UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS GRADUAÇÃO EM FÍSICA AMBIENTAL

"Modelagem da temperatura do ar em transecto urbano utilizando LSTM"

César Eduardo Guarienti

Orientador: **Prof. Dr. Raphael de Souza Rosa Gomes**
Coorientador: **Prof. Dr. Josiel Maimone de Figueiredo**

Cuiabá - MT
Fevereiro/2020

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS GRADUAÇÃO EM FÍSICA AMBIENTAL

Modelagem da temperatura do ar em transecto urbano utilizando LSTM

César Eduardo Guarienti

Tese apresentada ao Programa de Pós Graduação em Física Ambiental da Universidade Federal de Mato Grosso, como parte dos requisitos para obtenção do título de Doutor em Física Ambiental.

Prof. Dr. Raphael de Souza Rosa Gomes

Prof. Dr. Josiel Maimone de Figueiredo

Cuiabá, MT

Fevereiro/2020

DEDICATÓRIA

Aos meus pais, Osmar e Maria, à minha esposa Gracyeli, à nossa filha Marieli, aos meus sobrinhos Leandro e Rafaela, à minha irmã Rosângela.

Dados Internacionais de Catalogação na Fonte.

G915m Guarienti, César Eduardo.
Modelagem da temperatura do ar em transecto urbano utilizando LSTM / César Eduardo Guarienti. -- 2020
107 f. : il. color. ; 30 cm.

Orientador: Raphael de Souza Rosa Gomes.
Co-orientador: Josiel Maimone de Figueiredo.
Tese (doutorado) - Universidade Federal de Mato Grosso, Instituto de Física, Programa de Pós-Graduação em Física Ambiental, Cuiabá, 2020.
Inclui bibliografia.

1. dados ambientais. 2. predição de temperatura. 3. redes neurais recorrentes. 4. deep learning. 5. séries espaço-temporais. I. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
Programa de Pós-Graduação em Física Ambiental

FOLHA DE APROVAÇÃO

TÍTULO: **MODELAGEM DA TEMPERATURA DO AR EM TRANSECTO URBANO UTILIZANDO LSTM**

AUTOR: **CÉSAR EDUARDO GUARIENTI**

Tese de Doutorado defendida e aprovada em 13 de fevereiro de 2020, pela comissão julgadora:

Prof. Dr. Raphael de Souza Rosa Gomes
Orientador
Instituto de Computação – UFMT

Prof. Dr. Josiel Maimone de Figueiredo
Coorientador
Instituto de Computação – UFMT

Profa. Dra. Andréia Gentil Bonfante
Examinadora Externa
Instituto de Computação – UFMT

Prof. Dr. Carlo Ralph De Musis
Examinador Interno
Universidade de Cuiabá - UNIC

Prof. Dr. Clodoaldo Nunes
Examinador Externo
Instituto Federal de Mato Grosso - IFMT

Prof. Dr. Jonathan Willian Zangeski Novais
Examinador Externo
Universidade de Cuiabá – UNIC/Cuiabá

Agradecimentos

A Deus pelo dom da vida e por ser meu amigo de todas as horas.

Aos meus pais eu agradeço pelo apoio irrestrito em todos os momentos, pela doação, por cada luta enfrentada, pelas lições de vida e pelo exemplo. Ser pai me fez entender melhor o quanto se esforçaram por mim e minha irmã. Obrigado!

A minha esposa, pelo apoio, carinho e cuidado, que fez e faz diferença na minha vida. Obrigado pelo zelo com a nossa filha nesta ausência minha.

A minha filha amada: você é o maior presente que jamais mereci na vida. Deus foi muito generoso em me presentear como seu pai. É uma grande honra. O seu sorriso e a sua alegria trazem água ao meu deserto. Você é luz no meu caminho. Papai ama demais.

A minha irmã Rosângela e ao cunhado Thiago Ventura por todo apoio.

Ao meu sobrinho querido, Leandro. Você é o "gut gut" do titio. Foi um prazer acompanhar seu crescimento. Foi um privilégio cada momento.

