

Classificação do Índice de Satisfação de Clientes a partir de Indicadores de Rede Móvel e de Negócios utilizando Redes Neurais

Carlos A. Martins Jr., Lisandro Lovisolo e Marcelo G. Rubinstein

Resumo — Operadoras de telecomunicações estão sob concorrência. Para planejar a operação e prever rentabilidade métricas de negócios, tais como a taxa de rotatividade dos clientes (*churn*) e o NPS (*Net Promoter Score*) são utilizadas. Neste trabalho exploramos técnicas de aprendizado de máquina para previsão do NPS e, conseqüentemente, prover meios de reduzir o *churn*. Foi utilizada uma base amostral de 82.618 clientes de uma operadora brasileira que inclui informações sobre uso e cobrança. Mostramos que, sob uma ampla variedade de métricas de rede e de negócio, é possível treinar uma máquina para entender a percepção dos clientes e assim tomar ações capazes de mantê-los.

Palavras-Chave— *Aprendizado de Máquina, RNA, Churn, NPS.*

Abstract— *Competition in the telecommunications sector is a challenge. To maintain profitability, operators rely on business metrics, such as customer turnover rate (churn) and NPS (Net Promoter Score). Thus, in this work we explore machine learning techniques for NPS classification and, consequently, churn forecasting. A base with 82,618 customers from Brazil, which includes information on usage and collection, was used. We show that, under a wide variety of network and business metrics, it is possible to train a machine to understand customer perceptions aiming to keep them.*

Keywords— *Machine Learning, RNA, Churn, NPS.*

I. INTRODUÇÃO

As comunicações móveis estão em constante evolução e, com isso, grandes inovações têm surgido trazendo novas tecnologias e ferramentas no mercado. Essa evolução impacta fortemente a sociedade que, eventualmente, se torna dependente dessas novas tecnologias. Em linha com o constante aumento no consumo de dados, a Cisco previu um CAGR (*Compound Annual Growth Rate*) de 47% na demanda por banda 2016 e 2021 [1].

A constante evolução tecnológica obriga as operadoras de telecomunicações a acompanhá-los, visando manter a competitividade e a sua base de clientes. Conseqüentemente, as operadoras têm buscado formas de entender a experiência de seus usuários e retê-los [2]. Para tal, empregam-se ferramentas de *Performance Management, Service Quality Management e Customer Experience Management*, que possibilitam avaliar o desempenho do serviço prestado pela operadora. Busca-se com grande esforço a excelência do serviço prestado/ofertado, pois a manutenção de um cliente da base tem sempre um custo inferior ao da conquista de um novo cliente [3].

Operadoras de telecomunicações em todo o mundo visam minimizar a degradação da qualidade do serviço da rede, com o objetivo de melhorar a qualidade geral da experiência dos usuários, pois uma experiência ruim poderá frustrar o usuário e resultar em uma alta taxa de rotatividade dos clientes (*churn*) [4]. Assim, a gestão da rede e da qualidade de serviço das aplicações *Over The Top* (OTT) [5] torna-se cada dia mais complexa e desafiadora com o surgimento frequente de novas aplicações e demandas. O acesso a novas aplicações acaba tornando o usuário final mais exigente, aumentando suas expectativas quanto à confiabilidade, qualidade e acessibilidade dos serviços que em grande parte são unicamente atribuídos à operadora e não ao serviço OTT. Assim as empresas de telecomunicações sentem a necessidade de se tornarem mais competitivas e de elevarem cada vez mais a qualidade dos seus produtos e serviços.

Por conseqüência, a adequação do desempenho da rede é sempre um desafio para os administradores de redes de telecomunicações. No Brasil, a qualidade da Telefonia Móvel é regulamentada e fiscalizada pela ANATEL (Agência Nacional de Telecomunicações) através da regulamentação da Telefonia Móvel (Serviço Móvel Pessoal - SMP) [6].

Com o objetivo de entender a percepção dos clientes, tem-se empregado o NPS (*Net Promoter Score*), que se tornou uma métrica popular em vários setores para avaliar a experiência e a satisfação (percepção de qualidade), com o objetivo de aumentar a fidelização de clientes. O NPS foi introduzido como uma forma de aferir a probabilidade de um cliente indicar a empresa ou o produto para outro cliente [7]. Essa métrica traz visibilidade às reclamações dos clientes e, por isso, se torna interessante avaliar a sua correlação com problemas de rede e de negócios.

