

Comparação de Técnicas de Aprendizagem de Máquina para Previsão de Precipitações em Manaus

Rafaela dos Santos Sousa, Elloá B. Guedes, Maria Betânia Leal de Oliveira

¹Escola Superior de Tecnologia
Universidade do Estado do Amazonas
Av. Darcy Vargas, 1200 – Manaus – Amazonas

{rass.eng, ebgcosta, mloliveira}@uea.edu.br

Abstract. *This paper addresses the rainfall forecasting in Manaus with different Machine Learning techniques. Considering this goal, we used monthly rainfall historical data from 65 years as well as Niño indexes as input into four forecasting approaches: decision trees, random forests, artificial neural networks and nearest neighbors. More than 3800 models were proposed, trained and tested from which it was possible to see that the best results were provided by neural networks, with an F-score of 70%. The results obtained can be used by several institutions for strategic decision making in order to minimize the negative effects of precipitation in this city.*

Resumo. *Este trabalho endereça a previsão de precipitação em Manaus utilizando diferentes técnicas de Aprendizagem de Máquina. Para tanto, foram considerados registros históricos de 65 anos da precipitação mensal na cidade e índices Niño, segundo quatro abordagens de previsão: árvores de decisão, florestas aleatórias, redes neurais e vizinhos mais próximos. Mais de 3800 modelos foram propostos, treinados e testados, dentre os quais os melhores resultados foram verificados nas redes neurais, com F-score de 70%. Os resultados obtidos podem ser utilizados por diversos órgãos para tomada de decisões estratégicas para minimizar os efeitos negativos da precipitação nesta cidade.*

1. Introdução

Prever a ocorrência de precipitações na cidade de Manaus não é uma tarefa fácil. A maior parte das chuvas nesta cidade é resultado da influência de muitos sistemas de precipitações, tais como a Zona de Convergência Intertropical, Anticiclones Subtropicais, Alta da Bolívia, dentre outros, além da localização geográfica e topografia [da Silva 2012]. Apesar desta inerente dificuldade e complexidade, prever precipitações auxilia na tomada de decisões ligadas à Agricultura, prevenção de alagamentos e também na administração de recursos hídricos.

Neste contexto, a previsão de precipitações costuma ser endereçada segundo duas abordagens diferentes. A primeira delas envolve o estudo dos fenômenos de precipitação com o intuito de modelar as leis físicas que os regem. Porém, esta abordagem têm alguns fatores impeditivos, pois: (i) a precipitação é o produto final de processos atmosféricos complexos, que variam tanto no espaço quanto no tempo; (ii) mesmo que estes processos sejam descritos de maneira concisa e completa, o número de cálculos envolvidos pode ser proibitivo; e (iii) os dados que estão disponíveis para auxiliar nesta tarefa são limitados tanto nas dimensões espaciais quanto temporais. Na tentativa de contornar estas dificuldades é que se baseia a segunda abordagem, fundamentada no reconhecimento de padrões. Segundo ela, características relevantes tanto espaciais quanto temporais são extraídas de padrões históricos de precipitação, podendo ser utilizadas na previsão de novos eventos [Luk et al. 2001].

Considerando a abordagem baseada no reconhecimento de padrões e as precipitações na cidade de Manaus, este trabalho se propõe a investigar diferentes técnicas de Aprendizagem de Máquina para prever o volume anual de chuva na cidade em questão. Embora alguns trabalhos tenham feito previsões de precipitação para a cidade de Manaus, considerando um dia [Guedes et al. 2016] ou um mês à frente [de Araújo et al. 2017], não há registros de trabalhos com esta abordagem em previsões de mais longo prazo. Além disso, este trabalho também considera diferentes métodos para reconhecimento de padrões, enquanto a literatura para a cidade de Manaus foca primariamente nos resultados obtidos apenas com redes neurais artificiais.

