

Reconhecimento de ações judiciais relacionadas ao consumo de energia não registrado utilizando engenharia de características de relacionamento temporalizadas e rede LSTM**Recognition of lawsuits related to unregistered energy consumption using temporalized relationship characteristics engineering and LSTM network**

DOI:10.34115/basrv4n3-097

Recebimento dos originais:20/05/2020

Aceitação para publicação:17/06/2020

Pedro Thiago Cutrim

Bacharelado em Ciência da Computação pela Universidade Federal do Maranhão

Instituição: Universidade Federal do Maranhão

Endereço: Núcleo de Computação Aplicada, CCET, Av. do Portugueses, Campus do Bacanga, Bacana, São Luís - MA, Brasil

E-mail: thiago.cutrim@nca.ufma.br

João Vitor Ferreira França

Bacharel em Ciência da Computação pela Universidade Federal do Maranhão

Instituição: Universidade Federal do Maranhão

Endereço: Núcleo de Computação Aplicada, CCET, Av. do Portugueses, Campus do Bacanga, Bacana, São Luís - MA, Brasil

E-mail: jvitorfranca@nca.ufma.br

Geraldo Braz Junior

Doutor em Engenharia Elétrica pela Universidade Federal do Maranhão

Instituição: Universidade Federal do Maranhão

Endereço: Núcleo de Computação Aplicada, CCET, Av. do Portugueses, Campus do Bacanga, Bacana, São Luís - MA, Brasil

E-mail: geraldo@nca.ufma.br

Aristófanés Corrêa Silva

Doutor em Informática pela Pontifícia Universidade Católica

Instituição: Universidade Federal do Maranhão

Endereço: Núcleo de Computação Aplicada, CCET, Av. do Portugueses, Campus do Bacanga, Bacana, São Luís - MA, Brasil

E-mail: ari@nca.ufma.br

Erika Wilza Brito de Assis Lorenzo Alves

Bacharel em Direito pela Centro de Ensino Unificado de Teresina

Instituição: Equatorial Energia S/A

Endereço: Gerencia Corporativa de Contencioso e Compliance, Alameda A, Qd SQS, nº 100, Loteamento Quitandinha - Altos do Calhau, São Luís - MA, Brasil

E-mail: erika.assis@equatorialenergia.com.br

Milton Sales Lozano de Oliveira

Bacharel em Administração pela Faculdade do Estado do Maranhão - FACEM

Instituição: Equatorial Energia S/A

Endereço: Gerencia Corporativa de Contencioso e Compliance, Alameda A, Qd SQS, nº 100,
Loteamento Quitandinha - Altos do Calhau, São Luís - MA, Brasil
E-mail: milton.oliveira@equatorialenergia.com.br

Renato Ubaldo Moreira e Moraes

Mestre em Engenharia Elétrica pela Universidade Federal do Maranhão

Instituição: Equatorial Energia S/A

Endereço: Gerencia Corporativa de Contencioso e Compliance, Alameda A, Qd SQS, nº 100,
Loteamento Quitandinha - Altos do Calhau, São Luís - MA, Brasil
E-mail: renato.moraes@equatorialenergia.com.br

José Silva Sobral Neto

Bacharel em Direito pela Universidade Federal do Maranhão

Instituição: Equatorial Energia S/A

Endereço: Gerencia Corporativa de Contencioso e Compliance, Alameda A, Qd SQS, nº 100,
Loteamento Quitandinha - Altos do Calhau, São Luís - MA, Brasil
E-mail: jose.sobral@equatorialenergia.com.br

RESUMO

O grande número de ações judiciais contra empresas de distribuição de energia destaca o difícil problema de identificar e solucionar falhas de serviços neste setor. Este trabalho propõe uma metodologia para identificar novas ações judiciais no setor de energia baseado em informações do relacionamento cliente com a companhia, além da identificação de fatores correlacionados. A metodologia é basicamente dividida em 4 etapas: (a) aquisição de dados; (b) engenharia de características; (c) seleção de características; e (d) classificação usando LSTM. O método foi realizado em um banco de dados com mais de cinquenta mil consumidores e mostra-se robusto na tarefa de identificar a predição de ações judiciais de consumo de energia não registrada por meio de uma acurácia de 92,89%; especificidade de 94,27%; sensibilidade de 88,79%; e precisão de 83,84%. Assim, demonstra-se a viabilidade de usar o LSTM para resolver o problema da predição de processos judiciais de consumo de energia não registrados.