O NPS utilizado neste trabalho deriva de uma única pergunta: “De zero a dez, qual a probabilidade de você recomendar nossa empresa a um amigo ou colega?”. Dessa maneira, utiliza-se uma escala para avaliar sua disposição e intenção de recomendar a empresa a outra pessoa, como apresentado na Figura 1.

O NPS é calculado como a porcentagem de respondentes considerados “promotores”, ou seja, que responderam 9 ou 10 à pergunta, menos a porcentagem de respondentes considerados “detratores”, isto é, cuja resposta foi de 0 a 6. Acredita-se que as classificações mais baixas estejam associadas a experiências negativas, que podem impedir uma recomendação ou mesmo inviabilizar novos clientes em potencial.

0	1	2	3	4	5	6	7	8	9	10
Detrator							Neutro		Promotor	

Fig.1. Mapeamento da satisfação do cliente entre classes de promoção/divulgação de uma marca, produto ou serviço.

Os promotores são considerados clientes que “passaram a ter uma vida melhor” após o relacionamento com a marca e oferecem *feedbacks* positivos, promovendo um aumento no faturamento da empresa, direta e indiretamente. Os detratores são considerados como insatisfeitos com a compra do produto ou serviço prestado pela empresa. Clientes que respondem 7 ou 8 são caracterizados como “neutros”, ou seja, não são leais à marca ou empresa. Assim, caso uma concorrente ofereça um produto em melhores condições, o cliente provavelmente optará contribuindo para a taxa de *churn* [8].

Com a concorrência, entender a experiência do cliente se torna prioridade das operadoras móveis para reter assinantes.

Um dos desafios é a previsão do *churn*; com isso, técnicas estatísticas e de aprendizado de máquina têm sido utilizadas para criar um modelo capaz de prever o *churn* [9] [10]. Há diversas, diversas patentes também têm sido registradas com o mesmo objetivo [11] [12].

Neste trabalho empregamos uma pesquisa NPS visando entender o impacto que alguns indicadores mensuráveis podem ter sobre os clientes. Aplicamos as respostas da pesquisa NPS e indicadores chave de desempenho de rede num sistema classificador de forma a tentar avaliar se um cliente está satisfeito ou não. Isso permite sinalizar possíveis casos de *churn* e assim atuar no sentido de manter o cliente na base. Visa-se correlacionar as métricas de rede e de negócio apresentadas na Tabela 1 ao resultado do NPS de um cliente daria. A Tabela 1 contém o nome de cada indicador utilizado, a média (*mean*), o desvio padrão (*std*), o valor mínimo (*min*), os percentis 25 (25%), 50 (50%) e 75 (75%) e o valor máximo (*max*).

TABELA I. INDICADORES UTILIZADOS NO SISTEMA PARA PREVER O NPS DE CLIENTES DE UMA OPERADORA DE TELEFONIA MÓVEL.