Para apresentar os resultados obtidos, este trabalho está organizado como segue. Uma breve descrição a respeito das técnicas de Aprendizagem de Máquina utilizadas no trabalho é apresentada na Seção 2. Os trabalhos relacionados presentes na literatura e usados como base de estudo estão citados na Seção 3. A metodologia encontra-se descrita na Seção 4. Uma visão geral do conjunto de dados utilizado neste trabalho é mostrada na Seção 5. Os resultados obtidos são detalhados e discutidos na Seção 6. Por fim, as considerações finais são listadas na Seção 7.

2. Fundamentação Teórica

O *Aprendizado de Máquina* (AM) é uma subárea da Inteligência Artificial que lida com a elaboração de soluções empíricas com a utilização de modelos e métodos computacionais que aprendem a partir de dados [Witten et al. 2011]. Teorias estatísticas são utilizadas na elaboração destes modelos e métodos, cuja principal tarefa consiste na realização de inferências a partir de amostras [Alpaydın 2010].

O ferramental existente de modelos e métodos utilizados pela AM é extenso, especialmente se considerados os tipos de tarefas endereçados por esta área [Faceli et al. 2015]. No escopo deste trabalho, cuja tarefa é de classificação mediante aprendizado supervisionado, destacam-se as árvores de decisão, florestas aleatórias, redes neurais artificiais e k -vizinhos mais próximos. Estes métodos e técnicas são apresentados brevemente a seguir.

As *árvores de decisão* são uma estrutura de dados hierárquica do tipo grafo acíclico implementadas a partir de estratégias de divisão e conquista. Cada nó implementa uma função de testes com resultados discretos, que servem para rotular arestas. A partir de um conjunto de dados de entrada, um teste é aplicado para cada nó e uma das arestas é percorrida de acordo com o resultado do rótulo. Esse processo inicia no nó raiz e é repetido recursivamente até atingir um nó folha. Esse nó é o atributo alvo e descreve a saída [Alpaydın 2010].

As *florestas aleatórias* baseiam-se na utilização de árvores de decisão variadas e no resultado majoritário das árvores que compõem a floresta como estratégia de classificação. O uso deste método têm se popularizado bastante em virtude dos bons resultados obtidos em diversos problemas e da sua performance [Marsland 2015].

As *redes neurais artificiais*, por sua vez, são formalmente definidas como sendo sistemas distribuídos massivamente paralelos, formadas a partir de unidades de processamento simples e com uma propensão natural para o armazenamento e posterior disponibilização de conhecimento obtido a partir da experiência [Haykin 2009].

Por fim, o algoritmo *k -vizinhos mais próximos* (k -NN) está entre as técnicas de aprendizado de máquina mais simples para realizar o reconhecimento de padrões e difere das anteriores pela ausência do processo de treinamento, bastando apenas armazenar na memória as informações dos exemplos conhecidos [Haykin 2007]. De acordo com

Marsland, o princípio de funcionamento deste algoritmo consiste em analisar exemplos similares à nova amostra para determinar a qual classe ela pertence [Marsland 2015].

3. Trabalhos Relacionados

A previsão de precipitações com métodos de AM é um tema importante e considerado em diversas partes do mundo [Bushara and Abraham 2013, Nayak et al. 2013, Darji et al. 2015]. Dentre os resultados que podem ser observados na literatura, há uma predominância na previsão de precipitações utilizando redes neurais artificiais, com melhores resultados considerando intervalos anuais do que diários e mensais.

No caso de Manaus, em particular, alguns trabalhos anteriores endereçaram a questão da previsão de precipitações com AM, especialmente considerando o uso de redes neurais artificiais [Guedes et al. 2016, de Araújo et al. 2017, dos Santos Sousa et al. 2017]. Utilizando dados observacionais coletados na cidade de Manaus por meio de uma estação meteorológica automática, o trabalho de Guedes et al. considera a previsão da ocorrência de precipitações com redes neurais artificiais e séries temporais [Guedes et al. 2016]. Embora tenham obtidos bons resultados, este trabalho não considera o volume desses eventos de precipitação.