Palavras-chave: Previsão de ação judicial, LSTM, Distribuidor de energia

ABSTRACT

The great number of lawsuits against energy companies has highlighted the difficult problem of identifying and eliminating failures of services in the energy sector. This work proposes a methodology to predict the issue of new lawsuits in the energy sector on a client database and the identification of factors correlated factors. The methodology is divided into 4 stages: (a) data acquisition; (b) feature engineering; (c) feature selection; and (d) classification. The method was performed in a database with more than fifty thousand consumers and shows to be robust in the task of identifying the unregistered power consumption lawsuits prediction by achieved an accuracy of 92.89%; specificity of 94.27%; sensitivity of 88.79%; and precision of 83.84%. Thus, we demonstrate the feasibility of using LSTM to solve the problem of unregistered power consumption lawsuits prediction.

Keywords: Lawsuit filing prediction, LSTM, Power distributor

1 INTRODUÇÃO

No Brasil, a necessidade de mudanças no mercado de concessionárias de energia elétrica tem se tornado mais evidente nos últimos anos, principalmente devido ao surgimento de ações judiciais contra essas concessionárias. Somente no estado do Rio de Janeiro, aproximadamente 30 mil novas ações foram registradas a cada mês contra empresas de distribuição de energia, o que gera uma despesa aproximada de 100 milhões de reais (Almeida et al., 2014)

Quando se trata de uma concessionária de energia elétrica há grandes dificuldades na eliminação de falhas do serviço prestado (Johnston e Fern, 1999). Além disso, no Brasil, cada concessionária opera em um domínio extenso, muitas vezes abrangendo um estado completo da federação. Além de fornecer e administrar o suprimento de energia elétrica ao consumidor, ainda há uma gama de funções burocráticas ligadas a essa prestação de serviços que podem gerar insatisfação do cliente, resultando em custos judiciais para as concessionárias, como serviços: atendimento ao cliente; interrupção de energia; manutenção de equipamentos; entre outros.

Um dos problemas mais frequentes em ações judiciais é a cobrança do consumo de energia não registrado (CNR) e a regularização do faturamento de energia, causados quando há detecção de furto de energia nas instalações conectadas ao cliente. O problema é comum no Brasil, causando grandes prejuízos às empresas e à sociedade. Essa também é a realidade do Grupo Equatorial Energia, responsável pela distribuição de energia elétrica nos estados do Pará, Maranhão, Piauí e Alagoas, com mais de 5 milhões de clientes no Brasil, distribuídos em um vasto território.

Nesse contexto, é necessário que as empresas corrijam suas falhas de maneira rápida e eficiente, não apenas para evitar custos legais, mas também se preocupando com a reputação e a qualidade do serviço prestado. Além disso, entender o cliente (ou grupo de clientes) e os possíveis fatores correlatos às ações judiciais, facilita a decisão acerca de qual setor deve atuar de forma adequada para prevenir e/ou erradicar custos desnecessários.

Entender o consumidor de energia e prever seu comportamento pode contribuir significativamente para reduzir os processos judiciais, aumentando a qualidade e a eficiência da empresa, detectando o problema nos estágios iniciais e tomando as medidas corretas. Supõe-se que os dados históricos do relacionamento com o cliente da empresa podem ser um importante ponto de partida para a criação de mecanismos inteligentes de mineração e identificação de clientes que possam abrir processos legais para vários assuntos contra a empresa.

Portanto, o objetivo do presente trabalho é construir um método inteligente de identificação de clientes que possam estar relacionados à CNR, utilizando informações históricas de relação de consumo com a empresa e o classificador *Long Short-Term Memory Network* (LSTM).