A minha sobrinha Rafaela (Fifinha). Você é a gatinha do Dedé.

Aos meus padrinhos pelo apoio pela vida inteira.

Ao orientador Raphael: muito obrigado pela parceria!

Ao co-orientador Josiel por toda ajuda desde minha chegada ao IC!

A Diana Paula pela grande parceria neste estudo.

Aos alunos Lucas e Valfride pelo apoio nas nossas pesquisas.

Ao Paraná, Marta, Cesário e Soilce pelo gentil acolhimento no PPGFA.

Aos Professores da banca pela gentileza de contribuir com este estudo.

Aos meus professores pela enorme contribuição na minha caminhada.

A todos que, direta ou indiretamente, me ajudaram.

A UFMT pela oportunidade de estudar.

"Não importa onde você parou... Em que momento da vida você cansou... O que importa é que sempre é possível recomeçar. Recomeçar é dar uma nova chance a si mesmo... É renovar as esperanças na vida e, o mais importante... Acreditar em você de novo."

Carlos Drummond de Andrade

SUMÁRIO

LISTA DE FIGURAS **I**

LISTA DE TABELAS **IV**

LISTA DE ABREVIATURAS **V**

RESUMO **VII**

ABSTRACT **VIII**

1 Introdução **1**

1.1 Objetivo Geral 3

1.2 Objetivos Específicos 3

2 Fundamentação Teórica **4**

2.1 Conforto Ambiental 4

2.1.1 Clima urbano 4

2.1.1.1 Determinantes do clima urbano 6

2.1.1.1.1 Temperatura do ar 6

2.1.1.1.2 Umidade relativa do ar 6

2.1.1.1.3 Regime de vento 7

2.1.1.1.4 Precipitação 7

2.1.2 Anomalia térmica 7

2.1.2.1 Classificação e influência do revestimento de solo 8

2.1.2.2 Cobertura de solo 9

2.2 Redes Neurais Artificiais 10

2.2.1 Inteligência Artificial 10
2.2.1.1 Machine Learning 10
2.2.1.1.1 Funções de Ativação 16
2.2.1.1.2 Otimizadores 19
2.2.1.1.3 Loss Function 19
2.2.1.1.4 Redes Neurais Recorrentes 21
2.2.1.2 Deep Learning 22
2.2.1.3 Long-Short Term Memory 24
2.3 Estatística 30

3 Materiais e Métodos **33**
3.1 Dados Utilizados 33
3.1.1 Sensores 33
3.1.2 Localização 34
3.1.3 Período 35
3.1.4 Classificação da cobertura de solo 36
3.2 Desenvolvimento do Modelo Deep Learning 36
3.2.1 Ambiente de Desenvolvimento 36
3.2.2 Escolha da Rede Neural 37
3.2.3 Seleção das Variáveis 38
3.2.4 Preparação dos Dados 38
3.2.5 Amostragem 41
3.2.6 Open Science 42
3.2.7 Bibliotecas Python 42
3.2.8 Criação do Modelo LSTM 43
3.2.9 Modelos criados 45

4 Apresentação e Análise dos Resultados **47**
4.1 2011 prever 2016 47
4.1.1 Amostragem 90% de treino e 10% de teste 47
4.1.1.1 Modelo Inverno 47
4.1.1.2 Aplicação do melhor modelo de Inverno 49

