

## PREDIÇÃO DA EXPERIÊNCIA DE USO EM VIDEOCONFERÊNCIAS USANDO MODELOS DE RANDOM FOREST <sup>1</sup>

Gabriel Schattenberg Küter <sup>2</sup>, Marcelo de Souza <sup>3</sup>

<sup>1</sup> Vinculado ao projeto “Segmentação dos usuários e predição da experiência de uso do serviço de conferência web da RNP”.

<sup>2</sup> Acadêmico do Curso de Engenharia de Software – ESO – Bolsista.

<sup>3</sup> Orientador, Departamento de Engenharia de Software – ESO – marcelo.desouza@udesc.br.

Este trabalho explora modelos de *random forest* para classificar a experiência de usuários de videoconferências. Para isso, é usada uma base de dados do serviço ConferênciaWeb da Rede Nacional de Ensino e Pesquisa (RNP). Essa base de dados contém informações sobre os usuários, como seu papel na reunião (moderador ou visualizador), bem como sua interação durante a execução, como o uso dos recursos (e.g., microfone ou câmera) e o tempo de permanência na reunião. Além disso, foram computados dados sobre as reuniões, como se a reunião foi grava ou não, ou o tempo médio de permanência dos usuários da reunião.

Os usuários têm a possibilidade de avaliar sua experiência de uso ao final de cada reunião. Eles avaliam a experiência com uma nota de 0 a 10. Para o estudo proposto, essa nota é convertida usando os conceitos do *Net Promoter Score* (NPS; Reichheld, 2003). O NPS é uma métrica amplamente adotada para medir a lealdade dos clientes e sua disposição em recomendar produtos ou serviços. Os clientes atribuem um valor de 0 a 10 para a chance de indicar a empresa, produto ou serviço a um amigo ou familiar. Se a avaliação estiver entre 0 e 6, ele é classificado como um detrator. Caso seja 7 ou 8, ele é classificado como neutro. Caso seja 9 ou 10, ele é um promotor. O valor de NPS é calculado subtraindo o percentual de detratores do percentual de promotores. Para a predição da experiência de uso, os modelos de *random forest* são explorados para classificar usuários em detratores, neutros ou promotores. Os modelos também são usados para classificar reuniões nessas mesmas classes, i.e. com base na média das notas atribuídas pelos participantes da reunião. Para isso, foi adotado o modelo de *random forest* (Breiman, 2001), escolhido pela sua capacidade de lidar com conjuntos de dados complexos e de alta dimensionalidade, além de ser adequado para lidar com cenários de *overfitting*.

Novas variáveis foram criadas para armazenar a classe dos usuários e das reuniões segundo os conceitos do NPS. A maioria dos usuários que avaliam sua experiência na ferramenta o fazem de maneira positiva, sendo 10 e 9 as notas mais frequentes. Por consequência, a maioria dos usuários (quase 90%) são promotores, enquanto a maioria das reuniões (quase 90%) também pertencem a essa classe. Logo, a base de dados é desbalanceada, contendo muitos exemplos da classe promotor (majoritária) e poucos exemplos das demais classes (minoritárias). Tal desbalanceamento pode levar a modelos enviesados, que tendem a prever a classe majoritária de forma excessiva, ignorando a importância das classes minoritárias, as quais, por sua vez, podem fornecer um retorno relevante para a melhoria do serviço.

Para mitigar o impacto desse desbalanceamento, foram testadas duas abordagens principais: *undersampling* e *oversampling*. O *undersampling* foi realizado utilizando a técnica Near Miss (Mani e Zhang, 2003), que seleciona amostras da classe majoritária para remoção com base na proximidade entre as amostras das classes minoritárias e majoritárias. O objetivo dessa técnica é equilibrar as classes, reduzindo o número de exemplos da classe predominante sem sacrificar a diversidade dos dados. Por outro lado, o *oversampling* foi realizado através do

método SMOTE (Synthetic Minority Over-sampling Technique; Chawla et al., 2002), uma técnica que cria amostras sintéticas para as classes minoritárias. O SMOTE aumenta a representatividade dessas classes ao interpolar entre amostras existentes, o que preserva as relações estruturais nos dados e promove uma melhor generalização do modelo.