Função	Tipo	mean	std	min	25%	50%	75%	max
qt_dias_restante_fidelizacao	IN	131,23	147,39	-778,00	0,00	123,00	259,00	603,00
trafego_ransharing_sem_cota	IN	194.620.064,60	744.688.996,64	0,00	0,00	515.960,50	83.433.475,00	27.093.160.487,00
rend_percap_mensal	IN	0,04	0,07	0,00	0,00	0,01	0,04	0,60
vl_pag_sum_3m_tt	IN	128,57	171,10	0,00	0,00	70,00	208,58	3.658,38
vl_atr_sum_3m_tt	IN	16,73	119,90	0,00	0,00	0,00	0,00	6.029,85
mean_worst_rtt_operadora_3g	IQ	510,43	509,35	0,00	250,50	385,84	615,28	41.173,70
mean_worst_rtt_operadora_4g	IQ	106,86	253,02	0,00	0,00	0,00	137,36	14.171,88
mean_worst_rtt_ransharing_2g	IQ	227,03	1.446,49	0,00	0,00	0,00	0,00	82.500,00
mean_worst_rtt_ransharing_3g	IQ	196,72	729,57	0,00	0,00	0,00	0,00	30.604,66
mean_worst_rtt_ransharing_4g	IQ	47,71	250,28	0,00	0,00	0,00	0,00	16.363,67
pct_video_0_400	IQ	0,18	0,24	0,00	0,00	0,08	0,25	1,00
pct_video_400_700	IQ	0,07	0,11	0,00	0,00	0,03	0,11	1,00
pct_video_700_1000	IQ	0,06	0,10	0,00	0,00	0,03	0,09	1,00
pct_video_1000_2000	IQ	0,13	0,15	0,00	0,00	0,10	0,20	1,00
pct_video_2000_plus	IQ	0,45	0,32	0,00	0,14	0,47	0,72	1,00
worst_tp_mean	IQ	1.373,44	2.240,42	0,00	127,49	614,16	1.620,46	59.839,50
avg_changes_3g_4g	IR	30,81	29,48	0,00	8,26	25,40	44,78	607,26
avg_changes_operadora_ransharing	IR	5,44	13,79	0,00	0,00	0,41	5,26	1.058,00
total_pct_airtime	IR	0,04	0,08	0,00	0,00	0,01	0,03	1,00
trafego_total	IR	6.703.511.059,48	6.594.081.141,24	1.523,00	2.338.076.513,75	4.805.202.241,00	8.862.737.651,75	154.649.066.154,00
pct_uso_4g_operadora	IR	0,52	0,36	0,00	0,09	0,64	0,84	1,00
pct_uso_2g_ransharing	IR	0,00	0,01	0,00	0,00	0,00	0,00	1,00
pct_uso_3g_ransharing	IR	0,02	0,08	0,00	0,00	0,00	0,00	1,00
pct_uso_4g_ransharing	IR	0,03	0,09	0,00	0,00	0,00	0,01	1,00
pct_uso_3g_operadora	IR	0,43	0,35	0,00	0,13	0,30	0,77	1,00
traffic_volume_3g	IR	2.163.243.926,46	3.620.499.610,30	0,00	72.685.291,00	895.216.540,00	2.604.115.440,75	94.991.296.653,00
traffic_volume_4g	IR	2.810.684.054,39	4.614.125.546,60	0,00	0,00	792.459.646,50	3.945.031.733,00	119.351.430.243,00
classe	NPS	0,43	0,71	0,00	0,00	0,00	1,00	2,00

II. SISTEMA DE CLASSIFICAÇÃO NPS PROPOSTO

Podemos definir aprendizado de máquina a partir de: “Um programa de computador é dito aprender a partir de uma experiência E com respeito a uma classe de tarefas T e medida de desempenho P , se seu desempenho nas tarefas de T , segundo a medida P , melhora com a experiência E ” [13].

As técnicas de aprendizado de máquinas são aplicadas em classificação de classes, que pode ser dividida em não-supervisionada e supervisionada. A classificação supervisionada, utilizada neste trabalho, consiste na identificação prévia das classes de informação, que são representações do comportamento médio das classes que serão mapeadas automaticamente.

Alguns exemplos de aplicações focadas em modelos de classificação são: filtragem de *spam*, detecção de idioma, análise de sentimentos, reconhecimento de caracteres, detecção de fraude, risco de crédito entre outros. Neste trabalho, empregamos técnicas de aprendizado de máquina para classificar clientes de operadoras de telecomunicações na

escala NPS. Um classificador considera as três categorias da Figura 1. Alternativamente, testamos outro que considera os extremos da resposta ao NPS onde selecionou as respostas 0, 1 e 2 e 8, 9 e 10 dos clientes.

Para projetar os classificadores, são empregadas Redes Neurais Artificiais. Os atributos de entrada do classificador são IRs (Indicadores de Rede), IQs (Indicadores de Qualidade) e INs (Indicadores de Negócio). A saída do classificador é na escala NPS e no outro cenário dos extremos. Para obter uma base de respostas NPS para o treinamento do classificador, SMSs foram enviados para 850 mil clientes selecionados aleatoriamente. As respostas aos SMSs obtidas possui o total de 82.618, foram usadas para criar a base de dados empregada em nossos experimentos. O principal desafio foi mesclar decisões de negócio em telecomunicações com a ótica da inteligência artificial.

Visa-se avaliar a satisfação de clientes da operadora de telecomunicações em função dos indicadores de rede, indicadores de qualidade e os indicadores de negócio com uso de aprendizado de máquina.

III. SISTEMA DE APRENDIZADO PROPOSTO

Algoritmos de classificação associam padrões de entrada às classes, visando indicar o grupo de pertinência mais provável para um padrão de atributos. Tal mapeamento entre entradas e saídas pode ser “aprendido” e o algoritmo de classificação aplicará o aprendido (a função de classificação) a uma amostra desconhecida.

