

RAFAELA DOS SANTOS SOUSA

**PREVISÃO ANUAL DE PRECIPITAÇÕES EM MANAUS, AMAZONAS: UM
COMPARATIVO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA**

Trabalho de Conclusão de Curso apresentado
à banca avaliadora do Curso de Engenharia
de Computação, da Escola Superior de
Tecnologia, da Universidade do Estado do
Amazonas, como pré-requisito para obtenção
do título de Engenheiro de Computação.

Orientadora: Profa. Dra. Elloá B. Guedes

Co-Orientadora: Profa. Dra. Maria Betânia Leal de Oliveira

Manaus – Dezembro – 2019

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).
Sistema Integrado de Bibliotecas da Universidade do Estado do Amazonas.

725p Sousa, Rafaela dos Santos
Previsão Anual de Precipitações em Manaus, Amazonas:
Um comparativo de técnicas de Aprendizado de
Máquina / Rafaela dos Santos Sousa. Manaus : [s.n],
2019.
16 f.: il.; 30 cm.

TCC - Graduação em Engenharia de Computação -
Universidade do Estado do Amazonas, Manaus, 2019.
Inclui bibliografia
Orientador: Guedes, Elloá B.
Coorientador: Oliveira, Maria Betânia Leal de

1. Aprendizado de Máquina. 2. Previsão. 3.
Precipitações. 4. Manaus. I. Guedes, Elloá B. (Orient.). II.
Oliveira, Maria Betânia Leal de (Coorient.). III.
Universidade do Estado do Amazonas. IV. Previsão Anual
de Precipitações em Manaus, Amazonas: Um comparativo
de técnicas de Aprendizado de Máquina

Elaborado por Jeane Macelino Galves - CRB-11/463

RAFAELA DOS SANTOS SOUSA

**PREVISÃO ANUAL DE PRECIPITAÇÕES EM MANAUS, AMAZONAS: UM
COMPARATIVO DE TÉCNICAS DE APRENDIZADO DE MÁQUINA**

Trabalho de Conclusão de Curso apresentado à banca
avaliadora do Curso de Engenharia de Computação, da
Escola Superior de Tecnologia, da Universidade do Es-
tado do Amazonas, como pré-requisito para obtenção do
título de Engenheiro de Computação.

Aprovado em: 06/12/2019

BANCA EXAMINADORA



Profa. Dra. Elloá B. Guedes

UNIVERSIDADE DO ESTADO DO AMAZONAS



Profa. Dra. Maria Betânia Leal de Oliveira

UNIVERSIDADE DO ESTADO DO AMAZONAS



Profa. M.Sc. Márcia Sampaio Lima

UNIVERSIDADE DO ESTADO DO AMAZONAS



Prof. M.Sc. Mário Augusto Bessa de Figueiredo

UNIVERSIDADE DO ESTADO DO AMAZONAS

Coragem é fazer uma coisa mesmo estando derrotado antes de começar – prosseguiu Articus.
– E mesmo assim ir até o fim, apesar de tudo. Você raramente vai vencer, mas às vezes vai conseguir..

Harper Lee

O Sol é para Todos

Agradecimentos

À minha mãe, Margarida dos Santos Sousa, que sempre me incetivou e me inspirou durante toda a minha jornada, seja ela acadêmica quanto pessoal. Agradeço por todos os sacrifícios realizados em nome da minha educação e bem estar. Obrigada por ser meu porto seguro, obrigada por ser forte, por você e por mim, e não me deixar desistir.

Aos meus familiares, por celebrarem as minhas conquistas junto comigo. Em especial à minha irmã, Diana, e meu pai, José, que por mais durões que sejam, sempre me apoiaram e estiveram ao meu lado.

Aos meus amigos, Janderson, Luiz, Giovana, Marcos, Nicoli, Liliana, Rodrigo Souza, Rodrigo Moraes, Lucas, Emanuel, Paulo Roberto e Leticia, pelo companheirismo constante durante a graduação (e fora dela).