As principais contribuições deste trabalho são: um método totalmente automático para prever o consumo de energia não registrado e o método aplicável industrialmente; auxilia a empresa à estabelecer o risco de receber procedimentos judiciais; ajuda os gerentes a entender a motivação dos processos judiciais e gerar ações de planejamento e prevenção; e melhorar a prestação de serviço e resposta ao cliente baseado em um método inteligente.

2 TRABALHOS RELACIONADOS

Existem trabalhos na literatura que tratam de satisfação de clientes, análise de perfil de consumidor, e rotatividade de consumidor em empresas.

Alejandro et al. (2011) propõe um modelo para medir intenções comportamentais dos clientes em uma determinada empresa tendo como resultados que a qualidade das transações surge como uma característica dominante, com um efeito forte e positivo e que não há efeitos significativos dos custos de mudança na satisfação e no comprometimento dos clientes.

A relação entre os eventos anteriores, que poderiam ser tratados individualmente, com a qualidade e fidelidade do relacionamento é o objetivo do trabalho proposto por Francisco-Maffezzolli et al. (2011) em uma empresa de telefonia móvel. O estudo foi realizado com um quadro quantitativo não probabilístico e contou com o método hipotético-dedutivo. O estudo confirmou (direta e indiretamente) as premissas de um efeito positivo e significativo entre satisfação, confiança e comprometimento e seus antecedentes.

De acordo com Siu et al. (2013), os clientes reclamam porque querem ser tratados de forma justa pela empresa quando ocorre uma falha no serviço. Este estudo investiga o papel da justiça na retenção de clientes que tiveram experiências falhas com o serviço de restaurante. Como resultado, os autores confirmam a relação de justiça com a satisfação anterior e posterior, queixa aos clientes.

He et al. (2014) preveem a rotatividade de clientes de banco comercial com base no modelo SVM e usa o método de amostragem aleatória para melhorar o modelo SVM. Quando a relação é de 1:10 (clientes com rotatividade : clientes sem rotatividade), as taxas de recall, precisão e acurácia, respectivamente, são de 56,09%, 29,50% e 98,39%.

Métodos preditivos foram construídos com uma meta correlacionada de manutenção do cliente. A previsão de mudança de serviço é um problema comum, não apenas em empresas de distribuição de energia, mas também em muitas outras empresas em outros setores, como telecomunicações (Amin et al., 2019) e bancos (Keramati et al., 2016).

O método proposto neste trabalho tem como premissa a construção de um modelo preditivo. Para tanto, utiliza-se com características extraídas e geradas em dados de relacionamento temporal de consumidores de uma empresa de energia elétrica. O objetivo é detectar, de antemão, os casos em que

pode ocorrer roubo de energia e, assim, fornecer informações preventivas para que gerentes e técnicos lidem da melhor maneira possível para resolver o problema. O objetivo é promover um ambiente preventivo que possa corroborar com a redução de processos dessa natureza.

3 MATERIAIS E MÉTODO

O método proposto neste trabalho, tem como principal objetivo a predição de ações judiciais de consumo de energia não registrados no setor de energia elétrica. As próximas subseções explicam cada etapa do método proposto.

3.1 AQUISIÇÃO DE DADOS

A base de dados utilizada neste trabalho é uma base privada, proveniente da companhia de distribuição de energia Companhia Energética do Maranhão (CEMAR). A base consiste em informações de consumo de cliente e dados temporal, adquiridos dos mais diversos setores da companhia. Esta base de dados inclui informações de: histórico de consumo, alteração de consumo, descontinuidade de fornecimento de energia, informação financeira, notas de serviço, reclamações por cliente, tráfego de energia, entre outros. Assim, percebe-se que a informação do perfil de um cliente está separada em vários domínios e banco de dados de vários setores.

Na arquitetura do banco de dados da companhia, os dados estão distribuídos em inúmeras tabelas. A ferramenta Talend (Bowen, 2012) foi utilizada para que junções entre essas tabelas fossem possíveis e a geração de um arquivo *comma-separated values* (.csv), o qual é mais viável a análise de dados. Assim, após a fase de aquisição de dados, a próxima etapa é validar quais características nessa gama de dados será de importância para a predição de CNR.