O sistema de classificação utilizado neste trabalho pode ser desmembrado para fins explicativos em três partes: entrada dos dados, classificador e saída. Os dados de entrada, como apresentados na Figura 2, são compostos pelos Indicadores de Rede, Indicadores de Qualidade, Indicadores de Negócio. Conjuntamente, os dados da pesquisa NPS de cada cliente são utilizados no treinamento supervisionado da Rede Neural Artificial.

Em relação aos indicadores, cumpre dizer que indicadores de Rede aferem o funcionamento dos elementos de rede. Os Indicadores de Qualidade aferem como está a qualidade no uso de aplicações e os Indicadores de Negócio aferem questões de negócio como, por exemplo, quantos dias faltam para terminar o plano de fidelização de um cliente.

Utilizamos um classificador MLP (*Multi Layer Perceptron*) cujos resultados são medidos através das métricas precisão, acurácia, *recall* e *f1-score*. Todas essas métricas estão na faixa de 0 a 100% e quanto maior for cada uma delas, melhor será o classificador subjacente.

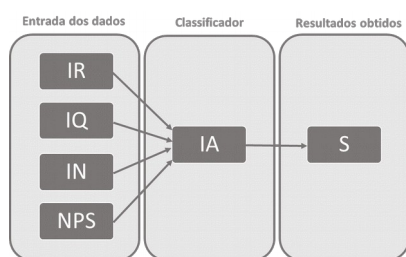


Fig.2 Sistema proposto em diagrama de blocos.

IV. DADOS USADOS PARA CLASSIFICAÇÃO

O conjunto de IR, IQ, IN e das respostas da pesquisa NPS de cada cliente foi empregado para treinar o classificador. A coleta de dados é realizada em todas as camadas da rede. Todos os dados foram obtidos a partir de uma ferramenta de *SQM-Service Quality Management* que são fornecidas por diversas empresas como Ericsson, Huawei e Nokia. Através do *SQM* selecionamos os indicadores de IR e IQ, posteriormente foram armazenados em um *Data Lake*. Os atributos de entrada do sistema proposto são apresentados na Tabela 1.

Normalmente, utiliza-se métodos para selecionar quais as funções são as melhores para serem utilizadas em um modelo de *Machine Learning* e pode-se dizer que não é um procedimento simples. Neste trabalho para selecionar quais as funções são as mais importantes foram utilizadas diferentes técnicas como a Matriz de Correlação, *Filter Methods*, *Wrapper Methods* e *Embedded Methods*. Em todas as técnicas selecionou-se 5, 10, 15 e 20 funções. Para cada caso o modelo foi treinado e aferiu-se os resultados, contudo, utilizando todas as funções obteve-se o melhor resultado.

O IN designado *qt_dias_restante_fidelizacao* com a unidade de medida em dias apresenta a quantidade de dias restantes para o término do período contratual. Caso o indicador seja igual ou inferior a zero, significa que o cliente não se encontra mais em período de fidelização e pode cancelar o plano sem cobrança de multa. Caso essa função seja maior que zero, ocorrerá incidência de multa por cancelamento do contrato.

O IN designado *trafego_ransharing_sem_cota* com a unidade de medidas em Bytes afere o tráfego utilizado pelo cliente através de rede compartilhada de outra operadora, após ter consumido todo o pacote que foi contratado.

O IN designado *rend_percap_mensal* mostra a renda mensal do cliente conforme informada no sistema de CRM da operadora.

O IN designado *valor_pag_sum_tt* é o valor faturado do cliente durante o período de coleta.

O IN designado *valor_atr_sum_tt* é o valor a faturar pelo cliente durante o período de coleta.

Os IQs designados *mean_worst_rtt* com a unidade de medidas em milissegundos representam a média dos piores RTTs (*Round Trip Times*) mapeados dos clientes utilizando a rede da operadora ou a rede compartilhada de diferentes tecnologias (gerações) durante o período de coleta.

Os IQs designados *pct_video* com valores normalizados aferem o *throughput* dos clientes utilizando os serviços de vídeo no YouTube separados por 5 faixas, sendo: 0 a 400 kbps, 400 a 700 kbps, 700 a 1.000 kbps, 1.000 a 2.000 kbps e acima de 2.000 kbps.

O IR *avg_changes_3g_4g* afere a quantidade de mudanças dos clientes entre as tecnologias 3G e 4G. Quando esse indicador tem um valor muito alto para um cliente, é possível inferir que o mesmo pode ter problemas de cobertura.

