

Modelos para Previsão Tributária Utilizando Redes Neurais LSTM

LSTM Artificial Neural Network Models for Tax Forecasting

Arthur F. Dornelas¹, Luciana D. Campos¹, Karla Figueiredo²

¹Departamento de Ciência da Computação – Universidade Federal de Juiz de Fora
(UFJF) – Juiz de Fora – MG – Brazil

²Departamento de Informática e Ciência da Computação / IME
(UERJ) – Rio de Janeiro – RJ – Brazil

arthurdornelas99@gmail.com, luciana.campos@ice.ufjf.br,
karlafigueiredo@ime.uerj.br

Abstract. *ICMS (Imposto sobre Circulação de Mercadorias e Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação) is one of the main taxes collected by Brazilian states, being its important value in the management and planning of the government, especially for the state of Rio de Janeiro, that presents itself in economic crisis and since the year 2020 is in a Tax Recovery Regime. Due to uncertainties and external and internal changes in the state, the forecast of this collected value has a characteristic of non-linearity, being necessary the application of non-linear models that can consider these changes in the amounts collected over time. Therefore, the work described here aims to use models of Recurrent Neural Networks Long Short-Term Memory (LSTM) and compare the approaches Multivariate Multi-step and Univariate Multi-step, in an attempt to generate an annual forecast of the state's tax collection that is superior to other approaches, and can be used as parameters for decision-making by government authorities.*

Resumo. *O ICMS (Imposto sobre Circulação de Mercadorias e Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação) é um dos principais impostos arrecadados pelos estados brasileiros, sendo seu valor importante na gestão e planejamento do governo, em especial para o estado do Rio de Janeiro, que se apresenta em crise econômica e desde o ano 2020 está em Regime de Recuperação Fiscal, necessitando de uma constante atualização da previsão de seus valores de receita e gastos. Devido às incertezas e mudanças externas e internas no estado carioca, a previsão desse valor coletado possui característica de não-linearidade, sendo necessário a aplicação de modelos não lineares que possam considerar essas mudanças nos valores arrecadados ao longo do tempo. Por conseguinte, o trabalho aqui descrito visa utilizar modelos de Redes Neurais Recorrentes Long Short-Term Memory (LSTM) e comparar as abordagens Multivariate Multi-step e Univariate Multi-step, na tentativa de gerar uma previsão anual da arrecadação tributária do estado superior à de outras abordagens, podendo ser utilizados como parâmetros para a tomada de decisões das autoridades governamentais.*

1. Introdução

A organização e planejamento nas contas públicas é essencial para o bom funcionamento do estado. A Lei de Responsabilidade Fiscal (Lei Complementar nº 101, de 04/05/2000) [Planalto 2000] propõe parâmetros e condições para os gastos públicos dos estados e municípios com o objetivo de que seja possível evitar situações econômicas indesejadas que possam tornar as entidades não funcionais financeiramente, além de proporcionar a herança administrativa para as próximas gestões. Caso ocorram infrações quanto à lei, sanções e penalidades podem ser aplicadas, como multas, cassação de mandatos e até reclusão dos responsáveis.

Para realizar o planejamento fiscal, deve-se estimar as receitas, que são os recursos financeiros arrecadados que englobam impostos, taxas, contribuições, entre outros, e as despesas, que são todos os gastos realizados, seja na aquisição de produtos, serviços, obras ou compras, sendo eles fixados para garantir que não ocorra mais gasto do que arrecadação [Governo 2020]. Em questão das receitas, o Imposto sobre Circulação de Mercadorias e Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação (ICMS), que foi regulamentado pela Lei Kandir [Planalto 1996], é um tributo estadual que incide na circulação de produtos e serviços tributáveis nas cidades, estados ou de pessoas jurídicas para pessoas físicas. É o principal meio de arrecadação de tributos do estado, sendo parte principal no montante arrecadado das receitas e consequentemente de extrema importância para o planejamento e organização da gestão [SEFAZ-RJ 2022].