3.2 ENGENHARIA DE CARACTERÍSTICAS

Nesta seção são descritas os conjuntos de características extraídas para avaliar o perfil de um cliente de uma concessionária de energia elétrica. Ademais, também serão apresentadas algumas formas adotadas para refinar tais características, gerando novas características mais representativas e relacionadas com a temporalidade do cliente na empresa.

As seguintes características foram extraídas de toda a base de dados proveniente da etapa anterior:

- Informações gerais: informações de características individuais de cada cliente, como localização, tipo de cliente (residencial ou comercial).
- Consumo de energia: características que lidam com o perfil de consumo de cada cliente, como média e desvio padrão de consumo e consumo mês a mês.
- Falta de energia: características relacionadas à falta de energia de um transformador conectado a um cliente, frequência, ocorrência.

- Histórico de desligamentos: características relacionadas a desligamentos no fornecimento de energia de clientes.
- Faturas: características que mostram se houve faturas causadas por ter um CNR com esse cliente e histórico de pagamento de faturas.
- Ações judiciais: características relacionadas a ações judiciais anteriores tomadas pelo cliente.

É importante notar que as características extraídas do banco de dados vêm sem tratamento, às vezes categórica e não normalizada. Além disso, simular períodos temporais para entender o comportamento de um cliente ao longo do tempo, torna-se uma informação aliada para predição do seu comportamento. Assim, além da necessidade transformar essas variáveis, é importante simular intervalos temporais das mesmas. As técnicas adotadas para tratar estas variáveis foram modeladas com o Feature Tools ¹estão descritos a seguir:

- Normalização de variáveis numéricas. A normalização consiste em escalar e converter valores das diferentes variáveis selecionadas para um único intervalo de valores;
- Codificação de variável categórica para one-hot code (Buckman et al., 2018) ou para inteiro. Variáveis categóricas são geralmente não numéricas. Para habilitar o processamento por algoritmos de aprendizado de máquina, o processo de codificação transforma essas variáveis em valores numéricos;
- Geração de estatísticas como média e desvio padrão de variáveis numéricas e contadores para variáveis categóricas. Tais estatísticas podem facilitar a visualização e identificação de padrões nos dados. As estatísticas foram geradas para diferentes períodos do histórico do consumidor: de 01/01/2016 até 02/28/2019 (histórico completo), 01/02/2019 até 02/28/2019 (fevereiro), 01/01/2019 até 01/31/2019 (janeiro), 01/03/2019 até 02/07/2019 (entre as últimas oito e quatro semanas do histórico), 09/30/2018 até 01/04/2019 (entre os últimos cinco meses e oito semanas do histórico).

Assim, os dados brutos foram processados para agregar informações temporais usando estatísticas. Dessa forma, novas características foram geradas com base nos dados originais extraídos do banco de dados. Isso é importante, pois mostra o comportamento dos indivíduos em vários intervalos de tempo, além do fato de agregar as informações categóricas em colunas numéricas, sendo possível analisar diversos recursos com o aprendizado de máquina.

Depois que as características são geradas e pré-processadas, o próximo passo é usar aprendizado de máquina para gerar um modelo que possa reconhecer se um indivíduo tem ou não propensão a entrar em justiça contra a empresa. Vale ressaltar que um passo importante para um classificador é a seleção

¹ Featuretools - Uma ferramenta em Python de software livre para engenharia de características automatizada, disponível em: www.featuretools.com

de variáveis relevantes (características). Na próxima subseção, as técnicas de seleção de características serão apresentadas.

3.3 SELEÇÃO DE CARACTERÍSTICAS

Com a gama de características adquiridas e geradas na etapa de engenharia de características, há a possibilidade de certas características não serem tão representativas e/ou redundantes. Para tanto, propõe-se uma etapa de seleção de características, as técnicas de seleção tem mostrados considerável importância nas últimas décadas em diversas áreas, incluindo reconhecimento de padrões, análise de dados, recuperação de informação, biometria, sensoriamento remoto, visão computacional, processamento de imagens, aprendizado de máquinas e mineração de dados (Li et al., 2018).