À minha orientadora, Elloá Guedes, por todo conhecimento técnico, paciência e compreensão. Agradeço todos os dias por ter tido a oportunidade de ser sua aluna e poder ter visto de perto o seu amor pela educação.

Ao Bruno, que me deu forças em um dos momentos mais difíceis da minha vida. Obrigada pela paciência, pelo carinho e por deixar os meus dias mais leves. Obrigada por me entender, me conhecer e, ainda assim, me amar.

Ao membros da comunidade PyLadies e PyData, principalmente Allex, Juliany e Jailson, por me ensinarem que conhecimento é tudo e que compartilhar é melhor ainda. Vocês mudaram a minha vida.

Agradeço também à Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) que, por meio dos Projetos PROTI Pesquisa e PPP 04/2017, colaborou para a consolidação da infraestrutura física e tecnológica do Laboratório de Sistemas Inteligentes da Escola Superior de Tecnologia da Universidade do Estado do Amazonas. Este trabalho de conclusão de curso é um dos produtos destes projetos, pois foi desenvolvido no referido laboratório, fez uso dos recursos computacionais ali disponíveis e foi melhorado graças às discussões e interações com o grupo de pesquisa nele sediado.

Previsão Anual de Precipitações em Manaus, Amazonas: Um comparativo de técnicas de Aprendizado de Máquina

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, ebgsocosta, mloliveira}@uea.edu.br

Abstract. *Rainfall in Manaus, Amazonas results from a nontrivial combination of several variables and affects the population in many aspects, such as in economics and health. For these reasons, this work aimed at forecasting the annual precipitation volume in this city. Considering the Machine Learning perspective, we proposed more than 4000 distinct models from which it was possible to highlight a neural network with F-score higher than 0.7 as the most appropriated model for this task. The results obtained can be used by authorities to mitigate damages caused by rainfall, for a better management of water resources and to plan rainfall dependent activities, such as agriculture.*

Resumo. *As precipitações em Manaus, Amazonas decorrem de uma combinação não trivial de diversas variáveis e afetam a população em muitos aspectos, como de natureza econômica e de saúde. Por essas razões, este trabalho teve como objetivo prever o volume anual de precipitações nesta cidade. Considerando a perspectiva do Aprendizado de Máquina, foram propostos mais de 4000 modelos diversos, a partir dos quais foi possível eleger uma rede neural que obteve F-score superior a 0.7 como modelo apropriado. Os resultados obtidos podem ser utilizados por autoridades para minimização dos danos causados pelas chuvas, para um melhor manejo dos recursos hídricos e para planejamento de atividades dependentes da precipitação, a exemplo da agricultura.*

1. Introdução

Manaus, capital do Estado do Amazonas e sétima cidade mais populosa do Brasil, tem como uma de suas características particulares o clima, de característica equatorial quente e úmido, naturalmente desconfortável [da Silva 2009]. De acordo com a classificação de Köppen, este clima é classificado como sendo *tropical de monções*, em que a chuva é abundante, mas não uniforme. As temperaturas são consistentemente altas durante todo o ano, com intensas precipitações de Outubro a Junho e com Agosto como mês mais seco, com menos de 60 mm de precipitação [Alavares et al. 2014].

As chuvas em Manaus influenciam a vida dos habitantes desta cidade de diversas formas. Essa influência se dá, por exemplo, em termos sociais e econômicos para as famílias que praticam a agricultura de subsistência, fortemente dependente do regime de precipitações [Adams et al. 2005]. Em termos de saúde, quando o volume de precipitações é elevado, a população sofre com agravos de veiculação hídrica, como leptospirose e hepatites virais, doenças diarreicas e doenças transmitidas por vetores, que tendem a sofrer um aumento [Silva et al. 2010]. De maneira bastante abrangente, a precipitação influencia em aspectos relacionados ao conforto térmico da população, dadas as características típicas desta região equatorial [Sioli 1991, da Silva 2012]. Por essas razões,

pesquisadores acreditam que Manaus possui uma combinação particular de ingredientes sociais, geográficos e de clima que por si só configuram um quadro social de atenção, especialmente considerando a relação de seus moradores e a água que banha a cidade [Silva et al. 2010].