Na tentativa de superar estas limitações, o trabalho de Araújo et al. considerou dados observacionais e também índices climáticos da temperatura do Oceano Pacífico para prever a anomalia do volume mensal de precipitações, concebendo redes neurais para os diferentes meses do ano [de Araújo et al. 2017]. Embora tenham se restringido apenas à utilização de um modelo de Aprendizagem de Máquina, os resultados obtidos mostram uma boa adequação dos atributos preditores ao domínio do problema considerado, que também será adotado no escopo deste trabalho.

Um terceiro trabalho que colaborou para a previsão de precipitações em Manaus considerou a antecipação de curtíssimo prazo na ocorrência de precipitações na região central da cidade a partir de precipitações em estações meteorológicas localizadas na região metropolitana da cidade [dos Santos Sousa et al. 2017]. Os resultados obtidos mostraram a dificuldade de previsão desta tarefa, a qual também considerou apenas as redes neurais como modelo de aprendizagem.

Desta maneira, é possível perceber duas lacunas: a ausência de trabalhos que considerem o volume anual de precipitações em Manaus e também a necessidade de utilização de outros métodos e técnicas da AM para esta tarefa, com o intuito de fornecer um panorama comparativo. Apesar das limitações, aspectos positivos destes trabalhos relacionados devem ser enfatizados, pois a utilização de AM, de dados observacionais e de índices climáticos como atributos preditores e a organização destes segundo séries temporais também serão utilizados neste trabalho.

4. Metodologia

Esta seção tem por objetivo caracterizar a metodologia utilizada para conduzir as atividades deste trabalho. O primeiro passo consistiu em caracterizar o atributo alvo a ser previsto: a *anomalia da precipitação anual*. Ao invés de prever a ocorrência de um determinado volume de precipitação anual, o que viria a requerer um trabalho de altíssima precisão, penalizando erros na casa de milímetros, optou-se por seguir uma abordagem comum à Meteorologia: considerar a variação da precipitação anual em relação à média histórica, se para mais ou para menos. Deste modo, o problema a ser tratado foi abordado como sendo de previsão um ano à frente em duas classes binárias do atributo alvo (precipitação anual acima ou abaixo da média histórica).

Como atributos preditores foram consideradas medidas de precipitação mensal obtidas de estações meteorológicas automáticas localizadas em Manaus [INMET 2018] e também a temperatura na superfície do mar em diversos pontos do Oceano Pacífico, caracterizando os índices Niño 1+2, 3, 3.4 e 4. Estes índices estão fortemente associados à variabilidade da chuva na bacia amazônica [Fisch et al. 1998], razão pela qual foram considerados no escopo deste trabalho. Um outro atributo preditor também utilizado foi a anomalia da precipitação dos anos anteriores.

Os atributos preditores foram defasados segundo janelas de tempo de 1, 2, 3 e 4 anos anteriores. A utilização de diferentes janelas de tempo auxiliam a determinar qual a quantidade de dados preditores é mais adequada para fornecer de um histórico suficiente na previsão da variável desejada.

De acordo com estes passos, foi possível então consolidar um conjunto de dados, cuja descrição será apresentada detalhadamente na Seção 5. Este conjunto de dados foi então particionado em duas partes, uma para treino com 70% do total de amostras, e outra para testes, com o restante das amostras, utilizadas para o aprendizado dos padrões pelos seguintes modelos: árvores de decisão, florestas aleatórias, redes neurais artificiais e k -vizinhos mais próximos.