Em 2020, o estado do Rio de Janeiro entrou no Regime de Recuperação Fiscal (Resolução CSRRF nº33 2020) [SEFAZ-RJ 2020], possibilitando o estado que está em um momento de grave crise financeira a reajustar suas contas na busca de tornar-se novamente sustentável. Uma das necessidades desse regime é a constante atualização dos dados de receitas e despesas do estado, de forma que seja possível acompanhar a evolução da recuperação.

Por conseguinte, possuir uma forma melhor de realizar a previsão da arrecadação do ICMS, através da análise das séries temporais, com uma menor taxa de erro é de suma importância para o estado, sendo providencial para a tomada de decisões e ações que têm impacto na vida dos contribuintes e na organização da gestão.

Devido a mudança contínua do ambiente externo e interno do estado, existe uma não-linearidade no problema, o que invalida uma série de modelos estatísticos-matemáticos que são propostos para problemas que tem critério de linearidade. Portanto, a utilização desses modelos não-lineares de *Machine Learning*, como as Redes Neurais Artificiais, podem apresentar resultados mais eficientes pois esses modelos conseguem capturar as mudanças não lineares dos dados analisados.

Este trabalho tem como objetivo principal ajustar modelos de previsão utilizando séries temporais para realizar a previsão anual do valor arrecadado do ICMS no estado do Rio de Janeiro. Em segundo plano, o trabalho visa treinar e avaliar modelos de Redes Neurais LSTM (*Long Short-Term Memory*) com apenas uma variável (*Univariate*) e no contexto de múltiplas variáveis (*Multivariate*). Os melhores resultados obtidos de cada modelo são comparados com o atual método de previsão utilizado pelo governo do estado, sendo possível avaliar a capacidade das soluções utilizando Redes Neurais LSTM em

relação às outras abordagens, e conseqüentemente, gerando uma forma de melhorar as previsões e informações disponíveis para os órgãos governamentais.

2. Fundamentação Teórica

2.1. Rede Neural Artificial

Rede Neural Artificial (RNA) tem como objetivo modelar o funcionamento do cérebro humano de forma digital, na tentativa de atingir a capacidade de processamento paralelo, não linear e altamente complexo do órgão animal [Haykin 1999].

2.2. Rede *Long Short-Term Memory*

A rede neural recorrente tradicional possui um problema em relação às memórias de longo prazo. Em longos intervalos de tempo o custo é muito elevado em questão de tempo e aprendizado. Dessa forma, em [Hochreiter and Schmidhuber 1997] foi proposto o modelo arquitetural para redes neurais recorrentes chamado de *Long Short-Term Memory*, pautada em um apropriado aprendizado a partir de gradiente, com o propósito de conseguir acessar essas memórias longas.

Na rede recorrente tradicional, o algoritmo de backpropagation, responsável pela alteração dos pesos e conseqüentemente pelo aprendizado, sofre o “problema do gradiente de fuga”. Assim, o gradiente de longo prazo que é usado para atualizar os pesos começa a diminuir, e se esse valor se torna muito pequeno, sua interferência no aprendizado também torna-se muito baixa. Portanto, as informações relacionadas a esses períodos distantes são perdidas, gerando uma alta dependência aos valores mais atuais [Hochreiter and Schmidhuber 1997].

Por conseguinte, a LSTM tem como propósito melhorar esse tipo de aprendizado de longo prazo através da utilização de células de memória e portões. A arquitetura da célula de memória funciona como o caminho para toda a sequência que a informação será processada. Ela apresenta três portões, que são pequenas redes neurais que decidem quais informações terão progresso durante o fluxo na célula. Esses portões são: o *forget gate* que é responsável por classificar a informação necessária dos passos anteriores, o *input gate* por decidir qual informação do passo atual deve ser agregada, e o *output gate* tem o papel de selecionar qual o próximo estado escondido. Dessa forma, várias células são conectadas gerando toda a forma de aprendizado da rede.