Resumidamente, seleção de características é o problema de selecionar um subconjunto N a partir de um conjunto D de características baseados em algumas técnicas, podendo ser de otimização, genética, remoção, correlação entre outras. Os subconjuntos de características devem ser suficientes para descrever as classes de destino.

Neste trabalho, dois métodos de seleção de características são avaliados separadamente, o algoritmo de *Particle Swarm Optimization* (PSO) e o algoritmo de eliminação de características recursivo Recursive Feature Elimination algorithm (RFE). Usando dois tipos diferentes de algoritmos de seleção destacam a representatividade das características selecionadas e o desempenho da classificação. Estes algoritmos são explicados nas seções que seguem.

3.3.1 Algoritmo Particle Swarm Optimization (PSO)

O algoritmo do PSO é uma técnica evolutiva inspirada pelo comportamento de enxame ou colaborativo de populações biológicas (Kennedy e Eberhart, 1995). Este procura uma solução ideal alterando iterativamente a velocidade e a posição com base na experiência de voo da própria partícula e do grupo em direção à localização de G_{best} e P_{best} nas próximas iterações. G_{best} corresponde ao melhor valor de aptidão da população que qualquer partícula alcançou enquanto P_{best} corresponde ao melhor valor de aptidão da partícula que alcançou até então.

No contexto da seleção de características, existem várias variantes do algoritmo PSO (Xue et al., 2013). Utilizou-se o PSO binário (BPSO) em nosso trabalho (Vieira et al., 2013) devido à sua simplicidade. Ao usar o BPSO para seleção, a dimensionalidade do espaço de pesquisa é o número total de recursos no conjunto de dados. “1” no vetor de posição significa que a característica correspondente é selecionada e “0” caso contrário. A aptidão de uma partícula, definida por Eq. 1, é avaliada através da métrica recordatória obtida por uma técnica de regressão logística simples para realizar a classificação.

$$aptidao = \alpha(1 - sensibilidade) + (1 - \alpha)(1 - \frac{N}{D}) \quad (1)$$

onde α é o hiper-parâmetro que decide a troca entre a métrica de recall e o tamanho do subset de características N respeitando o total de números de características D em um conjunto de dados.

3.3.2 Algoritmo Recursive feature elimination (RFE)

O algoritmo RFE foi proposto por Guyon et al., 2002) para seleção de genes no problema de classificação usando máquinas de vetores de suporte. De acordo com RFE, um bom critério de classificação de características não é necessariamente um bom critério de classificação de subconjunto de características. Assim, dado um estimador externo, que pesa as características de um problema (pode ser SVM, Regressor Linear, entre outros), o objetivo do RFE é selecionar as características recursivamente considerando subconjuntos menor reduzindo o subconjunto atual.

Resumidamente, o estimador externo é treinado com um conjunto inicial de características e a importância de cada característica é obtida pela atribuição de um coeficiente para cada um. As características menos importantes são então removidas do conjunto atual de características. Este procedimento é repetido recursivamente no conjunto até que o número desejado de recursos a serem selecionados seja atingido. Observa-se que, por ser uma eliminação recursiva e iterativa, é necessário ter um parâmetro que indique o número final de recursos a serem escolhidos no final do processo de RFE. Com as características mais representativas selecionadas pelos algoritmos de seleção, a próxima etapa é classificar um indivíduo na possibilidade de ele ser ou não CNR. A próxima seção explica como foi feito esse processo de classificação utilizando o algoritmo LSTM.

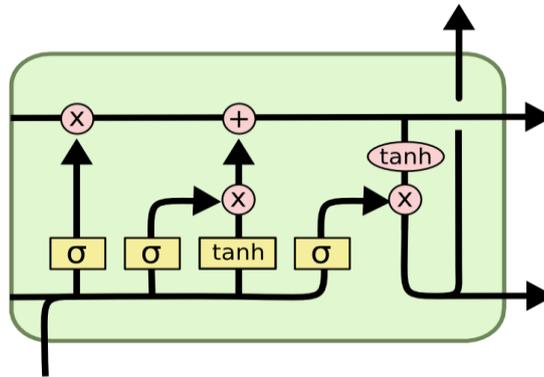
3.4 CLASSIFICAÇÃO

Na etapa de classificação, primeiramente a base de características é dividida em treino e teste. No treinamento é criado um modelo capaz de aprender padrões a partir dos dados treinos aqui realizado por uma rede LSTM e a partir disto classificar novos indivíduos provenientes da base de teste.