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 [da Silva 2012]. Além desses sistemas, somam-se a localização geográfica, topografia e também as mudanças climáticas, que têm alterado a ocorrência e a severidade de eventos hidroclimatológicos [Stocker et al. 2013]. Assim, dadas essas características, pode-se verificar que prever a ocorrência de precipitações em Manaus não é uma tarefa simples.

Apesar das dificuldades, endereçar a previsão da precipitação em Manaus pode colaborar para o planejamento de atividades ligadas à Agricultura e Pecuária, para a minimização de danos às populações em áreas de riscos, para o manejo apropriado de recursos hídricos, para a tomada de decisões estratégicas pelos setores governamentais de caráter ambiental, saúde pública, infraestrutura, dentre outros.

Embora a abordagem mais amplamente difundida para previsão de precipitações leve em conta a modelagem das leis físicas que regem estes eventos, é possível endereçar esta questão por meio de modelos computacionais que reconhecem padrões nos dados das séries históricas de precipitações, como proposto pela área de Aprendizado de Máquina. Segundo essa abordagem, 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, portanto, a abordagem proposta pelo Aprendizado de Máquina, este trabalho investigou e comparou diferentes técnicas para prever o volume anual de chuva na cidade em questão. Foram construídos, treinados e testados mais de 4000 modelos com parâmetros distintos, respeitando as práticas de validação cruzada e utilizando janelas de tempo, produzindo diferentes métricas de desempenho. Após uma comparação criteriosa, foi verificado que as redes neurais artificiais endereçaram melhor esta tarefa, em que o modelo mais adequado obteve F -score superior a 0.7. Os resultados obtidos consolidam um modelo de previsão de precipitações anual para Manaus e reforçam a adequação das redes neurais nesse tipo de tarefa, conforme também verificado por outros autores da literatura.

Para apresentar os resultados obtidos, este trabalho está organizado como segue. As pesquisas afins disponíveis na literatura encontram-se detalhadas na Seção 2. Os materiais e métodos utilizados para condução das atividades deste trabalho encontram-se descritos na Seção 3. Os resultados obtidos são mostrados e discutidos na Seção 4. Por fim, as considerações finais são listadas na Seção 5.

2. Trabalhos Relacionados

A previsão de precipitações com métodos de Aprendizado de Máquina é um tema importante e considerado em diversas partes do mundo [Darji et al. 2015, Bushara and Abraham 2013, Gardner and Dorling 1998, Nayak et al. 2013]. 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 Aprendizado de Máquina, 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, esse trabalho não considera o volume desses eventos de precipitação.

Na tentativa de superar essas 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 único modelo de Aprendizagem de Máquina, os resultados obtidos colaboraram para a elaboração de modelos de previsão de volume e também mostraram uma boa adequação dos atributos preditores ao domínio do problema considerado.

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.

Dessa 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 para esta tarefa, com o intuito de fornecer um panorama comparativo. Apesar das limitações, aspectos positivos destes trabalhos devem ser destacados, tais como a utilização de Aprendizado de Máquina, de dados observacionais e de índices climáticos como atributos preditores e também a organização destes segundo séries temporais. Dadas as vantagens observadas pelos outros autores, os aspectos positivos observados também serão adotados no escopo deste trabalho.

3. Materiais e Métodos

Esta seção tem por objetivo caracterizar os materiais e métodos utilizados para conduzir as atividades deste trabalho. De maneira geral, esta pesquisa seguiu o fluxo de atividades que a literatura preconiza no tocante ao processo de Aprendizagem de Máquina [Brink et al. 2017].