O IR *avg_changes_operadora_ransharing* afere a quantidade de vezes que o cliente muda entre a rede da operadora e o Ran Sharing o compartilhamento de redes entre operadoras. Um valor pequeno significa pouco uso de rede compartilhada.

O IR *total_pct_airtime* representa o tempo que cada cliente ficou conectado à rede móvel. Com essa função é possível identificar quanto tempo e em quais as células o cliente fica mais tempo conectado, sabendo quais são os pontos de maior interesse do cliente, como sua casa e seu trabalho, por exemplo.

O IR *trafego_total* representa o tráfego total em Bytes utilizado pelo cliente durante o período de coleta.

O IR *pct_uso* mede a porcentagem de uso do cliente em cada tecnologia e na rede compartilhada. Como exemplo, se o valor do *pct_uso_4G* for o maior, significa que esse cliente acessou a rede de acesso com a tecnologia 4G e, teoricamente, deveria ter uma melhor experiência de uso.

O IR indicador *traffic_volume* representa o volume de tráfego em Bytes enviado e recebido nas redes 4G e 3G da operadora.

A. A Pesquisa NPS

O resultado da pesquisa é apresentado na Figura 3. Essa base de dados foi utilizada no classificador de três classes (Detrator, Neutro e Promotor).

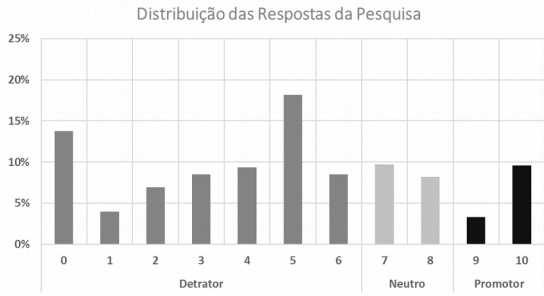


Fig.3. Distribuição das respostas aos SMSs da NPS pesquisa realizada.

Observando a distribuição das respostas, aproximadamente 54% dos clientes responderam entre 3 e 7. Clientes que respondem 0, 1 ou 2 são aqueles com pior percepção de experiência da empresa, sendo possíveis fontes de *churn* e as respostas 8, 9 e 10 correspondem aos clientes que possuem menor chance de *churn*. Considerando esse cenário, tem-se 38.945 respostas válidas nesses dois grupos extremos, para os quais obtêm-se a distribuição apresentada na Figura 4. Essa base de dados foi utilizada no classificador de duas classes.

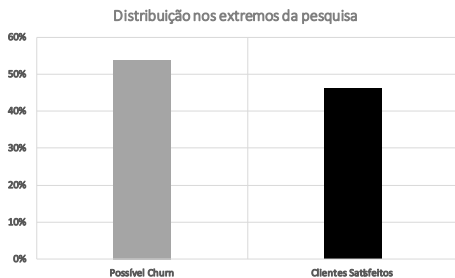


Fig.4. Proposta de pesquisa para respostas das extremidades.

V. REDE NEURAL ARTIFICIAL PARA CLASSIFICAÇÃO

Uma RNA simples pode ser definida como um processador massivo e paralelamente distribuído que foi modelado de forma simplificada com base no funcionamento do cérebro humano, exemplificado na Figura 5 [14].

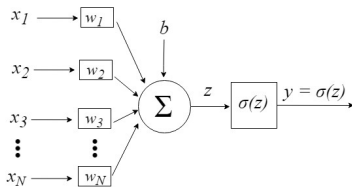


Fig.5. Modelo de neurônio artificial.

É possível resumir o modelo de um neurônio artificial em quatro passos simples descritos a seguir.

Na camada de entrada, como apresentado na Figura 5, temos os sinais de entrada (x_1 a x_N). Cada sinal de entrada é multiplicado por um peso (w_1 a w_N) que indica sua influência na saída do modelo. Somam-se então os sinais e tem-se:

$$z = \sum_{i=1}^N x_i w_i + b \quad (1)$$

Nessa equação, b provê um grau de liberdade a mais, através de um viés. O valor z é aplicado a uma função de ativação σ , resultando no valor final de saída do neurônio y .