A *Long-Short Term Memory* (LSTM) foi proposta por Hochreiter and Schmidhuber (1997), porém somente com o advento das placas gráficas foi possível utilizar ela em longa escala e produzir resultados promissores em diversos domínios de problemas (Li and Wu, 2015)

A LSTM é um tipo especial de rede neural recorrente utilizada no campo de aprendizado profundo. Diferente das redes *feedforward* convencionais, estas possuem um módulo de repetição composto por uma única camada de rede neural que se retroalimenta. Esses módulos são compostos por quatro camadas que interagem de forma especial, como pode ser visto na Figura 1.

Figura 1 Composição da LSTM



A LSTM é composta por um estado de célula, que é a linha horizontal que atravessa o topo do diagrama. Ela atravessa a cadeia inteira sofrendo apenas pequenas alterações causadas pelas estruturas conhecidas como "gates". Esses são meios que opcionalmente deixam a informação atravessar a cadeia, e são compostos por uma camada sigmoide e uma operação de multiplicação. O sinal sigmoide tem como saída valores de 0 a 1, que significa passar o sinal inteiramente ou não passar. Uma LSTM possui três gates como esse para proteger e controlar o estado da célula.

Resumidamente, o primeiro gate é o *forget*, que olha os valores de x e a e gera um número de 0 a 1 para cada número no estado da célula. O próximo passo é para decidir qual nova informação será armazenada no estado da célula. Para isso há duas partes: o *input gate* que decide quais valores serão atualizados, e a segundo *tanh gate* que cria um vetor de novos valores candidatos. Finalmente, é decidido qual será o *output*: primeiro, roda-se uma camada sigmoide que decide quais partes do estado farão parte da saída. Depois submete-se o estado da célula a uma função tanh (resultando em valores de -1 a 1) e multiplica-se pela saída da função sigmoide.

Neste trabalho, as características de clientes com e sem CNR são apresentados para rede e um modelo de predição será gerado para que seja capaz de classificar novos indivíduos de possíveis consumidores em entrar com ações judiciais. Para validar o método, métricas de validações são extraídas do modelo. Esse trabalho usa métricas comuns em problemas de classificação binária. As métricas de validação de desempenho são acurácia, precisão, sensibilidade, especificidade e F-measure (Amin et al., 2017).

4 RESULTADOS E DISCUSSÃO

Para avaliação, o banco de dados adquirido foi dividido em dois conjuntos: treinamento e teste. Os consumidores CNR foram obtidos a partir de dados históricos de ações judiciais com esse assunto. Consumidores sem CNR foram selecionados aleatoriamente entre outros consumidores. Manteve-se uma proporção de 1 CNR para 3 clientes sem CNR, a fim de realizar experimentos com dados

desequilibrados que são esperados para o problema. Um total de 56.894 clientes foram analisados. A proporção entre clientes com CNR e sem CNR no conjunto de dados de treinamento e teste está descrita na Tabela 1.

Tabela 1. Proporção entre clientes com CNR e sem CNR no conjunto de dados de treinamento e teste

Base de dados	Cliente com CNR	Cliente sem CNR	Total
Treino	8.562	25.554	34.116
Teste	5.712	17.066	22.778
Total	14.274	42.620	56.894

Como apresentado na Tabela 1, a base de dados é muito heterogênea quando se trata do número de clientes. Analisou-se mais de 50 mil clientes com a metodologia proposta, tanto para criação de um modelo no treinamento, quanto para o teste, há um desbalanceamento da base.