4.1.1.3 Modelo Outono 50
4.1.1.4 Aplicação do melhor modelo de Outono 52
4.1.1.5 Modelo Primavera 52
4.1.1.6 Aplicação do melhor modelo Primavera 54
4.1.1.7 Modelo Verão 55
4.1.1.8 Aplicação do melhor modelo Verão 56
4.1.1.9 Tabela do Coeficiente de Determinação 57
4.2 2016 prever 2016 57
4.2.1 Amostragem 90% de treino e 10% de teste 58
4.2.1.1 Modelo Inverno 58
4.2.1.2 Modelo Outono 59
4.2.1.3 Modelo Primavera 61
4.2.1.4 Modelo Verão 63
4.2.1.5 Tabela do Coeficiente de Determinação 65
4.3 Previsão 2021 66
4.3.1 Inverno 2021 66
4.3.2 Outono 2021 67
4.3.3 Primavera 2021 68
4.3.4 Verão 2021 69
4.4 Dados de 2011, 2016 e previsão 2021 70
4.4.1 Comparativo Inverno 2011, 2016, 2021 70
4.4.2 Comparativo Outono 2011, 2016, 2021 71
4.4.3 Comparativo Primavera 2011, 2016, 2021 72
4.4.4 Comparativo Verão 2011, 2016, 2021 72

5 Considerações Finais **74**
5.1 Contribuições 75
5.2 Trabalhos Futuros 75

REFERÊNCIAS **77**

LISTA DE FIGURAS

1 Neurônios com suas principais partes. Fonte: Ferneda (2006). . . . 13
2 Neurônios dispostos em duas camadas. 14
3 Neurônios dispostos em camadas de entrada, oculta e saída. . . . 14
4 Funcionamento do neurônio artificial. 15
5 Função linear. 17
6 Função sigmoid. 17
7 Função Tangente. 18
8 Função ReLU. 18
9 Exemplo de Rede Neural Recorrente. 22
10 Representação do modelo Stacked LTSM. 25
11 Representação dos três gates LTSM. 26
12 Linha com estado da célula. 26
13 *Primeiro gate LSTM: forget gate.* 26
14 *Segundo gate LSTM: input gate.* 27
15 *Terceiro gate LSTM: output gate.* 27
16 Modelo Vanilla LTSM. 28
17 Modelo Stacked LTSM. Fonte: Elaborado pelo próprio autor. . . . 29
18 Bidirectional LTSM. Fonte: Elaborado pelo próprio autor. 29

19 Sensor utilizado no transecto. Adaptado de Zavitoski, 2018. . . . 34
20 Área de estudo com 91 pontos em Cuiabá-MT. 35
21 Conjunto de etapas na manipulação do banco de dados. 39
22 Fluxo de etapas. Fonte: Adaptado de Ferrari et al. 2019. 44
23 Diagrama com formato do modelo de entrada do Keras. 45

24 Gráfico com melhor modelo de inverno 2011 para prever inverno 2016. 48
25 Aplicação do melhor modelo de inverno 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão. 49
26 Gráfico com melhor modelo de outono. 51
27 Aplicação do melhor modelo de outono 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão. 52
28 Gráfico com melhor modelo da primavera. 53
29 Aplicação do melhor modelo de primavera 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão. 54
30 Gráfico com melhor modelo de verão. 55
31 Aplicação do melhor modelo de verão 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão. 56
32 Modelo inverno 2016 aplicado em dados de inverno 2016. 58
33 Aplicação do melhor modelo de inverno 2016 na previsão das outras estações do mesmo ano. 59
34 Modelo outono 2016. 60
35 Aplicação do melhor modelo de outono 2016 no conjunto inteiro de dados de inverno, outono, primavera, verão. 61
36 Modelo primavera 2016... 62
37 Aplicação do melhor modelo da primavera de 2016 nos dados das outras estações do mesmo ano. 63
38 Modelo verão 2016. 64
39 Aplicação do melhor modelo de verão 2016 no conjunto inteiro de dados das outras estações de 2016 65
40 Previsão de temperatura para o ano de 2021. 67
41 Previsão de temperatura para outono de 2021. 68
42 Previsão de temperatura para primavera de 2021. 69
43 Previsão de temperatura para verão de 2021. 70

44 Comparativo da temperatura de inverno. 71
45 Comparativo da temperatura de outono. 71
46 Comparativo da temperatura da primavera. 72
47 Comparativo da temperatura do verão. 73