2.3. Trabalhos Relacionados

Devido a importância do tema na organização dos estados, temos uma sequência de trabalhos realizados utilizando o conceito de redes neurais aplicados em séries temporais na tentativa de atingir previsões melhores.

Em [Silva and Figueiredo 2020] foi realizado o trabalho de previsão do ICMS do estado do Rio de Janeiro, com a utilização da variável endógena, que é apenas o valor do imposto ao longo do tempo, e o conceito de Machine Learning chamado de *Long Short-Term Memory*(LSTM), comparando com abordagens de Rede Neural *Multi-Layer Perceptron* (MLP). O trabalho descreve a utilização de previsão de múltiplos passos auto-regressivo e janelamento deslizante de 12 meses, além da comparação entre tratamento de dados utilizando apenas normalização e diferenciação e a realização de manipulações

matemáticas. Como conclusão, a utilização de rede LSTM foi superior a MLP, com a utilização de apenas normalização e diferenciação no pré-processamento dos conjuntos.

Em [Castanho 2011] foi utilizado os dados de ICMS do estado do Espírito Santo de 2000 até 2009 na tentativa de prever a arrecadação do ano 2010. Foi utilizado o modelo Holt-Winters unido a metodologia de Box-Jenkins e modelo econométrico misto com regressão múltipla e apenas uma equação comportamental. Foram analisados as diferentes variáveis e setores que têm impacto no tributo e influenciam na arrecadação. Foi utilizado o erro percentual absoluto médio (MAPE) para decidir qual modelo apresentou o melhor desempenho, chegando ao resultado de que para previsões de curto prazo os modelos de Holt-Winters foram superiores, para médio prazo os melhores modelos foram de Box-Jenkins, enquanto os modelos econométricos são melhores adaptados a longo prazo.

Em [Pessoa et al. 2011] escolheram a arrecadação do período de 1998 até 2011 no estado de Minas Gerais e foi realizada a comparação entre os modelos ARIMA e ARFIMA para a previsão do período. O modelo ARIMA (1, 0, 1) apresentou resultados superiores considerando as métricas de Raiz Quadrada do Erro Quadrado Médio de Previsão (RQEMP), Erro Absoluto Médio de Previsão (EAMP) e Coeficiente de Desigualdade de Theiler (CDT). Já o modelo ARFIMA (1, 0.36, 1) foi superior em relação ao Erro Absoluto Médio Percentual de Previsão (EAMPP). Como conclusão, os dois modelos apresentaram capacidade de realizar a previsão e atingir bons resultados para a tomada de decisão dos órgãos públicos.

Em [Scheffer et al. 2014] o trabalho foi baseado na previsão do ICMS do estado do Rio Grande do Sul do período de 1998 até 2014. Utilizaram como base para previsão os modelos Box-Jenkins com a implementação de modelos ARIMA, chegando no melhor resultando sendo um modelo ajustado SARIMA, em que a base da análise foi de resíduo e testes de significância dos parâmetros, sendo comparado com diversos modelos concorrentes.

Dessa forma, vemos que existe a necessidade por trabalhos que tem como objetivo a previsão do tributo, porém temos muitas técnicas novas de redes neurais e machine learning que ainda podem ser aplicadas e comparadas, visando atingir resultados melhores que possam auxiliar os governantes e sociedade.

3. Metodologia

A metodologia utilizada nesse trabalho envolve os modelos de rede neural artificial LSTM *Univariate* e *Multivariate* para a previsão anual do valor do ICMS do estado do Rio de Janeiro.

3.1. Pré Processamento dos Dados

Os dados da série temporal do ICMS foram retirados do site da SEFAZ-RJ ¹, sendo eles os valores mensais arrecadados do período de janeiro de 2002 a dezembro de 2019. Os dados referentes ao PIB (Produto Interno Bruto) foram retirados do site IpeaData ² e os dados do consumo de energia elétrica comercial da região sudeste foram retirados do site

¹<http://www.fazenda.rj.gov.br/sefaz/>

²<http://www.ipeadata.gov.br/>

Empresa de Pesquisa Energética³, sendo eles também referentes ao período de janeiro de 2002 a dezembro de 2019.