Em primeiro lugar, o problema considerado foi de *classificação binária* com vistas a prever, com um ano de antecedência, o atributo-alvo denominado *anomalia da precipitação anual*. Este atributo-alvo é obtido a partir do volume anual de precipitação em Manaus da seguinte forma:

$$\text{anomalia}(y_i) = \begin{cases} +1, & \text{se } y_i - \bar{y} \geq 0 \\ -1, & \text{em caso contrário,} \end{cases} \quad (1)$$

em que y representa o volume anual de precipitações, i denota um ano específico e \bar{y} é a média histórica anual do volume de precipitações. 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 (anomalia positiva) ou para menos (anomalia negativa), conforme

ilustrado na Figura 1. É importante observar também que o atributo-alvo é balanceado para as classes consideradas.

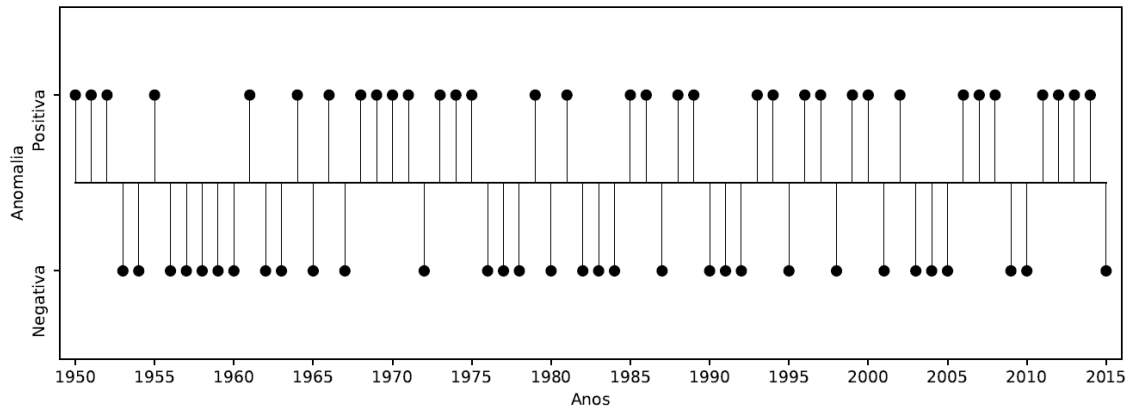


Figura 1: Anomalia da precipitação anual.

Para conduzir esta tarefa de Aprendizagem de Máquina, os dados experimentais, modelos de Aprendizagem de Máquina e métricas de desempenho considerados no escopo deste trabalho encontram-se detalhados nas subseções a seguir.

3.1. Dados Experimentais

Os dados experimentais para a realização deste trabalho são de natureza observacional, coletados a partir de estações meteorológicas automáticas localizadas em Manaus [INMET 2018] e também a partir da temperatura mensal 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. Estes dados, que compreendem os anos de 1950 a 2015, passaram então a compor os atributos preditores da tarefa de classificação. A Tabela 1 apresenta a síntese da estatística descritiva dos dados experimentais.

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

Esses dados foram organizados segundo uma série temporal e defasados segundo janelas de tempo w 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.

3.2. Seleção de Modelos e Ajustes de Parâmetros

Os modelos de Aprendizado de Máquina considerados nesta tarefa foram: árvores de decisão, florestas aleatórias, k -vizinhos mais próximos, máquinas de vetores de suporte,

redes neurais artificiais e regressão logística, por se adequarem à problemas de classificação e serem compatíveis com os tipos dos atributos disponíveis.

Cada modelo, dentre os escolhidos, possui diferentes parâmetros de configuração. Por exemplo, para o k -vizinhos mais próximos, diferentes valores de k podem ser considerados. Já para as redes neurais, por sua vez, é necessário especificar a arquitetura, função de ativação, taxa de aprendizado, número de épocas, dentre outros. Apesar das particularidades de cada modelo, é preciso assegurar que a escolha dos respectivos parâmetros seja otimizada.