Na etapa de engenharia de características, para cada indivíduo foram adquiridas 1100 características, destas 245 eram características nativas, i.e., retiradas diretamente da base de dados. Outras 855 características foram geradas pelo processo de variação temporal e transformação de variáveis categóricas em *one-hot encode*. Assim, o método mostra-se robusto, tanto na análise de 50 mil indivíduos quanto o fato de cada um estar relacionado à um número expressivo de características.

Com os conjuntos de dados definidos e as características geradas, a próxima etapa é a seleção de características. Nesta etapa, os algoritmos PSO e RFE serão usados para selecionar as melhores características entre os 1100. O algoritmo PSO ao término de sua execução selecionou 619 características. Por sua vez, o RFE necessita de um parâmetro que indica quantas características haveriam ao fim de sua execução. Para contornar o problema de parametrização, separou-se 10% da base de treinamento e a cada iteração do algoritmo, validavam-se as métricas nesta base separada, o algoritmo parava sua execução quando não havia melhora significativa nas métricas. Assim, ao fim do algoritmo RFE, foram selecionadas 595 características.

Finalmente, na etapa de classificação, o conjunto de dados de treinamento foi usado para gerar o modelo de predição via LSTM. Este modelo foi testado com os dados dos consumidores no conjunto de dados de teste. A Tabela 2 resume os resultados, incluindo a precisão (Acc), precisão (Prec), sensibilidade (Sens), especificidade (Spec) e F-measure. Estes resultados foram obtidos usando o conjunto de testes.

Tabela 2. Resultado da classificação utilizando LSTM

Experimentos	Características	Acc (%)	Prec (%)	Sens (%)	Spec (%)	F-measure (%)
LSTM	1100	92,46	83,31	87,50	94,12	85,35
LSTM + PSO	619	91,52	77,94	92,36	91,24	84,54
LSTM + RFE	595	92,89	83,84	88,79	94,27	86,25

Observando a Tabela 2, os melhores resultados alcançados com LSTM foi em conjunto com o algoritmo RFE. Utilizando esta combinação, consegue-se uma redução significativa de características com métricas balanceadas e promissoras, com uma acurácia de 92,89%, precisão de 83,84%, sensibilidade de 88,79%, especificidade de 94,27% e F-measure de 86,25%. Quando se trata de taxa de sucesso em acertar consumidores com CNR, deve-se observar a métrica de sensibilidade, onde a combinação de LSTM + PSO alcançou os melhores resultados, contudo, as outras métricas não foram tão altas.

Ao analisar todos os experimentos, nota-se que tanto as características escolhidas, quanto o LSTM são robustos na tarefa de classificação de consumidores CNR, demonstrando um grande potencial para ser utilizado como ferramenta no setor de energia elétrica a fim prevenir possíveis ações judiciais, quanto melhorar a satisfação do cliente. Também vale ressaltar que, mesmo com o desequilíbrio da base de dados, o LSTM conseguiu alcançar uma maior sensibilidade do que especificidade, garantindo robustez em acertar os consumidores que possam a ter ações de CNR.

Vale ressaltar que medidas preventivas são cada vez mais comuns e necessárias para melhorar os serviços das empresas de energia elétrica. Além disso, entender o comportamento do cliente ajuda a empresa a responder de forma rápida e eficaz, não apenas ao lidar com ações judiciais, mas também com preocupação pela reputação da empresa em relação ao consumidor e pela prestação de serviço de qualidade. Assim, observa-se que o método proposto neste trabalho que atingiu uma precisão de 83,84% e sensibilidade de 88,79%, desempenha um papel fundamental se empregado por empresas de energia elétrica pode ajudar tanto a empresa na redução de custos de processos judiciais, quando uma tomada de decisão rápida, aumentando a satisfação do consumidor.

5 CONCLUSÃO

Este trabalho investigou o desempenho do modelo LSTM em conjunto com técnicas de seleção de características (PSO e RFE) para previsão de processos de consumo de energia não registrados no setor elétrico com base em dados de clientes. Os melhores resultados foram obtidos usando LSTM e o algoritmo de RFE, apresentando uma acurácia de 92,89%, precisão de 83,84%, sensibilidade de 88,79%,

especificidade de 94,27% e F-measure de 86,25%. Esses resultados mostram o quanto o método é robusto, e que, se empregado por companhias do setor elétrico, pode tanto auxiliar na redução de ações judiciais quanto na melhoria da satisfação do cliente.

