

Seleção de Provedores de Nuvem através do Auxílio de Aprendizado de Máquina Automatizado¹

Kauã Bernardo Hopfer², Adriano Fiorese³

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

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

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

Conforme a evolução das tecnologias da informação, as organizações passaram a fazer uso das mesmas para facilitar, diminuir custos e agilizar seus serviços. Assim surge a necessidade de formas acessíveis, sob-demanda e eficientes para o acesso a recursos computacionais. Desta forma, a computação em nuvem vem se tornando uma solução adequada para o problema, sendo capaz de cumprir as necessidades listadas, hospedando e fornecendo serviços computacionais ao redor do mundo via Internet.

Assim, devido ao crescimento da demanda para serviços de computação em nuvem, as principais empresas como Google, Amazon e Microsoft desenvolveram seus provedores proprietários para serviços em nuvem e assim influenciaram outras empresas menores a também entrarem no mercado oferecendo serviços semelhantes. Entretanto, isso introduz um novo dilema, o problema de tomada de decisão acerca do provedor de nuvem mais adequado para um indivíduo ou organização.

Por essa razão, o objetivo do trabalho consiste na automatização desse processo de escolha através 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. Assim, o provedor que possuir a maior pontuação final será teoricamente o provedor de nuvem mais apropriado para aqueles requisitos. Para alcançar esses resultados, foi utilizada uma abordagem ao problema através da aplicação de modelos de *Automated Machine Learning* (AutoML) com a principal ferramenta sendo o *framework* H2O Flow. Tais modelos baseiam-se no treinamento supervisionado dos resultados obtidos pela solução de algoritmo de Matching proposto no trabalho (Moraes, 2017, p. 9-11). O uso da ferramenta AutoML H2O Flow, nos permite realizar 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.

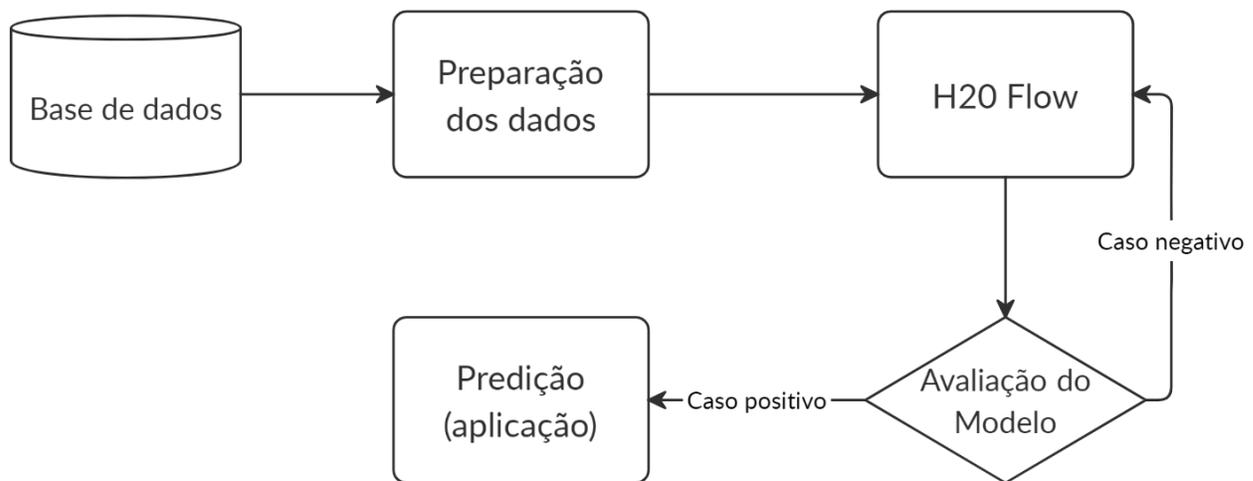
Para isso, a primeira etapa de desenvolvimento é a criação de uma base de dados de indicadores de provedores de nuvem para que possamos realizar os testes. Para o nosso caso, é gerada uma tabela randômica de indicadores de provedores de nuvem que são utilizados para descrever os provedores de nuvem participantes bem como são os indicadores que o usuário pode requerer com os valores que ele necessita. Isso visa maior riqueza e amplitude de dados. Os indicadores que compõem a base de dados de indicadores de provedores de nuvem são: quantidade de RAM, armazenamento (disco), CPU, seu nível de disponibilidade, segurança, e seu custo total.

Em seguida, a base de dados é utilizada para alimentar o algoritmo de *Matching* para que cada provedor da base de dados receba a sua pontuação (*output*). Essa pontuação será posteriormente usada para treinar o modelo de aprendizado de máquina automático (AutoML). Em

seguida, a base de dados de indicadores com os respectivos *outputs* é importada para a ferramenta H2O Flow para que possa ser iniciada a criação de um modelo de aprendizado (regressão, classificação ou agrupamento). Após o modelo auto-criado pela ferramenta, é realizada uma avaliação a fim de analisar se de fato o treinamento foi efetivo e o modelo está pronto para ser aplicado.

Enfim, se o resultado for positivo o modelo está pronto para ser aplicado em uma nova base de dados (semelhante) para efetuar a predição de pontuação dos provedores de nuvem. Entretanto, caso os resultados forem negativos deve-se voltar ao passo anterior como é ilustrado na Figura 1.

Figura 1. Fluxograma de desenvolvimento com H2O Flow



Por fim, a partir de outra base de dados gerada aleatoriamente, foi aplicado o modelo anteriormente treinado e quando comparado aos resultados do algoritmo de *Matching*, os resultados de predição do modelo obtiveram 0,61% de desvio padrão; 0,37% de variância e 99,78% de acurácia média total. Já em relação ao tempo de execução, em um teste realizado com uma base de dados de 5000 provedores de nuvem, o algoritmo de *Matching* realizou a tarefa em 7945 milissegundos (ms) enquanto isso, o modelo AutoML H2O Flow realizou sua predição em apenas 254 ms.

Assim, através desses números é possível comprovar a eficácia da abordagem proposta para o problema de tomada de decisão de escolha de provedores de nuvem para um usuário. Dessa forma a proposta se torna um meio mais rápido e simples de se obter as pontuações finais dos provedores de nuvem com um alto nível de confiabilidade.

Palavras-chave: Provedores de nuvem, *Automated Machine Learning*, 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.