Desta maneira, para otimizar os parâmetros dos modelos utilizados, a técnica de *busca em grid* (do inglês, *grid search*) foi considerada, a qual realiza uma busca exaustiva sobre as possíveis combinações de cada um dos parâmetros de cada modelo [Brink et al. 2017]. Isto implicou em um número considerável de modelos a serem treinados e testados, o que aproxima as chances de encontrar a melhor solução.

3.3. Avaliação de Desempenho

Na tarefa de classificação binária levada em consideração, os possíveis resultados são os valores verdadeiros positivos (tp), verdadeiros negativos (tn), falsos positivos (fp) e falsos negativos (fn). Assim, após a execução de cada um dos modelos, foram derivadas as seguintes *métricas de desempenho*: acurácia, precisão, revocação e F -score. Essas medidas são calculadas da seguinte forma:

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

Para avaliar a capacidade de generalização dos modelos propostos, propiciar uma comparação justa e minimizar as chances de escolha de modelos com *overfitting*, o método k -fold de *validação cruzada* foi utilizado para validar os resultados obtidos, aumentando a confiança nos mesmos [Marsland 2015]. Em particular, considerando que o conjunto de dados compreende apenas 65 anos, as restrições advindas dos aspectos temporais dos dados e os custos computacionais envolvidos, foram utilizados 3 *folds* neste processo de validação.

3.4. Tecnologias Utilizadas

Para viabilizar o pré-processamento dos dados, o treinamento e teste dos modelos e a posterior comparação dos resultados obtidos, foi utilizada a linguagem de programação Python¹, por ser flexível, rápida e dispor de várias bibliotecas para Aprendizado de Máquina. Em particular, o *framework* `sci-kit learn`² e a biblioteca `pandas`³ possuíam papel central nos algoritmos elaborados.

4. Resultados e Discussão

Ao conduzir as atividades conforme descrito na seção anterior, a organização dos dados experimentais em quatro janelas de tempo produziu quatro conjunto de dados, com as seguintes características apresentadas na Tabela 2.

¹<http://python.org>

²<http://scikit-learn.org/>

³<http://pandas.pydata.org/>

Tabela 2: Conjuntos de dados derivados a partir do defasamento temporal.

Tamanho da Janela	Atributos Preditores	Exemplos Disponíveis
$w = 1$	62	65
$w = 2$	123	63
$w = 3$	184	62
$w = 4$	245	61

Os parâmetros utilizados variaram conforme o modelo, da seguinte forma:

1. **Árvores de Decisão.** *Critério de divisão:* entropia de Shannon, impureza de Gini; *Profundidade da Árvore:* 1 a 9; *Quantidade máxima de nós folha:* 1 a 9;
2. **Florestas Aleatórias.** *Critério de divisão:* entropia de Shannon, impureza de Gini; *Número de estimadores:* 23 valores distintos;
3. **k -vizinhos Mais Próximos.** *Quantidade de vizinhos:* 30 valores distintos; *Pesos:* baseados em distância ou uniforme;
4. **Redes Neurais.** *Camadas Ocultas:* 1 ou 2; *Quantidade de Neurônios por Camada:* conforme utilização da regra da pirâmide geométrica; *Função de Ativação:* identidade, logística, tangente hiperbólica e retificação linear; *Método de Otimização:* gradiente descendente, L-BFGS ou Adam.
5. **Regressão Logística.** *Penalidade:* L1 e L2; *Método de Otimização:* Para a penalidade L1, liblinear e saga, para a penalidade L2, newton-cg, lbfgs e sag.
6. **Máquinas de Vetores de Suporte.** *Kernel:* linear, RBF e sigmóide; *Gau:* de 3 a 9; *Gamma:* 0.5, 1, 2 e 10.

Considerando esses parâmetros, a utilização de quatro janelas temporais e a busca em *grid*, obtiveram-se 2888 árvores de decisão, 184 florestas aleatórias, 528 redes neurais, 240 valores para o algoritmo de k -vizinhos mais próximos, 16 modelos de regressão logística e 336 de máquinas de vetores de suporte. Cada um destes modelos foi então treinado e testado por três vezes, utilizando cada um dos *folds* disponíveis, conforme recomendado pelo método de validação cruzada adotado.