Em uma tarefa de classificação binária, os possíveis resultados são os valores verdadeiros positivos (tp), verdadeiros negativos (tn), falsos positivos (fp) e falsos negativos (fn). Levando isto em consideração, da execução de cada modelo de aprendizagem de máquina foram derivadas as seguintes métricas de desempenho: acurácia, precisão, revocação e F -score. Estas medidas são obtidas da seguinte forma:

$$\begin{aligned} \text{Acurácia} &= \frac{tp + tn}{tp + tn + fp + fn} & \text{Precisão} &= \frac{tp}{tp + fp} \\ \text{Revocação} &= \frac{tp}{tp + fn} & \text{F-score} &= \frac{2 \cdot tp}{2 \cdot tp + fp + fn} \end{aligned}$$

No contexto deste trabalho, as métricas de desempenho adotadas auxiliam na avaliação comparativa dos diferentes modelos de aprendizagem de máquina utilizados e podem ser interpretadas como segue. A acurácia reflete a proporção de resultados com anomalia acima da média histórica dentre o total de resultados obtidos. A precisão, por sua vez, indica a chance de uma previsão de anomalia positiva estar correta. A revocação pode ser compreendida intuitivamente como a habilidade do modelo encontrar todos os exemplos positivos. E, por fim, o F -score é a média harmônica entre precisão e revocação, com melhores resultados quando próxima de 1 e piores resultados quando próxima de 0.

Para evitar *overfitting* no treinamento dos modelos ou que os resultados obtidos por estes fossem influenciados por algum viés da aleatoriedade na apresentação dos dados ao modelo ou na partição entre os conjuntos de treino e teste, foi utilizada a técnica de *validação cruzada* com $k = 3$ *folds* para validar os resultados. Esta técnica aumenta a confiança nos resultados obtidos, refletindo adequadamente o poder de generalização de cada modelo [Marsland 2015].

5. Visão Geral do Conjunto de Dados

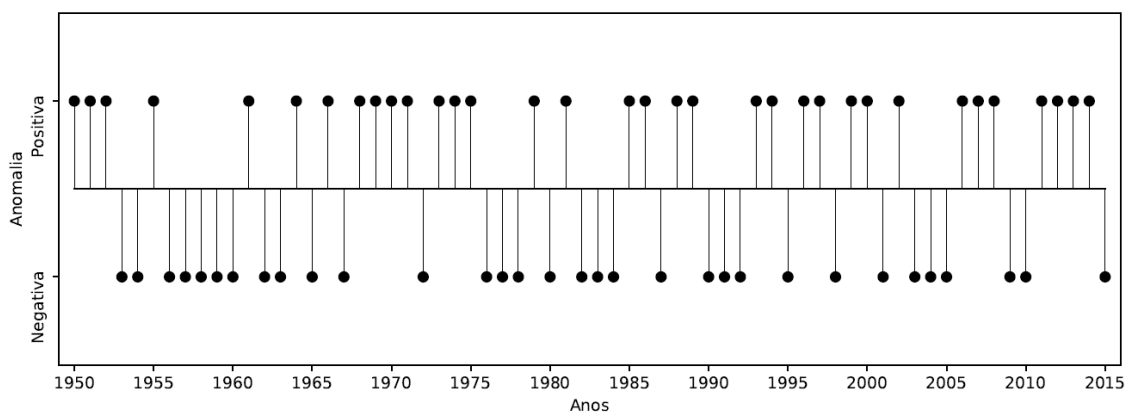
De acordo com a metodologia descrita, foram coletados os dados que vieram a compor o conjunto de dados para treino e testes dos modelos de aprendizagem de máquina para

previsão do volume anual de precipitações em Manaus. Estes dados contemplam os anos de 1950 a 2015 e uma breve estatística descritiva dos mesmos pode ser vista na Tabela 3.