A divisão da base de dados foi realizada de acordo com o método *holdout* [Campos], sendo separados em três conjuntos diferentes: treinamento, validação (*early stopping*) e teste, sendo eles respectivamente de janeiro de 2019 a dezembro de 2019, de janeiro de 2017 a dezembro de 2018 e de janeiro de 2002 a dezembro de 2016. No tratamento dos dados foram aplicados apenas o processo de normalização, em cada uma das variáveis utilizadas.

Para a organização dos dados de entrada para a utilização em um aprendizado supervisionado, foram utilizados como entrada os dados referentes aos 24 meses anteriores aos 12 meses de previsão, sendo realizado um processo de janelamento deslizante [Silva and Figueiredo 2020] para a progressão dos meses de previsão na etapa de treinamento.

Foi utilizado também o método *multi-step* direto [Chevillon 2007], de forma que seja feita a previsão de 12 meses utilizando o mesmo conjunto de entradas de 24 meses anteriores, sem nenhum tipo de processo auto-regressivo ou de atualização das entradas.

3.2. Seleção das Variáveis Exógenas

Foram avaliados diversos dados de produtos e serviços que possuíam a cobrança de ICMS e selecionados os que apresentavam uma maior correlação. Para realizar esse processo foi utilizado a função de correlação da biblioteca Pandas⁴, em que utilizando o método de correlação de Pearson, o coeficiente gerado possui um valor de -1 a 1. Quanto mais próximo de -1 menor é a relação entre as variáveis e quanto mais próximo de 1 maior é a relação das mesmas [Pearson 1896]. Sendo assim é possível quantificar a semelhança entre as variáveis externas com a variável de interesse que é a arrecadação do ICMS.

A partir da análise foi possível concluir que as variáveis de PIB e Consumo de Energia Elétrica Comercial da Região Sudeste possuem alto grau de correlação com o valor mensal arrecadado de ICMS, chegando a níveis de coeficientes maiores que 0.95. Dessa forma, elas foram selecionadas para serem entradas da rede neural Multivariate, por questão de simplificação nas análises e por possuírem um alto grau de correlação.

3.3. Avaliação

Para a avaliação dos modelos foi utilizado a métrica de erro percentual médio absoluto (do inglês, Mean Absolute Percentage Error - MAPE)(equação 1), referentes aos dados dos anos de 2017 e 2018.

$$MAPE = \frac{1}{n} * \sum_{t=1}^n \frac{(y_t - \hat{y}_t)}{y_t} * 100 \quad (1)$$

Cada cenário de teste foi executado 10 vezes e a média de MAPE considerada para selecionar qual obteve melhor resultado, sendo o menor valor o de melhor acurácia. Sendo assim, é possível comparar os parâmetros utilizados para cada topologia e selecionar qual atingiu melhor capacidade de previsão para o problema.

³<http://www.epe.gov.br/>

⁴<https://pandas.pydata.org/>

4. Estudo de Casos

Nesta seção são apresentados o estudo de casos realizados e os resultados obtidos comparando as arquiteturas LSTM *Univariate* e LSTM *Multivariate* para a previsão anual do ICMS no estado do Rio de Janeiro.

Para todos os cenários de teste foram feitas 10 execuções, sendo as métricas de erro a média final. Para a LSTM *Univariate*, como mostrado na Tabela 1, foi utilizado como parâmetros o *mini-batch* igual a 32, os algoritmos otimizadores comparados foram a ADAM com taxa de aprendizado 0.001 e SGD com taxa de aprendizado 0.01. Já em relação as camadas, a primeira camada LSTM foi testada em todos os casos com os números de neurônios 400, 500, 550, 600 ou 700, e a segunda camada *Dense* sempre com 12 neurônios, pois serão 12 meses de previsão. Por fim, foi utilizado as funções de ativação Sigmóide e Relu.

