ordem cronológica. Feito isso, o próximo passo foi simular escalonamentos utilizando as bases de dados HPC2N e SDSC-SP2, reconhecidas pela comunidade essas políticas e utilizar os resultados para passar como métricas à IA. Entretanto, aqui está o próximo desafio: escolher quais métricas passar para a máquina, em outras palavras, escolher quais são os objetivos da IA -minimizar o tempo em fila de cada tarefa, o gasto de energia, a maior parte dos atributos que tornam a fila ineficiente etc. Assim, foram escolhidos quatro objetivos: minimizar a média de energia gasta, a média de tempo em fila, a média dos atributos negativos e a diferença entre o tempo de fila estimado e o real.

À medida que o progresso se consolidou, o último passo foi escolher o tipo de aprendizado, visto que existem diversos tipos, com os quais são possíveis chegar a resultados parecidos: Supervisionado (AS), Profundo (Deep Learning), Não Supervisionado etc. Assim, visto que o Aprendizado Supervisionado foi o que fez mais sentido com a proposta da pesquisa, este foi escolhido. Além disso, dentre os vários tipos de AS, foi escolhido o por meio de classificação pelo método de GNB, visto que este apresentou os resultados mais consistentes. Em última análise, o desenvolvimento da IA trouxe à tona um desafio final e fascinante: algumas métricas necessárias

para prever os resultados do escalonamento só seriam obtidas após a conclusão do escalonamento em si. Para contornar essa complexidade, uma solução inovadora emergiu: a integração de métodos de regressão, que consigam fazer uma predição dos dados ausentes, possibilitando que a IA consiga fazer a otimização do escalonamento. Esse desdobramento da pesquisa, repleto de conquistas e inovações, destaca não apenas a relevância do aprendizado supervisionado em um contexto de HPC, mas também como o casamento inteligente entre técnicas consagradas e conceitos visionários pode conduzir a soluções práticas e significativas para desafios complexos.

Para melhor visualizar os escalonadores tutores, foi feito um gráfico de tipo *Cumulative Distribution Function* (CDF - Figura 1), que no eixo X representa o *stretch* (tempo que a tarefa ficou no sistema dividido pelo tempo de execução) de cada tarefa e no eixo Y representa a porcentagem de tarefas (0.0 até 1). Foi escolhido essa medida para o eixo X, visto que ele aponta qual escalonador tem um aproveitamento melhor do sistema, de modo que quando o *stretch* é maior que 1 indica que as tarefas estão demorando mais do que o esperado para serem concluídas, e isso pode levar a atrasos e ineficiências no sistema; *stretch* igual a 1 significa que não houveram atrasos, enquanto que quando este é menor que 1, mostra que a tarefa foi concluída mais rapidamente do que o seu tempo de execução real. Isso é geralmente uma situação desejável em computação de alto desempenho e em sistemas de escalonamento, pois significa que as tarefas estão sendo processadas de forma eficiente e com baixo tempo de espera. O gráfico aponta que FCFS é o escalonado com melhor aproveitamento do sistema, visto que mais de 60% (demonstrado no eixo Y) estão abaixo de 1 no eixo X enquanto nos outros escalonadores, menos de 60% das tarefas estão abaixo de 1.

Figura 1. Gráfico que mostra o *stretch*.



Palavras-chave: Escalonamento. Aprendizado Supervisionado. Escalonadores Tutores.