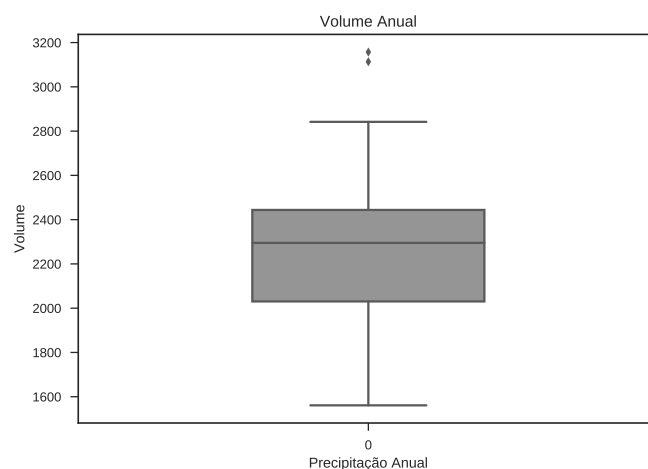
Tabela 1: Estatística descritiva dos atributos preditores.

	Média	Mediana	Desvio Padrão
Precipitação anual	2272.21	2295.0	318.72
Niño 1+2	23.01	22.8	0.9
Niño 3	25.76	25.66	0.68
Niño 3.4	26.88	26.87	0.66
Niño 4	28.35	28.39	0.56

Em relação ao atributo alvo, a anomalia da precipitação anual, não foi possível observar a existência de padrões no mesmo, conforme Figura 1, o que evidencia a não-trivialidade desta tarefa. Um outro gráfico ilustrando a distribuição do volume de precipitações anuais, ilustrado na Figura 2, indica uma dispersão dos valores e até mesmo a existência de *outliers*.



(a) Anomalia da precipitação anual.



(b) Boxplot do volume anual de precipitações.

Figura 1: Gráficos da precipitação anual em Manaus.

Os dados foram sistematicamente organizados segundo séries temporais e defasados segundo as quatro janelas de tempo estabelecidas na metodologia. O passo seguinte consistiu, portanto, no treinamento e testes dos diferentes modelos, cujos resultados são apresentados e discutidos na seção a seguir.

6. Resultados e Discussão

As técnicas propostas para o cenário em questão foram implementadas, treinadas e testadas com ferramenta da linguagem de programação Python. A escolha da linguagem Python se deu pela ampla quantidade de bibliotecas desenvolvidas para Aprendizagem de Máquina, a exemplo do *framework* `scit-kit learn`¹ e da biblioteca `pandas`².

Levando em conta as diferentes técnicas de AM a serem exploradas no escopo deste trabalho, o passo seguinte foi a definição de diferentes parâmetros e hiperparâmetros para proposição de diferentes modelos dentre as técnicas consideradas. No caso das árvores de decisão, por exemplo, utilizou-se diferentes métricas de entropia para proposição das árvores, as quais foram posteriormente comparadas de acordo com os critérios de desempenho especificados na Seção 4. A proposição dos diferentes modelos dentro de cada técnica seguiu a técnica de *grid search*, comumente adotada para otimização de parâmetros.

Ao final da etapa de proposição de modelos, obteve-se 2888 árvores de decisão, 184 florestas aleatórias, 528 redes neurais e 240 valores para o algoritmo de k -vizinhos mais próximos. Cada um destes modelos foi então submetido aos procedimentos de treinamento e teste. Deste último, após a realização da validação cruzada visando diminuir vieses nos resultados, obteve-se as métricas de desempenho. Para uma melhor compreensão destes resultados, a Figura 2 e a Tabela 2 sintetizam a métrica F -score. Na Figura 2, esta métrica é ilustrada considerando o agrupamento por métodos para cada janela de tempo. A Tabela 2, por sua vez, sumariza as melhores métricas obtidas por técnica e por janela de tempo, ilustrando os resultados dos melhores modelos identificados.