Os modelos de random forest foram construídos considerando uma divisão de 60% dos dados para treinamento, 20% para validação e 20% para teste. O conjunto de treinamento é usado para a construção dos modelos, enquanto o conjunto de validação é usado para o ajuste dos seus hiperparâmetros, que são responsáveis por determinar a estrutura e o comportamento do modelo. Os hiperparâmetros considerados incluíram o número de estimadores (que define a quantidade de árvores de decisão na floresta), a profundidade máxima das árvores (que controla até que ponto cada árvore pode se especializar nos dados) e, por fim, o critério de divisão (que determina o critério de divisão dos dados ao criar um nodo). A busca aleatória é uma técnica eficaz para otimizar hiperparâmetros de modelos de aprendizagem de máquina, pois permite testar uma ampla gama de combinações de forma eficiente, maximizando o desempenho dos modelos para a tarefa preditiva de interesse.

A Tabela 1 apresenta o conjunto de hiperparâmetros dos modelos de random forest, seus tipos e possíveis valores. São apresentados os melhores valores encontrados pela busca aleatória considerando a classificação de usuários e reuniões usando as estratégias de undersampling (US) e oversampling (OS) para balanceamento das classes. Os melhores valores dos modelos que usam undersampling apresentam menos estimadores e menor profundidade máxima, em comparação com os modelos que usam oversampling.

**Tabela 1.** Hiperparâmetros e melhores valores encontrados.

Cenário		Estimadores int: [64, 256]	Profundidade int: [16, 128]	Critério de divisão cat: { <i>gini</i> , <i>entropy</i> , <i>log_loss</i> }
Usuários	RF+US	186	17	<i>log_loss</i>
Usuários	RF+OS	230	45	<i>gini</i>
Reuniões	RF+US	120	16	<i>entropy</i>
Reuniões	RF+OS	251	83	<i>log_loss</i>

O modelo configurado, i.e. com seus hiperparâmetros ajustados, foi então avaliado no conjunto de teste. A Tabela 2 apresenta o desempenho dos modelos de random forest nos cenários estudados. São apresentados os valores de acurácia, precisão, recall e medida F. Os valores das últimas três métricas são apresentados separadamente para as classes detrator, neutro e promotor, respectivamente. Os melhores resultados para cada cenário estão destacados em negrito. É possível perceber que os modelos de random forest são eficazes na classificação de usuários e reuniões, apresentando acurácia superior a 70% em todos os casos. Os valores de precisão e recall e medida F também são satisfatórios.

Observou-se, também, que as estratégias de *oversampling* resultaram em um desempenho significativamente melhor em termos de precisão, recall, e medida F, especialmente para as classes minoritárias. Por conseguinte, isso sugere que o SMOTE, ao aumentar a representatividade das classes detratores e neutros, conseguiu equilibrar melhor o modelo, tornando-o mais sensível a variações entre as diferentes classes e, portanto, mais adequado para a identificação de áreas onde a ferramenta de videoconferência pode ser aprimorada.

**Tabela 2.** Desempenho dos modelos de random forest.

Cenário	Acurácia	Precisão	Recall	Medida F
Usuários RF+US	72,3	(69,5; 59,2; <b>87,8</b> )	(44,6; 72,3; <b>99,9</b> )	(54,3; 65,1; <b>93,5</b> )
Usuários RF+OS	<b>79,6</b>	( <b>79,9</b> ; <b>79,7</b> ; 79,0)	( <b>84,8</b> ; <b>83,2</b> ; 70,6)	( <b>82,3</b> ; <b>81,4</b> ; 74,6)
Reuniões RF+US	77,2	(73,0; 66,3; <b>91,8</b> )	(60,3; 72,3; <b>99,6</b> )	(66,0; 69,1; <b>95,5</b> )
Reuniões RF+OS	<b>86,1</b>	( <b>84,9</b> ; <b>85,8</b> ; 87,4)	( <b>83,4</b> ; <b>79,1</b> ; 95,8)	( <b>84,2</b> ; <b>82,3</b> ; 91,4)

Como conclusões, os modelos de random forest se mostraram eficazes para classificar usuários e reuniões, apresentando alto desempenho para essa tarefa. Tais modelos são importantes ferramentas de análise de dados, pois permitem classificar a experiência de usuários e reuniões em tempo real. Ao identificar usuários detratores, é possível aplicar instrumentos para obter maiores informações sobre a sua percepção da ferramenta e sugestões de melhorias (bem como outras abordagens de intervenção). Além disso, ao identificar uma reunião detratora é possível repassar essa informação ao moderador, juntamente com sugestões para a melhoria da reunião.

**Palavras-chave:** Predição da experiência de uso. Random Forest. Net Promoter Score.

## REFERÊNCIAS

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard Business Review*, 81(12), 46-55.

Mani, I., & Zhang, I. (2003). kNN approach to unbalanced data distributions: A case study involving information extraction. *Proceedings of the Workshop on Learning from Imbalanced Datasets* (pp. 1-7).

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 16, 321-357.

Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13, 281-305.