48 Etapas da criação do modelo criado de Rede Neural do tipo LSTM. 90

LISTA DE TABELAS

1 Escalas espaciais. 6

2 Período de medição por transectos móveis em 2011/2012. 35

3 Período de medição por transectos móveis em 2016. 36

4 Grupo de variáveis de um determinado ponto. 40

5 Quatro conjuntos de variáveis que servem para dar entrada na Rede Neural. Exemplo de um *timestamp*. 41

6 Exemplo de combinação dos parâmetros de treinamento da Rede Neural 44

7 Valores de R^2 com dados de 2011 para prever 2016 com 90% dos dados para treinamento. 57

8 Valores de R^2 com dados de 2016 para prever outras estações de 2016 com 90% dos dados para treinamento. 66

LISTA DE ABREVIATURAS

ADALINE	Adaptive Linear Elemtent	12
ARIMA	Autoregressive Integrated Moving Average	2
BPMN	Business Process Modeling Notation	36
CNN	Convolutional Neural NetworkAverage	27
CPU	Center Process Unit	43
CSV	Comma Separated Value	36
DL	Deep Learning	22
IA	Inteligência Artificial	10
IDE	Integrated Development Environment	32
KNN	K-nearest Neighbor	23
LDA	Linear Discriminant Analysis	23
LSTM	Long Short-Term Memory	3
MAE	Mean Absolute Error	19
ML	Machine Learning	10
MLP	Multilayer Perceptron	14
MLPR	Multilayer Perceptron Regressor	14
MSE	Mean Squared Error	19

PPGFA Programa de Pós-Gradução em Física Ambiental 30

ReLU Rectied Linear Units 13

RMSE Root Mean Squared Error 19

RN Redes Neurais 2

RNA Redes Neurais Artificiais 2

RNN Redes Neurais Recorrentes 21

SARIMA Seasonal Autoregressive Integrated Moving Average 2

SMO Sequential Minimal Optimization 23

SQL Structured Query Language 35

UR Umidade Relativa 1

UTM Universal Tranversa de Mercator 31

RESUMO

GUARIENTI, C. E. Modelagem da temperatura do ar em transecto urbano utilizando LSTM. Cuiabá, 2020, 107f. Tese (Doutorado em Física Ambiental), Instituto de Física, Universidade Federal de Mato Grosso.

Nas cidades existe um processo de armazenamento de calor durante o dia e devolução para atmosfera desse calor durante a noite. Dessa forma, é necessário compreender como ocorrem as modificações de temperatura do ar ao longo do tempo e espaço para fornecer informações para tomada de decisão. Assim sendo, esta pesquisa teve como objetivo fazer a predição da temperatura do ar utilizando Rede Neural Recorrente, do tipo *Long-Short Term Memory* (LSTM). Utilizou-se como dados de entrada para Rede Neural variáveis ambientais e de cobertura do solo, coletadas em transecto móvel noturno na cidade de Cuiabá-MT no ano de 2011 e 2016, separados por estação: inverno, outono, primavera e verão. Na preparação de dados para uso da LSTM foi aplicado cálculo de distância que considerou a questão temporal e espacial. Para prever a temperatura do ar foram testadas várias combinações de parâmetros da LSTM e dados de 2011 foram utilizados para prever 2016, e com isso foram criados modelos. Para validar os testes e comparar dados reais com dados previstos foram utilizadas as estatísticas Coeficiente de Determinação (R^2) e *Root Mean Squared Error* (RMSE). Os resultados obtidos com R^2 superiores a 0,9, em grande parte dos testes, apontam para viabilidade do uso de LSTM para previsão em séries espaço-temporais. É possível concluir que o modelo LSTM foi capaz de prever a temperatura do ar em 2016 e simular para o ano de 2021.

Palavras-chaves: dados ambientais, predição de temperatura; redes neurais recorrentes; *deep learning*; séries espaço-temporais.

ABSTRACT

GUARIENTI, C. E. Modeling air temperature in urban transects using LSTM. Cuiabá, 2020, 107f. Thesis (PhD in Environmental Physics), Institute of Physics, Federal University of Mato Grosso.