Após a condução da etapa de testes, as métricas de desempenho foram então obtidas. Organizando os resultados obtidos por janela de tempo e considerando a ordenação por F -score, devido ao fato dessa métrica levar em conta simultaneamente a precisão e a revocação, tem-se os resultados sintetizados na Tabela 3.

A partir dos resultados obtidos, foi possível observar que o maior valor da métrica de F -score observado residiu nas redes neurais, especialmente para um modelo específico que considerou a janela de tempo $w = 2$. Porém, analisando conjuntamente todas as redes construídas, treinadas e testadas para esta tarefa, foi possível verificar um F -score médio de 0.4351 com desvio padrão igual a 0.11. Isso denota que o comportamento das redes neurais não foi homogêneo para a tarefa de previsão considerada, que não é possível afirmar que as redes neurais são superiores aos demais modelos de Aprendizagem de Máquina para a precipitação em Manaus e que a busca em *grid* foi particularmente importante na otimização dos parâmetros, permitindo identificar um modelo com bom desempenho.

A rede neural com melhor desempenho, verificado para $w = 2$, possui arquitetura (123, 1, 4, 1), isto é, possui duas camadas ocultas com 1 e 4 neurônios, respectivamente.

Tabela 3: Melhores resultados para as 6 técnicas de Aprendizagem de Máquina, ordenados por F-score.

Técnica	Janela de Tempo	Acurácia	Precisão	Revocação	F-score
Árvores de Decisão	$w = 1$	0.4012	0.4339	0.6667	0.5241
	$w = 2$	0.4742	0.4361	0.5394	0.5643
	$w = 3$	0.4451	0.4889	0.5152	0.5293
	$w = 4$	0.5182	0.4902	0.6424	0.5701
Florestas Aleatórias	$w = 1$	0.4755	0.4629	0.5152	0.5422
	$w = 2$	0.5333	0.4861	0.4697	0.6224
	$w = 3$	0.3959	0.4815	0.6515	0.5946
	$w = 4$	0.4212	0.4923	0.6121	0.5771
k -vizinhos mais próximos	$w = 1$	0.4452	0.4635	0.5455	0.4974
	$w = 2$	0.5318	0.5579	0.5030	0.5140
	$w = 3$	0.5745	0.5743	0.5818	0.5741
	$w = 4$	0.5621	0.5640	0.6121	0.5820
Máquinas de Vetores de Suporte	$w = 1$	0.5079	0.5079	1.0	0.6736
	$w = 2$	0.5	0.5	1.0	0.6667
	$w = 3$	0.6026	0.5804	0.7455	0.6477
	$w = 4$	0.5	0.5	1.0	0.6667
Redes Neurais Artificiais	$w = 1$	0.5079	0.5334	0.3333	0.7063
	$w = 2$	0.3712	0.3889	0.7485	0.7396
	$w = 3$	0.5714	0.5556	0.1212	0.7247
	$w = 4$	0.5439	0.4111	0.5455	0.7190
Regressão Logística	$w = 1$	0.3528	0.3968	0.4242	0.3997
	$w = 2$	0.4545	0.4524	0.5333	0.4970
	$w = 3$	0.5102	0.5099	0.5485	0.5259
	$w = 4$	0.4409	0.4762	0.4848	0.4615

Esta rede neural usa a função de ativação identidade e método de otimização Adam, um otimizador estocástico. A Tabela 4 mostra a matriz de confusão produzida por esta rede neural para um dos conjunto de testes, com valores percentuais.

Tabela 4: Matriz de confusão para as previsões do conjunto de testes da rede neural (123, 1, 4, 1). No sentido vertical encontram-se os valores esperados e no sentido horizontal encontram-se os valores obtidos pela rede.

