

Uso de aprendizado de Máquina Automatizado para Seleção de Provedores de Nuvem¹

Kauã Bernardo Hopfer², Adriano Fiorese³

¹ Vinculado ao projeto “Seleção de Provedores de Nuvem Computacional”

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

³ Orientador, Departamento de Ciência da Computação – CCT – adriano.fiorese@udesc.br

A computação em nuvem é um tipo de serviço que possibilita o acesso a recursos computacionais como armazenamento de dados, hospedagem de sistemas, criação, execução e distribuição de software pela Internet. Dessa forma, com o uso da nuvem o usuário não precisa se preocupar com a contratação ou compra de servidores e hardwares específicos para dispor desses recursos, atentando-se apenas à escolha do provedor de serviço de nuvem ideal.

Entretanto, dado o crescimento da demanda para serviços de computação de nuvem, veio o aumento da oferta. Com essa questão, é introduzido o problema que o trabalho de IC busca resolver, ou seja, qual provedor de nuvem é o mais adequado para um indivíduo ou organização.

Assim, o objetivo do trabalho consiste na automatização desse processo de escolha com o emprego de um sistema de ranqueamento de provedores de serviços de nuvem. Esse sistema de ranqueamento deve produzir uma pontuação para cada provedor avaliado a partir da qualificação dos requisitos exigidos pelo usuário. Então, o provedor que possuir a maior pontuação final será teoricamente o provedor de nuvem mais apropriado para dado usuário/organização, conforme os requisitos exigidos.

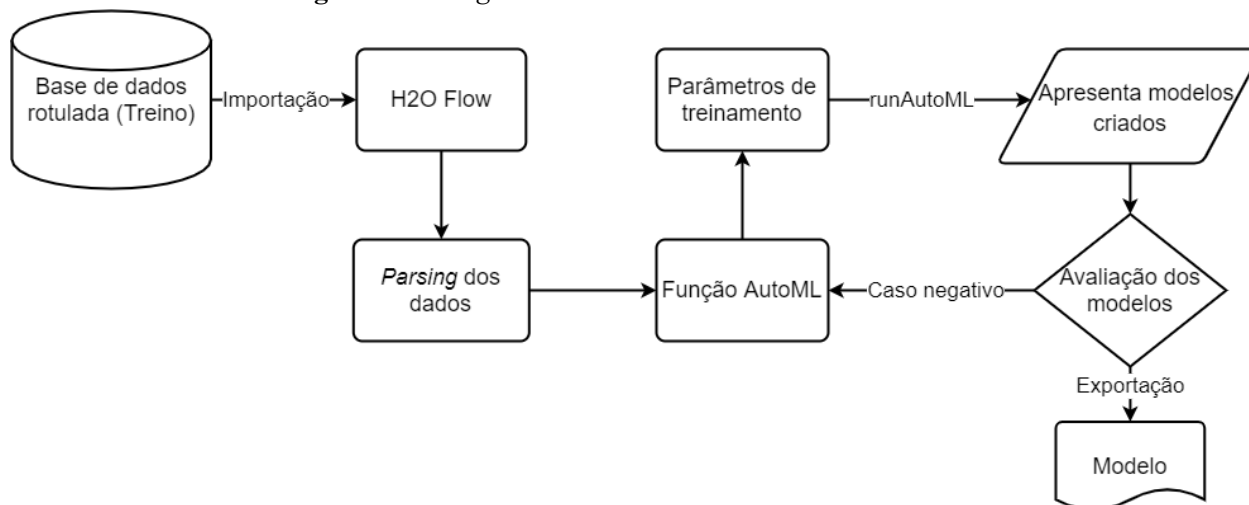
Para atingir esse objetivo, foi utilizada a abordagem ao problema com a aplicação de modelos de regressão por meio de Aprendizado de Máquina Automatizado, do inglês, *Automated Machine Learning* (AutoML), obtidos com o uso da plataforma de aprendizado de máquina H2O. Tais modelos baseiam-se no treinamento supervisionado embasado nos resultados obtidos pela solução do algoritmo de Matching proposto no trabalho [de Moraes et al. 2017].

Com o uso da plataforma AutoML H2O, é possível o treinamento, análise e criação de modelos AutoML que após treinados, são capazes de realizar previsões de pontuações baseadas implicitamente na lógica do algoritmo de Matching, de forma mais rápida e simples para a obtenção dos resultados. Além disso, com o emprego da interface gráfica baseada em web H2O Flow, é possível a visualização e análise dos resultados de forma intuitiva devido a sua capacidade gráfica e organização de dados.

O processo de treinamento pode ser ilustrado pela Figura 1. Nela está o fluxograma de treinamento do modelo AutoML. Assim, o processo de treinamento é iniciado com a importação da base de dados rotulada para o H2O flow, para em seguida realizar o *parsing* dos dados. Em seguida, no ciclo de treinamento, definimos os parâmetros de treinamento, como tempo limite de execução e exclusão de algoritmos a serem testados, e então é executada a função `runAutoML`. Após o fim da execução, o treinamento também é finalizado. Dessa forma, a função apresenta os modelos criados. Caso um modelo satisfatório seja encontrado, pode-se realizar a exportação do mesmo, salvando o modelo para ser utilizado a qualquer momento. Caso contrário, retorna-se ao

ciclo de treinamento, para que os parâmetros possam ser alterados em busca de resultados diferentes.

Figura 1. Fluxograma de desenvolvimento com H2O Flow



Com todos os passos concluídos, o usuário pode realizar previsões acerca de outras bases de dados de indicadores de provedores de nuvem, por meio do ambiente H2O Flow, com o uso do modelo previamente treinado.

Os resultados obtidos indicam que o algoritmo determinístico de Matching, realiza a tarefa de pontuação/ranqueamento dos provedores de nuvem em 7945 milissegundos(ms). Enquanto isso, por meio do modelo criado com o AutoML, o tempo de execução para a mesma tarefa foi de apenas 254 ms, apresentando uma melhora de 96,8% no tempo de execução quando é empregado com uma base de dados de volume elevado.

Em relação a precisão dos resultados, com o uso do AutoML e utilizando uma base de dados de provedores de nuvem reais, foi possível manter um nível de acurácia média dos resultados de 99,78% quando comparado com o algoritmo de Matching.

Assim, com os resultados obtidos, foi possível comprovar que a abordagem adotada no trabalho conseguiu ser significativamente mais rápida em comparação com a metodologia determinística. E alcançou estes resultados conservando um alto nível de acurácia sobre as previsões de pontuações dos provedores de nuvem.

Palavras-chave: Provedores de nuvem, Aprendizado de Máquina Automatizado, Nuvem computacional.

Referências Bibliográficas: Borges de Moraes, Lucas, Fiorese, Adriano, Matos, Fernando. A Multi-criteria Scoring Method based on Performance Indicators for Cloud Computing Provider Selection. SCITEPRESS, 2017.