In cities there is a process of storing heat during the day and returning that heat to the atmosphere at night. Thus, it is necessary to understand how the temperature changes in time and space to provide information for decision making. Therefore, this research aimed to predict the temperature of the use of the Recurrent Neural Network, type LSTM (Long-Short Term Memory). Used as input data for the Neural Network, environmental and soil cover variables, collected in the night mobile project in the city of Cuiabá-MT, in the year 2011 and 2016, separated by seasons: winter, autumn, spring and summer. The preparation of data for use of the LSTM was applied to a distance calculation that considers a temporal and spatial issue. To predict air temperature, various combinations of LSTM parameters and 2011 data were tested and used for the 2016 forecast, and with these models created. To validate the tests and compare real data with data, the Coefficient of Determination (R^2) and Mean Square Quadratic Error (RMSE) were used as statistics. The results obtained with an R^2 greater than 0.9, in most of the tests, point to the feasibility of using LSTM for forecasting time-space series. It is possible to conclude that the LSTM model was able to predict the temperature in 2016 and similar for the year 2021..

Keywords: environmental data, temperature prediction; recurrent neural networks; deep learning, space-time series.

Capítulo 1

Introdução

Conforto ambiental refere-se à medida de satisfação do ser humano em determinado espaço e é uma preocupação constante, por exemplo, ao se planejar edificações. O conforto térmico relaciona-se com estado de satisfação com as condições térmicas do ambiente ao redor. No mundo, atualmente, existem vários estudos a respeito do conforto ambiental, que buscam, por exemplo, compreender como ocorrem as modificações que contribuem para o armazenamento de calor durante o dia na cidade e sua devolução para atmosfera a noite. Dessa forma, é fundamental entender como ocorrem essas alterações na temperatura do ar ao longo do tempo para formulação de políticas públicas que possam prevenir esse fenômeno ou pelo menos atenuar possíveis consequências.

A tecnologia é utilizada como ferramenta de apoio ao método científico. Por exemplo, na área de Conforto Ambiental é possível realizar determinado percurso (transecto) coletando dados como temperatura do ar e Umidade Relativa do Ar (UR). Com esses dados é possível realizar analises, simulações, modelagens, e desse modo, contribuir com a produção científica e gerar conhecimento.

No Brasil existem estudos que buscam encontrar zonas de conforto térmico adequadas para o ser humano. Pesquisas buscam compreender de que forma o uso e ocupação do solo podem contribuir com as mudanças térmicas. Nessa linha, existem estudos que buscam identificar, por exemplo, cidades com grandes variações de temperatura do ar em diferentes pontos geográficos do seu perímetro.

Em Cuiabá existem estudos que procuram identificar diferenças de temperatura e umidade intraurbana, que verificam diferentes formas de cobertura do solo por meio de transectos móveis. O método de transectos móveis é utilizado para coletar informações de variáveis como temperatura do ar, altitude e velocidade do vento. Esse tipo de pesquisa pode contribuir para melhor entendimento de como as características urbanas e o tipo de ocupação do solo podem alte-

rar elementos climáticos que compõe a atmosfera local. Para analisar os dados coletados no percurso realizado é necessário fazer uso de tecnologias computacionais, por exemplo, para compreender como ocorrem as alterações climáticas que acontecem no meio urbano.

As tecnologias computacionais têm contribuído nos mais variados campos da atividade humana. A sua contribuição na produção de ciência tem ganhado destaque no cenário mundial. É possível citar, por exemplo, os benefícios conseguidos por meio do uso de ferramentas computacionais no auxílio da automatização em atividades de pesquisa. A computação facilita e auxilia os cientistas na proposição, teste de hipóteses e na simulação de modelos.

O estudo das variáveis meteorológicas em um local é feito por um determinado período e isso gera determinada quantidade de dados. O conjunto das informações de variáveis formam séries de dados, que podem ser estudadas de forma espacial ou temporal. Por meio de técnicas computacionais é possível analisar séries espaço-temporais e a partir disso tentar fazer predição, isto é, prever dados.