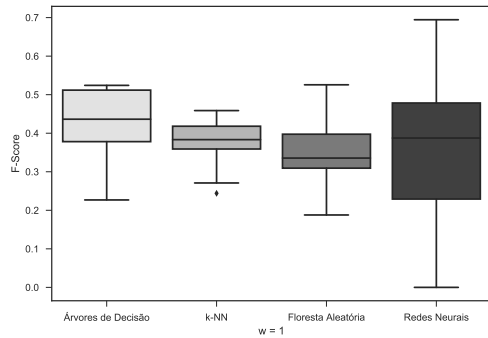
Interpretando os resultados obtidos, é possível perceber que o maior valor da métrica de F -score observado residiu nas redes neurais, cujo F -score médio foi de 0.3925 com desvio padrão igual a 0.15. De acordo com estes dados, é importante enfatizar que, mesmo com melhores resultados quando comparados às demais técnicas (árvores de decisão, florestas aleatórias, k -NN), não houve uma uniformidade nas taxas de acerto destas redes, o que pode ser verificado pela medida de dispersão (desvio padrão).

Analisando mais atentamente os resultados obtidos pelas redes neurais, conforme Figura 3, é possível observar que o melhor desempenho foi verificado na janela de tempo igual a $w = 4$, obtida pela rede neural com arquitetura (245, 11, 5, 1), isto é, duas camadas ocultas com 11 e 5 neurônios, respectivamente. A Tabela 3 mostra a matriz de confusão produzida por esta rede neural para o conjunto de testes.

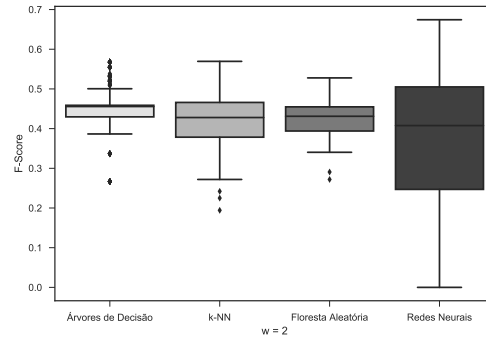
Os resultados obtidos desta comparação de diferentes técnicas e métodos de Aprendizagem de Máquina corroboram a percepção da literatura, conforme Seção 3, a respeito da adequação das redes neurais para este domínio de problema. Algumas características deste modelo em si podem estar mais diretamente relacionadas à esta adequação, em especial: (i) boa captura da relações não-lineares; (ii) resistência ao ruído; (iii) aproximação universal de qualquer função; (iv) o fato de serem dirigidas à dados, etc.

¹<http://scikit-learn.org/>

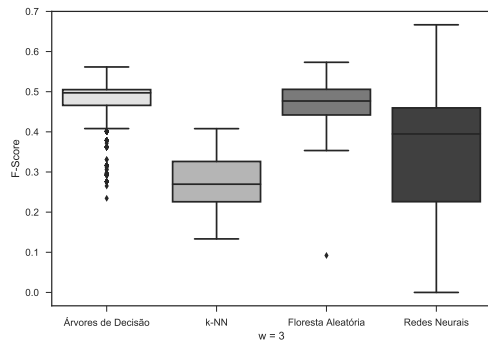
²<http://pandas.pydata.org/>



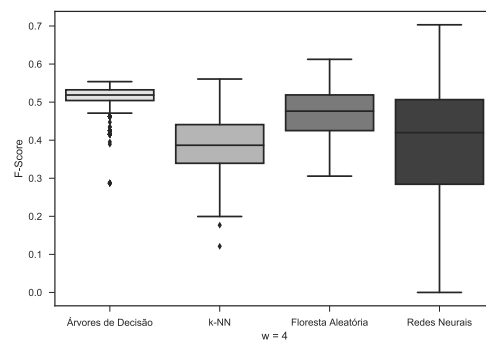
(a) Janela de tempo $w = 1$.



(b) Janela de tempo $w = 2$.



(c) Janela de tempo $w = 3$.



(d) Janela de tempo $w = 4$.