As RNAs de vários níveis são constituídas basicamente por uma camada de entrada, uma ou n camadas ocultas e uma camada de saída. Os sinais são propagados entre camadas sucessivas. As redes de Perceptrons de Múltiplas Camadas (MLP – Multi Layer Perceptron), como a ilustrada na Figura 6, foram propostas para resolver problemas mais complexos que aqueles resolvidos pelo modelo neural basal apresentado na Figura 5. A Figura 6 apresenta uma RNA MLP que conta com uma camada de entrada com quatro neurônios, uma camada oculta com três neurônios e uma camada de saída com cinco neurônios, resultando em uma arquitetura 4-3-5.

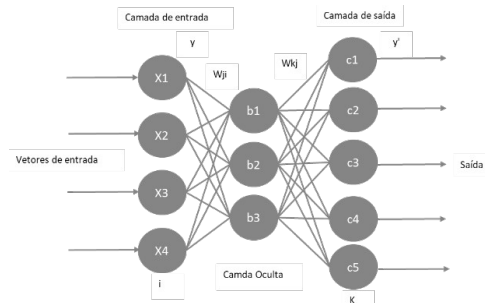


Fig.6. Arquitetura básica de RNA.

Para realizar o treinamento de uma rede neural para classificação supervisionada, é necessário fornecer amostras das classes que se pretende que a rede reconheça. Assim, um padrão X_i correspondente à classe Y_i é aplicado à rede, produzindo uma saída \hat{Y}_i . A partir disso, obtêm-se:

$$E(y, \hat{y}) = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (2)$$

Esse erro é (retro)propagado da saída para a entrada de forma a ajustar os pesos dos neurônios [14].

A. Rede Neural artificial para classificação

Neste trabalho fez-se uso de RNAs para obter o classificador de três classes com a resposta à pesquisa NPS apresentada na Figura 3 e outro classificador de duas classes com os grupos extremos da pesquisa apresentados na Figura 4, cujas entradas são os indicadores apresentados na Tabela 1.

A função de ativação empregada nas RNAs deste trabalho é a ReLU [15]. A função ReLU é comumente utilizada para treinar Redes Neurais devido à facilidade computacional que ela propicia. A camada de saída utiliza a função de ativação SoftMax [15]. A função SoftMax é muito útil em casos com problemas de classificação para mapear as saídas no intervalo entre 0 e 1. A implementação da RNA foi realizada utilizando-se Python [16].

VI. RESULTADOS OBTIDOS

Para a avaliação do desempenho dos classificadores neste trabalho, foram utilizadas as métricas Acurácia, Recall, Precisão, F1-Score e Curva ROC (*Receiver Operating Characteristic Curve*). Os dois classificadores treinados possuem as mesmas características, a menos da quantidade de classes (duas ou três). A arquitetura do modelo utilizado foi 4-3-1. Para chegar à essa arquitetura foi-se acrescentando neurônios nas camadas até que o resultado obtido não melhorasse mais.

O de duas classes classifica os clientes entre churn ou no-churn, enquanto o de três classes o faz entre detratores, neutros ou promotores. Os conjuntos de amostras foram apresentados nas figuras 3 e 4 e a proporção de treino foi com 70% da base de dados e 30% para testes. A quantidade de épocas usadas para separar o treinamento em fases distintas também foi a mesma utilizada nos dois classificadores. Os classificadores utilizam a entropia cruzada e, para aferir os resultados da RNA utilizada, foram calculados os resultados das métricas apresentadas no sistema proposto.

Percebe-se que o modelo com apenas duas classes obteve os melhores resultados, pois considera apenas os clientes nos extremos.

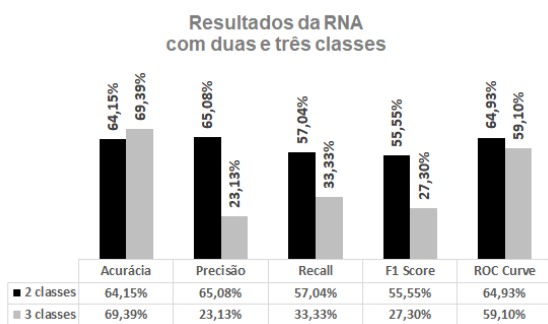


Fig. 7. Resultados obtidos com a RNA considerando as duas e três classes.

Nessa comparação, percebe-se que no caso de três classes o modelo apresentou maior acurácia; isso se deve provavelmente à maior quantidade de amostras para detratores que representam 69,15% de todos os pesquisados. Com isso, a operadora pode adicionar os novos pesquisados para melhoria contínua do modelo. O nível de acurácia apresentado ainda não é satisfatório para modelo preditivos. Contudo, a acurácia pode melhorar com novos clientes que são pesquisados mensalmente.