Atualmente existem vários estudos sobre predição de dados utilizando diferentes técnicas. É possível citar, por exemplo, ARIMA (*Autoregressive Integrated Moving Average*) e SARIMA (*Seasonal Autoregressive Integrated Moving Average*). Hoje em dia percebe-se que as técnicas que utilizam Redes Neurais Artificiais (RNA) têm ganhado bastante atenção, pois as RNA oferecem uma gama de opções. Especificamente para fazer predição em séries as Redes Neurais Artificiais se destacam, por exemplo, por características como o uso de funções matemáticas para decidir se uma informação deve ser passada adiante ou não.

Neste estudo a problemática está relacionada com a necessidade de criar modelos preditivos precisos, que considerem a influência dos parâmetros climáticos e físicos, de modo que possa fornecer uma informação quantitativa sobre a temperatura do ar considerando dados de cobertura do solo no perímetro urbano na cidade de Cuiabá-MT.

Justifica-se a necessidade deste estudo pela falta de abordagens específicas utilizando Redes Neurais Artificiais na predição de séries espaço-temporais com dados ambientais juntamente com dados de cobertura do solo. Estudos recentes colocam RNA como uma das técnicas mais promissoras para fazer predição.

Então, a disponibilização de códigos que utilizam RNA para o estudo de espaço-temporais podem auxiliar pesquisadores que possuem dados capturados pelos mais diversos tipos de sensores. Dessa forma, pesquisadores podem utilizar esses códigos no estudo de predição, o que pode facilitar essa importante etapa

da pesquisa, pois podem utilizar um código já validado, e podem fazer ajustes como alteração de amostragem e variação de parâmetros, que já estão definidos, classificados.

Portanto, com o uso de Redes Neurais Artificiais na predição de dados em séries espaço-temporais procurou-se justificar a viabilidade dos códigos criados, auxiliando e facilitando etapas do método científico como a criação de algoritmos e aplicação de técnica estatística.

Este estudo apresenta códigos em linguagem computacional Python para criar modelos e fazer predição de temperatura do ar utilizando técnica computacional de Rede Neural.

1.1 Objetivo Geral

Realizar predição da temperatura do ar tendo como base séries espaço-temporais, com dados ambientais e de cobertura do solo, utilizando técnica computacional de Rede Neural Artificial.

1.2 Objetivos Específicos

Para atingir o objetivo geral deste trabalho, foram definidos os seguintes objetivos específicos:

- Preparar dados para serem utilizados como entrada para Rede Neural Artificial *Long-Short Term Memory* (LSTM).
- Criar e validar modelo de Rede Neural Artificial com dados de 2011 para prever temperatura do ar no ano de 2016 em Cuiabá-MT.
- Utilizar estatística para comparar dados reais e dados previstos.
- Simular previsão de dados para o ano de 2021 utilizando modelo já validado.

Na sequência deste estudo verifica-se que o capítulo 2 apresenta uma revisão bibliográfica de Redes Neurais Artificiais, variáveis ambientais, de cobertura do solo e Estatística. O capítulo 3 contém a metodologia e tecnologias utilizadas. No capítulo 4 apresenta-se os resultados obtidos juntamente com a discussão destes resultados. Por fim, no capítulo 5 visualizam-se as conclusões finais.

Capítulo 2

Fundamentação Teórica

Este capítulo apresenta uma revisão bibliográfica com teorias e conceitos necessários para a realização desta pesquisa e está organizado em 3 subcapítulos. O primeiro subcapítulo trata sobre a área de Conforto Ambiental, o segundo sobre Redes Neurais Artificias, o terceiro sobre Estatística.

2.1 Conforto Ambiental

O conforto ambiental é uma preocupação nos dias atuais, por exemplo, ao se construir espaços arquitetônicos, diferente de outros tempos, em que se atentava somente às questões de exigência técnica (NOGUEIRA et al., 2005).