Figura 2: Boxplot da medida F -score das 4 técnicas de AM para as quatro janelas de tempo.

Tabela 2: Melhores resultados para as 4 técnicas de AM baseados em Acurácia, Precisão, Revocação e F -score.

Técnica	Janela de Tempo	Acurácia	Precisão	Revocação	F -score
Árvores de Decisão	$w = 1$	0.4012	0.4339	0.6667	0.5241
	$w = 2$	0.5333	0.5238	0.6303	0.5676
	$w = 3$	0.4784	0.4778	0.6818	0.5616
	$w = 4$	0.5212	0.5333	0.5818	0.5538
Redes Neurais Artificiais	$w = 1$	0.4921	0.1667	0.6667	0.6944
	$w = 2$	0.4212	0.4111	0.6970	0.6743
	$w = 3$	0.4921	0.1587	0.5576	0.6667
	$w = 4$	0.5	0.3125	0.4091	0.7032
K-vizinhos mais próximos	$w = 1$	0.3860	0.4230	0.5152	0.4587
	$w = 2$	0.5606	0.5472	0.6455	0.5694
	$w = 3$	0.5253	0.6807	0.4091	0.4079
	$w = 4$	0.6152	0.7111	0.5091	0.5607
Florestas Aleatórias	$w = 1$	0.3219	0.3882	0.4242	0.5256
	$w = 2$	0.5167	0.5333	0.4758	0.5278
	$w = 3$	0.5579	0.4739	0.4182	0.5732
	$w = 4$	0.4848	0.4996	0.4515	0.6124

Tabela 3: Matriz de confusão para as previsões do conjunto de testes da rede neural (245, 11, 5, 1). No sentido vertical encontram-se os valores esperados e no sentido horizontal encontram-se os valores obtidos pela rede.

	Positivo	Negativo
Positivo	7	3
Negativo	5	4

7. Considerações Finais

Neste trabalho foi ilustrada a utilização e comparação de diferentes técnicas de Aprendizagem de Máquina ao problema de previsão da anomalia anual de precipitações da cidade de Manaus, Amazonas. O intuito desta previsão foi descobrir antecipadamente características do regime de chuva anual, se o volume deste seria maior ou menor que a média histórica para esta localidade. Nesta tarefa de previsão foram utilizados dados reais coletados por estações meteorológicas em Manaus.

Embora alguns trabalhos da literatura já tivessem endereçado a previsão de variáveis meteorológicas de Manaus com redes neurais, a previsão a longo prazo, de caráter anual, ainda não havia sido contemplada. Além disso, não havia análise comparativa com outros métodos e técnicas do corpo de conhecimento da Aprendizagem de Máquina que também podiam endereçar esta lacuna. Assim, este trabalho considerou estas duas perspectivas.

Como resultado, foram propostos e treinados mais de 3800 modelos diferentes, contemplando as técnicas de árvore de decisão, floresta aleatória, redes neurais e k -vizinhos mais próximos, com diferentes parâmetros e hiperparâmetros. Também foi considerada a organização dos dados segundo séries temporais com quatro janelas anuais de tempo.

Os resultados obtidos mostraram que, embora todos os modelos contemplados pudessem endereçar o problema de previsão, as melhores métricas de desempenho foram percebidas nas redes neurais artificiais. Em particular, o modelo com maior F -score utilizou uma janela de tempo de quatro anos e duas camadas ocultas em sua arquitetura. Estes resultados colaboram para uma melhor previsão do regime anual de precipitações em Manaus, podendo ser utilizados como referência por outros pesquisadores e também por instituições públicas (Defesa Civil, Bombeiros, etc.) para tomadas de decisão junto à população, seja de caráter informativo ou preventivo.

Embora diferentes técnicas tenham sido consideradas, os resultados obtidos pelas mesmas não superaram aqueles verificados nas redes neurais artificiais. Esta constatação corrobora resultados de outros trabalhos e reforça a importância deste modelo na realização desta tarefa não-trivial.