VII. CONCLUSÕES

No contexto dinâmico de rápidas mudanças tecnológicas em telecomunicações, o presente estudo fornece informações empíricas que podem ajudar nas decisões estratégicas do churn e do NPS. Para serem competitivos no mercado de telecomunicações, os provedores de serviços móveis precisam ser capazes de antever possíveis problemas, levando a ações proativas para reter seus clientes. Neste trabalho propusemos utilizar diferentes métricas de rede, de negócio e de aplicações

(*OTT*) para construir um modelo preditivo para o que pode ser útil para o churn.

Usando dados disponibilizados por uma operadora de telefonia móvel, validou-se essa proposta usando um classificador implementado usando rede neural artificial. Foi possível prever com alguma confiança o qual seria o resultado. A RNA foi selecionada como primeira opção e outras possibilidades serão testadas em trabalhos futuros.

Futuramente, pretendemos investigar quais são os indicadores que têm maior correlação com os problemas de a detração; com isso, para futuros trabalhos, outros modelos de classificadores que explicitem essa correlação podem ser utilizados como por exemplo o XGBoost.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio do CNPq e da FAPERJ. Agradeço à empresa que disponibilizou os dados para realizar a pesquisa, contudo, não podemos revelar a identidade.

REFERÊNCIAS

- [1] Index, Cisco Visual Networking. "Global mobile data traffic forecast update, 2016–2021." white paper 7 (2017).
- [2] Lemmens, Aurélie, and Sunil Gupta. "Managing churn to maximize profits." *Marketing Science*, Forthcoming (2020).
- [3] KOTLER, Philip, and Gary ARMSTRONG. "Philip. Administração de marketing." São Paulo: Atlas (2000).
- [4] Misra, Richa, Sonali Singh, and Renuka Mahajan. "An empirical study on the cellular subscribers churn, selection factors and satisfaction with the services." *International Journal of Product Development* 23.2-3 (2019): 105-121.
- [5] Dasgupta, Koustuv, et al. "Social ties and their relevance to churn in mobile telecom networks." *Proceedings of the 11th international conference on Extending database technology: Advances in database technology*. 2008.
- [6] Qualidade - Telefonia Móvel. Agência de Telecomunicações, 2020. Disponível em: <<https://www.anatel.gov.br/dados/controle-de-qualidade/controle-telefonia-movel>>. Acesso em: 18 de jun. de 2020.
- [7] Reichheld, Frederick F. "The one number you need to grow." *Harvard business review* 81.12 (2003): 46-55
- [8] Hung, Shin-Yuan, David C. Yen, and Hsiu-Yu Wang. "Applying data mining to telecom churn management." *Expert Systems with Applications* 31.3 (2006): 515-524.
- [9] Dasgupta, Koustuv, et al. "Social ties and their relevance to churn in mobile telecom networks." *Proceedings of the 11th international conference on Extending database technology: Advances in database technology*. 2008.
- [10] Mozer, Michael C., et al. "Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry." *IEEE Transactions on neural networks* 11.3 (2000): 690-696.
- [11] Sridhar, Kamakshi, et al. "System and method for predicting and reducing subscriber churn." U.S. Patent Application No. 16/598,112 April 2020.
- [12] Fang, Chunsheng, and Derek Lin. "Low false positive behavioral fraud detection." U.S. Patent No. 9,479,518. 25 Oct. 2016.
- [13] Mitchell, Tom M. "Machine learning. 1997." Burr Ridge, IL: McGraw Hill 45.37 (1997): 870-877.
- [14] Da Silva, I. N., D. Spatti, and R. Flauzino. *Redes Neurais Artificiais para engenharia e ciências aplicadas: curso prático*, Artliber Editora Ltda, São Paulo, SP, Brasil. ISBN978-85-88098-53-4, 2010.
- [15] Maas, Andrew L., Awni Y. Hannun, and Andrew Y. Ng. "Rectifier nonlinearities improve neural network acoustic models." *Proc. icml*. Vol. 30. No. 1. 2013.
- [16] TensorFlow Core v2.2.0, Just your regular densely-connected NN layer, 2020. Disponível em: https://www.tensorflow.org/api_docs/python/tf/keras/layers/Dense. Acesso em: 19 de jun. de 2020.