Para a LSTM *Multivariate*, como mostrado na Tabela 2, foram utilizados a mesma combinação de parâmetros, porém adicionado também o teste com a utilização de duas variáveis, sendo elas PIB e ICMS e com três variáveis, PIB, ICMS e Consumo de Energia Elétrica Comercial da Região Sudeste. Foi utilizado em ambas arquiteturas o *early stopping* a partir dos dados de validação para interromper o treinamento e evitar *overfitting*. Sendo os dados de treinamento correspondentes ao período entre 2002 a 2016 e os anos de 2017 e 2018 para a validação.

Tabela 1. Cenários de teste para LSTM *Univariate*.

Cenário	Otimizador	Função de Ativação	Número de Neurônios
1	Adam	Relu	400, 500, 550, 600, 700
2		Sigmóide	
3	Sgd	Relu	
4		Sigmóide	

Tabela 2. Cenários de teste para LSTM *Multivariate*.

Cenário	Otimizador	Função de Ativação	Número de Neurônios	Variáveis
5	Adam	Relu	400, 500, 550, 600, 700	ICMS e PIB
6		Sigmóide		
7	Sgd	Relu		
8		Sigmóide		
9	Adam	Relu	400, 500, 550, 600, 700	ICMS, PIB e Consumo de Energia Elétrica Comercial - Região Sudeste
10		Sigmóide		
11	Sgd	Relu		
12		Sigmóide		

Na Tabela 3 são apresentados os cenários de teste da arquitetura *Univariate* com seus respectivos erros médios de validação. O cenário 3 obteve o melhor resultado para o erro MAPE, sendo a topologia composta por 550 neurônios na 1ª camada e 12 neurônios na 2ª camada, função de ativação relu e algoritmo de otimização Sgd.

Tabela 3. Comparação dos resultados dos cenários de experimento da arquitetura LSTM Univariate.

Erro Médio de Validação				
Cenário	Número de Neurônios	Otimizador	Ativação	MAPE (%)
1	550	Adam	Relu	6.95
2	600	Adam	Sigmóide	7.18
3	550	Sgd	Relu	6.88
4	500	Sgd	Sigmóide	10.84

Na Tabela 4 são apresentados os cenários de teste da arquitetura *Univariate* com seus respectivos erros médios de validação. O cenário 6 obteve o melhor resultado para o erro MAPE nas arquiteturas *Multivariate* com 2 variáveis, sendo a topologia composta por 400 neurônios na 1ª camada e 12 neurônios na 2ª camada, função de ativação sigmóide e algoritmo de otimização Adam. O cenário 10 obteve o melhor resultado para o erro MAPE nas arquiteturas *Multivariate* com 3 variáveis, e consequentemente, a melhor acurácia entre todos os cenários de teste, sendo a topologia composta por 500 neurônios na 1ª camada e 12 neurônios na 2ª camada, função de ativação sigmóide e algoritmo de otimização Adam.

Tabela 4. Comparação dos resultados dos cenários de experimento da arquitetura LSTM Multivariate.

Erro Médio de Validação					
Cenário	Número de Variáveis	Número de Neurônios	Otimizador	Ativação	MAPE (%)
5	2	600	Adam	Relu	6.51
6	2	400	Adam	Sigmóide	6.17
7	2	550	Sgd	Relu	6.88
8	2	700	Sgd	Sigmóide	6.34
9	3	500	Adam	Relu	6.17
10	3	500	Adam	Sigmóide	5.21
11	3	600	Sgd	Relu	6.63
12	3	700	Sgd	Sigmóide	6.24

Portanto, podemos concluir que a LSTM *Multivariate* com a utilização de 3 variáveis, sendo elas o ICMS, PIB e Consumo de Energia Elétrica Comercial da Região Sudeste, conseguiu o melhor desempenho quanto a métrica de erro MAPE, com 5.21%, e por conseguinte, foi o melhor modelo encontrado. Seguido dela temos a LSTM *Multivariate* com a utilização de 2 variáveis, sendo elas o ICMS e o PIB, com erro de MAPE de 6.17%. E por último, temos o modelo da LSTM *Univariate* com a utilização de 1 variável, que é o valor endógeno do ICMS. Ela atingiu em seu melhor caso o erro de MAPE de 6.78%.