Em trabalhos futuros, almeja-se obter mais exemplos acerca da precipitação de Manaus, visando melhoria no treinamento dos modelos já consolidados. Além disso, outras técnicas do ferramental disponível da Aprendizagem de Máquina ainda podem ser exploradas, a exemplo das técnicas paramétricas (regressão polinomial, por exemplo) e de outras técnicas não-paramétricas (redes bayesianas, máquinas de vetores de suporte, etc.).

Agradecimentos

As autoras agradecem o suporte da Universidade do Estado do Amazonas (UEA) e o apoio financeiro provido pela Fundação de Amparo à Pesquisa do Estado do Amazonas

(FAPEAM). A autora Rafaela dos Santos Sousa é bolsista do Programa de Apoio à Iniciação Científica da UEA e FAPEAM edição 2017 – 2018.

Referências

- Alpaydin, E. (2010). *Introduction to Machine Learning*. The MIT Press, 2 edition.
- Bushara, N. O. and Abraham, A. (2013). Computational intelligence in weather forecasting: A review. *Journal of Network and Innovative Computing*, 1:320–331.
- da Silva, D. A. (2012). Função da precipitação no conforto do clima urbano da cidade de Manaus. *Revista Geonorte*, pages 22–40.
- Darji, M. P., Dabhi, V. K., and Prajapati, H. B. (2015). Rainfall forecasting using neural network: A survey. In *International Conference on Advances in Computer Engineering and Applications*, India. IMS Engineering College.
- de Araújo, N. P., Guedes, E. B., de Oliveira, M. B. L., and Andreoli, R. V. S. (2017). Previsão do volume mensal de precipitações em Manaus, Amazonas com redes neurais artificiais. In *IV Escola Regional de Informática Norte I*, pages 1–9, Manaus, Amazonas.
- dos Santos Sousa, R., Silva, L., Guedes, E. B., and de Oliveira, M. B. L. (2017). Redes neurais artificiais aplicadas à previsão antecipada de precipitações na região central de Manaus. In *IV Escola Regional de Informática Norte I*, pages 1–9, Manaus, Amazonas.
- Faceli, K., Lorena, A. C., Gama, J., and de Carvalho, A. C. P. L. F. (2015). *Inteligência Artificial – Uma abordagem de aprendizado de máquina*. Editora LTC, Rio de Janeiro.
- Fisch, G., Marengo, J. A., and Nobre, C. A. (1998). Uma revisão geral sobre o clima da Amazônia. *Acta Amaz.*, 28(2):101.
- Guedes, E. B., de Lima, P. M., and de Oliveira, M. B. L. (2016). Neural networks for time series rainfall forecasting: A case study in Manaus, Amazonas. In *Encontro Nacional de Inteligência Artificial e Computacional*, pages 337–348, Recife, Pernambuco.
- Haykin, S. (2007). *Redes neurais: princípios e prática*. Bookman Editora.
- Haykin, S. (2009). *Neural Networks and Learning Machines*. Pearson, Nova Jersey, 3 edition.
- INMET (2018). Instituto Nacional de Meteorologia. <http://www.inmet.gov.br>. Acessado em 24 de abril de 2018.
- Luk, K. C., Ball, J. E., and Sharma, A. (2001). An application of artificial neural networks for rainfall forecasting. *Mathematical and Computing Modelling*, 33:683–693.
- Marsland, S. (2015). *Machine Learning – An algorithmic perspective*. CRC Press, Estados Unidos, 2 edition.
- Nayak, D. R., Mahapatra, A., and Mishra, P. (2013). A survey on rainfall prediction using artificial neural network. *International Journal of Computer Applications*, 72(16):32–40.
- Witten, I. H., Frank, E., and Hall, M. A. (2011). *Data mining: practical machine learning tools and techniques*. Elsevier, 3 edition.