No futuro, pretendemos investigar teoricamente e experimentalmente o método proposto, considerando várias outras questões pertinentes. Em primeiro lugar, a base de dados é desequilibrada, pelo que o cliente com CNR (classe minoritária ou classe de interesse) contém um número menor de amostras em comparação com o cliente sem CNR (classe majoritária). Em segundo lugar, a eliminação e a detecção de outliers contribuiriam muito para fornecer melhores resultados. Finalmente, analise outras técnicas de seleção de recursos e faça uma comparação detalhada.

AGRADECIMENTOS

Este trabalho foi apoiado pelo projeto Sipaju, financiado pela Equatorial Energia e pela Agência Nacional de Energia Elétrica (ANEEL) através do P&D N° PD-00037-0031/2018.

REFERÊNCIAS

- Almeida, V. M. C. D., Figueiredo, O. H. D. S., & Silva, G. F. D. (2014). Propensão do Cliente à Proposição de Ações Judiciais: Proposição e Teste de Modelo Preditivo para o Setor de Energia Elétrica. *Revista Brasileira de Marketing*, 13(5), 01-18.
- Johnston, R., & Fern, A. (1999). Service recovery strategies for single and double deviation scenarios. *Service Industries Journal*, 19(2), 69-82.
- Alejandro, T. B., Vieira, V. A., Monteiro, P. R., & Veiga, R. T. (2011). Relationship marketing in supply chain: an empirical analysis in the Brazilian service sector. *Journal of Business & Industrial Marketing*.
- Francisco-Maffezzoli, E. C., Prado, P. H. M., da Silva, W. V., & Marchetti, R. Z. (2011). Evaluation of the customer relationship quality and propensity to change mobile telephone operators. *Brazilian Business Review*, 8(4), 1-22.
- Siu, N. Y. M., Zhang, T. J. F., & Yau, C. Y. J. (2013). The roles of justice and customer satisfaction in customer retention: A lesson from service recovery. *Journal of business ethics*, 114(4), 675-686.
- He, B., Shi, Y., Wan, Q., & Zhao, X. (2014). Prediction of customer attrition of commercial banks based on SVM model. *Procedia Computer Science*, 31, 423-430.
- Amin, A., Al-Obeidat, F., Shah, B., Adnan, A., Loo, J., & Anwar, S. (2019). Customer churn prediction in telecommunication industry using data certainty. *Journal of Business Research*, 94, 290-301.
- Keramati, A., Ghaneei, H., & Mirmohammadi, S. M. (2016). Developing a prediction model for customer churn from electronic banking services using data mining. *Financial Innovation*, 2(1), 10.

- Bowen, J. (2012). Getting started with talend open studio for data integration. Packt Publishing Ltd.
- Buckman, J., Roy, A., Raffel, C., & Goodfellow, I. (2018). Thermometer encoding: One hot way to resist adversarial examples.
- Li, J., Cheng, K., Wang, S., Morstatter, F., Trevino, R. P., Tang, J., & Liu, H. (2017). Feature selection: A data perspective. *ACM Computing Surveys (CSUR)*, 50(6), 1-45.
- Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-International Conference on Neural Networks (Vol. 4, pp. 1942-1948)*. IEEE.
- Xue, B., Zhang, M., & Browne, W. N. (2012). Particle swarm optimization for feature selection in classification: A multi-objective approach. *IEEE transactions on cybernetics*, 43(6), 1656-1671.
- Vieira, S. M., Mendonça, L. F., Farinha, G. J., & Sousa, J. M. (2013). Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients. *Applied Soft Computing*, 13(8), 3494-3504.
- Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1-3), 389-422.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- Li, X., & Wu, X. (2015, April). Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 4520-4524). IEEE.
- Amin, A., Anwar, S., Adnan, A., Nawaz, M., Alawfi, K., Hussain, A., & Huang, K. (2017). Customer churn prediction in the telecommunication sector using a rough set approach. *Neurocomputing*, 237, 242-254.