Dessa forma, conclui-se que a utilização de variáveis externas com alta correlação no modelo LSTM é o mais benéfico na previsão anual do valor do ICMS do estado do Rio de Janeiro, sendo melhor até 2 variáveis exógenas do que apenas 1.

4.1. Avaliação dos Melhores Modelos

Como comparativo, foi gerado o erro relativo das previsões para os anos de 2017 e 2018, que é o disponível pelo SEFAZ-RJ[SEFAZ-RJ 2022] e foi analisado o desempenho dos melhores modelos LSTM *Univariate*, LSTM *Multivariate* com 2 variáveis, LSTM *Multivariate* com 3 variáveis e a previsão feita pela SEFAZ-RJ.

O erro relativo é o valor absoluto da diferença do valor previsto pelo valor real, com o resultado sendo dividido pelo valor real.

A Tabela 5 mostra essa comparação, onde é possível concluir que as arquiteturas LSTM possuíram melhor desempenho que a atual previsão feita pela SEFAZ-RJ. Além disso, é possível analisar que o comportamento dos experimentos se manteve, sendo a LSTM *Multivariate* com as variáveis ICMS, PIB e Consumo de Energia Elétrica Comercial na Região Sudeste o melhor desempenho entre as redes neurais artificiais analisadas.

Tabela 5. Comparação dos resultados dos melhores cenários de experimento das arquiteturas LSTM com a previsão do SEFAZ-RJ, para os anos de 2017 e 2018.

Modelo	Número de Variáveis	Ano 2017		Ano 2018	
		Valor (R\$ mil)	Erro Relativo	Valor (R\$ mil)	Erro Relativo
Valor Real	-	32.362.495,90	-	35.836.058,09	-
SEFAZ-RJ	-	35.287.454,00	9.04 %	33.746.760,00	5.83 %
LSTM Univariate	1	31.975.661,56	1.19 %	34.653.904,89	3.29 %
LSTM Multivariate	2	32.514.114,30	0.46 %	34.962.612,54	2.43 %
LSTM Multivariate	3	32.495.584,51	0.41 %	35.242.139,26	1.65 %

É possível identificar a previsão para o ano de 2019 através do gráfico da Figura 1, onde observa-se que a LSTM *Multivariate* se aproxima mais do comportamento da série original no ano em relação a LSTM *Univariate*.

Por conseguinte, nota-se um comportamento mais acertivo dos dois modelos nos meses iniciais e que se perde ao decorrer do ano, o que era um comportamento esperado da abordagem *Multi-step* utilizada.

5. Conclusão e Trabalhos Futuros

O objetivo deste trabalho foi desenvolver uma rede neural artificial para previsão de séries temporais com o modelo *Long Short-Term Memory*, que é uma alternativa para a resolução do problema de memória de longo prazo das redes recorrentes tradicionais, através da implementação de três portões e células de memória.

O problema aplicado foi a necessidade de previsão anual do valor arrecadado do ICMS no estado do Rio de Janeiro. Como o estado está em situação de crise econômica e atualmente faz parte do Regime de Recuperação Fiscal, o planejamento fiscal é primordial, sendo necessária a previsão anual do valor do ICMS para a Lei Orçamentária Anual (LOA). A Secretaria de Estado de Fazenda do Rio de Janeiro (SEFAZ-RJ) já realiza a previsão anual, porém utilizando modelos estatísticos matemáticos.