		Resultados Observados		
		Positiva	Negativa	Total
Resultados Previstos	Positiva	0.2	0.15	0.35
	Negativa	0.25	0.4	0.65
Total		0.45	0.55	

Algumas características particulares das redes neurais podem estar mais diretamente relacionadas ao resultado obtido, em especial a boa captura das relações não lineares, a resistência ao ruído, e a capacidade de serem aproximadoras universais de qualquer função. Estes resultados corroboram a percepção da literatura no tocante à qualidade das redes neurais como preditoras de eventos de precipitação em diversas localidades no mundo [Darji et al. 2015, Bushara and Abraham 2013, Gardner and Dorling 1998, Nayak et al. 2013].

Embora alguns trabalhos tenham feito previsões de precipitação para a cidade de Manaus considerando meia hora [dos Santos Sousa et al. 2017], um dia [Guedes et al. 2016] ou um mês à frente [de Araújo et al. 2017], não foram encontrados registros com previsões de mais longo prazo para a localidade considerada, tal como abordado neste trabalho. Isso ressalta a importância dos resultados obtidos e motiva a realização de novas pesquisas, em que os resultados aqui apresentados podem atuar como base para comparações. 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.

5. Considerações Finais

Neste trabalho foram ilustradas a utilização e a comparação de diferentes técnicas de Aprendizagem de Máquina aplicadas ao problema de previsão da anomalia anual de precipitações da cidade de Manaus, Amazonas. O intuito dessa previsão foi descobrir *a priori* características do regime de chuva anual. Nesta tarefa de previsão foram utilizados dados reais coletados por estações meteorológicas em Manaus e também índices meteorológicos amplamente conhecidos.

Considerando diferentes técnicas e parâmetros, foram então propostos, treinados e testados mais de 4000 modelos distintos, contemplando as técnicas de árvores de decisão, florestas aleatórias, k -vizinhos mais próximos, máquinas de vetores de suporte, redes neurais artificiais e regressão logística. 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 classificação considerado, uma rede neural com duas camadas ocultas e que utilizou dados de dois anos anteriores obteve melhor desempenho. Por utilizar dados reais e ter sido acurado e validado com técnicas apropriadas, este modelo pode entrar em produção junto a órgãos como Defesa Civil, Corpo de Bombeiros, Prefeitura etc., para tomadas de decisão junto à população, seja de caráter informativo ou preventivo, visando minimizar os danos causados pelas precipitações em Manaus.

Em trabalhos futuros, almeja-se obter mais dados históricos acerca da precipitação de Manaus, com vistas a melhoria dos modelos elaborados e também a utilização de outras técnicas.

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

- Adams, C., Murrieta, R. S. S., and Sanches, R. A. (2005). Agricultura e alimentação em populações ribeirinhas das várzeas do Amazonas: novas perspectivas. *Ambiente & Sociedade*, 8(1):1–22.
- Alavares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., and Sparovek, G. (2014). Koppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6):711–728.
- Brink, H., Richards, J. W., and Fetherolf, M. (2017). *Real-World Machine Learning*. Manning Publications, Estados Unidos.
- 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. (2009). *A influência das áreas verdes no clima da cidade de Manaus*. PhD thesis, Universidade Federal do Amazonas.
- 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.
- Fisch, G., Marengo, J. A., and Nobre, C. A. (1998). Uma revisão geral sobre o clima da Amazônia. *Acta Amaz.*, 28(2):101.
- Gardner, M. W. and Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14/15):2627–2636.
- 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.
- INMET (2018). Instituto Nacional de Meteorologia. <http://www.inmet.gov.br>. Acessado em 3 de maio 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.

- Silva, D. X., Barcellos, C., and Bacuri, R. (2010). Diagnóstico preliminar da vulnerabilidade e as previsões dos efeitos das mudanças climáticas na saúde pública no município de Manaus. Technical report, Fundação Oswaldo Cruz.
- Sioli, H. (1991). *Amazônia: Fundamentos da ecologia da maior região de florestas tropicais*. Vozes, Rio de Janeiro.
- Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. (2013). Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change. Cambridge Univ. Press.