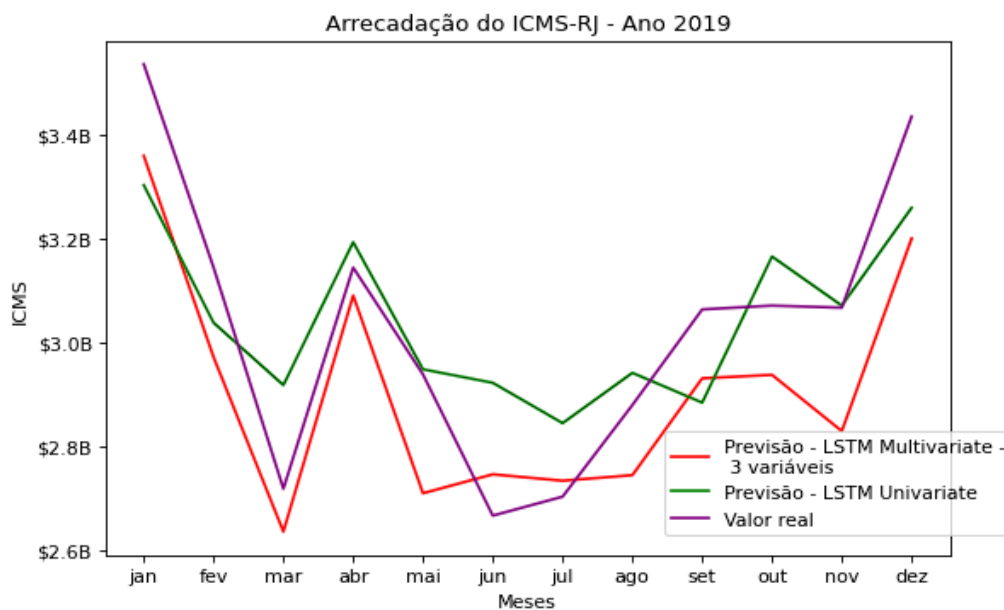


Figura 1. Gráfico com a arrecadação real do ICMS no ano de 2019 e as previsões da LSTM Multivariate e Univariate. Elaborado pelo autor.

De acordo com o trabalho de [Silva and Figueiredo 2020], os dados utilizados foram do período de janeiro de 2002 até dezembro de 2016 para a etapa de treinamento, janeiro de 2017 até dezembro de 2018 para a validação e janeiro de 2019 até dezembro de 2019 para o teste. Sendo assim, foram desenvolvidos dois modelos distintos de LSTM na tentativa de comparar a previsão anual do imposto entre eles e com o atual meio de previsão do estado.

O modelo LSTM *Univariate Multi-step* foi utilizado com sua única variável endógena sendo o valor arrecadado do ICMS e atingiu resultados superiores aos encontrados no atual modelo utilizado pelo estado carioca. Já o modelo LSTM *Multivariate Multi-step* foi separado de duas formas, uma utilizando apenas duas variáveis, que foram o ICMS e o PIB, e um com três variáveis, utilizando o ICMS, PIB e Consumo de Energia Elétrica Comercial da Região Sudeste. O modelo com duas variáveis apresentou melhor desempenho nos experimento que o modelo *Univariate*, e por conseguinte, também superou o atual modo de previsão do SEFAZ-RJ. O modelo com três variáveis foi o que apresentou melhor acurácia entre todos nos experimentos e também superou a previsão do órgão governamental carioca.

É importante ressaltar que a presença de mais dados da série temporal pode favorecer ainda mais o desempenho da rede neural. Foi feito a tentativa de adicionar mais camadas LSTM e camadas Dense, porém não atingiram resultados superiores. Também foi feita a tentativa de utilizar o modelo *Multi-step* auto regressivo, que atualiza os dados de previsão nas entradas das próximas previsões, porém não foi possível chegar a resultados satisfatórios com essa abordagem para a previsão anual. Foi elaborado também durante o trabalho um modelo de Multi-task Learning para as arquiteturas *Multivariate*, porém a necessidade de apenas prever os valores do ICMS e os resultados do modelo não ultrapassarem o modelo *Multivariate* implementado, mostram que é uma abordagem complexa para o cenário utilizado e que não gerou retornos compatíveis.

Portanto, concluímos que o modelo LSTM *Multivariate Multi-step* é mais eficiente na previsão das séries temporais anuais do ICMS no estado do Rio de Janeiro e pode ser utilizado pelos órgãos públicos nas previsões do imposto.

Em trabalhos futuros, podem ser implementados a utilização de mais variáveis exógenas de alta correlação, na tentativa de atingir melhor desempenho e acurácia, incluindo também números diferentes de horizontes de entrada e até de saída. Caso a previsão dos valores das variáveis além do ICMS também seja necessário, indica-se a utilização da abordagem *Multi-task Learning* [Zhang and Yang 2017], que pode gerar o aprendizado em compartilhado das variáveis ao mesmo tempo que as saídas sejam separadas. Outra abordagem indicada é a de utilização de outras técnicas de redes neurais artificiais, como a *Gated Recurrent Unit* (GRU) [Cho et al. 2014], sendo realizadas de forma separada ou em conjunto com as camadas LSTM.

Referências

- Campos, L. *Modelo Estocástico Periódico baseado em Redes Neurais*. PhD thesis, Pontífica Universidade Católica do Rio de Janeiro.
- Castanho, B. (2011). *Modelos para Previsão de Receitas Tributárias: O ICMS do Estado do Espírito Santo*. Universidade Federal do Espírito Santo.
- Chevillon, G. (2007). *Direct multi-step estimation and forecasting*. *Journal of Economic Surveys* 21(4).
- Cho, K., Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. (2014). *Learning phrase representations using RNN encoder–decoder for statistical machine translation*.
- Governo (2020). *Receitas e despesas*. <https://www.gov.br/economia/pt-br/aceso-a-informacao/receitas-e-despesas>.
- Haykin, S. (1999). In *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, Patparganj, India, 2nd edition.
- Hochreiter, S. and Schmidhuber, J. (1997). *Long short-term memory*. *Neural Computation Anais*.
- Pearson, K. (1896). *Mathematical contributions to the theory of evolution III*. *Philos Trans R Soc Lond Ser B Biol Sci*.
- Pessoa, F., Coronel, D., and Lima, J. (2011). *Previsão de Arrecadação de ICMS para o Estado de Minas Gerais: Uma Comparação entre modelos Arima e Arfima*. *Revista Brasileira de Gestão e Desenvolvimento Regional*.
- Planalto (1996). *Lei Complementar Nº 87, de 13 de maio de 1996. Institui o ICMS*. http://www.planalto.gov.br/ccivil_03/leis/LCP/Lcp87.htm.
- Planalto (2000). *Lei Complementar Nº 101, de 04 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal*. http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm.
- Scheffer, D., Souza, A., and Zanini, R. (2014). *Utilização de Modelos Arima Para Previsão da Arrecadação de ICMS do Estado do Rio Grande do Sul*.

SEFAZ-RJ (2020). Resolução nº 33 do conselho de supervisão do regime de recuperação fiscal do estado do rio de janeiro. http://www.fazenda.rj.gov.br/sefaz/content/conn/UCMServer/path/Contribution%20Folders/transparencia/RecuperacaoFiscal/docs/item%207/Resolucoes/SEI_ME%20-%209673492%20-%20Resoluc%cc%a7a%cc%83o%2033-2020.pdf?lve.

SEFAZ-RJ (2022). Transparência da receita estadual. <https://portal.fazenda.rj.gov.br/transparencia-da-receita-estadual/>.

Silva, P. F. and Figueiredo, K. (2020). *Aprendizado Profundo Aplicado na Previsão de Receita Tributária Utilizando Variáveis Endógenas*. ENIAC.

Zhang, Y. and Yang, Q. (2017). An overview of multi-task learning.