Existem diversos estudos que avaliam o conforto ambiental, alguns utilizando métodos que consideram, em específico, as características climáticas do Brasil. O estudo do conforto térmico é importante em países de clima predominantemente tropical, que em grande parte atinge temperaturas elevadas, como é o caso da cidade de Cuiabá, Mato Grosso (NOGUEIRA et al., 2005). Na capital de Mato Grosso são caracterizados dois períodos: quente-seco, que compreende os meses de abril a agosto e outro quente-úmido, que vai de setembro a março (MAITELLI et al., 1991). No meio urbano, em particular, o clima é resultado de modificações ocorridas no processo de urbanização, pois esse processo altera características da atmosfera (OKE, 1978).

2.1.1 Clima urbano

Os fenômenos que influenciam o clima urbano em vários setores da sociedade geram preocupação, pois trazem impactos diretos em áreas urbanas (BARBOSA e AMORIM, 2012). As demandas que aparecem com o crescimento das

cidades e sua população geram necessidades de modificações na infraestrutura, por exemplo, melhoramento na drenagem das águas pluviais e canalização de córregos (MENDONÇA, 2014). Muitas cidades têm um modelo climático que de certa forma se diferencia com relação a sua circunvizinhança, é um fenômeno que é caracterizado como anomalia térmica. A caracterização das cidades é formada pela quantidade de pessoas que ali vivem, que ocupam pequenos territórios e assim podem trazer algum tipo de degradação. As mudanças globais, que são relacionadas à deterioração de ecossistemas, causam impactos em grande proporção na área urbana. Ao analisar as consequências dessas mudanças deve-se incluir características regionais, que estão relacionadas com a dinâmica climática (LOMBARDO, 2009).

Para Amorim (2000) tais problemáticas são resultantes do aumento da população e modificam a paisagem da cidade, é possível citar alguns exemplos como: diminuição da vegetação, impermeabilização do solo, canalização de córregos e aumento do fluxo de veículos.

Em Nogueira e Santos (2012), o clima urbano é resultante da interação entre os variados fatores urbanos, como o clima regional e com o meio físico. Dessa forma é preciso realizar uma definição espaço-temporal com aplicação em áreas urbanas. A escala horizontal é definida por três escalas:

- O macroclima é caraterizado por elementos meteorológicos que são obtidos em postos ou estações meteorológicas (SENATORE, 2006).

- O mesoclima corresponde a uma situação particular do macroclima, que define a influência temporal da cidade e do clima, e está ligada a camada limite urbana sendo esta em nível dos telhados. O termo escala local ou topoclima é utilizado para dizer quanto um mesoclima constitui critérios para caracterizar o clima de áreas urbanas menores, com características parecidas (PAULA, 2017).

- Microclima corresponde às condições climáticas de uma superfície realmente pequena (NÓBREGA; VERÇOSA, 2011).

As escalas espaciais mais utilizadas estão listadas conforme Tabela 1, (MENDONÇA; DANNI-OLIVEIRA, 2017):

Para realização de pesquisas ambientais em nível de microclima, o estudo das variações urbanas é recomendado a utilização de medidas móveis como transectos, juntamente com registros contínuos para maior detalhamento da distribuição horizontal de dados ambientais (SILVEIRA, 2007).

Tabela 1: Escalas espaciais.

Ordem de Grandeza	Escala Horizontal	Escala Vertical
Macroclima	$>$ 2000 km	3 a 12 km
Mesoclima	10 a 2000 km	100 m a 10 km
Microclima	alguns m a 10 km	abaixo de 100 m

As áreas que são influenciadas por grandes massas de vegetação possuem suas próprias especificidades, e também possuem microclimas diferentes de outras áreas, de modo que impactam no balanço de energia, seja por questões relacionadas aos processos biológicos de fotossíntese ou evapotranspiração (PAULA, 2017; OLIVEIRA; COSTA, 2005).

2.1.1.1 Determinantes do clima urbano

Os elementos climáticos podem definir a qualidade e fornecer componentes do clima, e são definidos pelas características físicas que trazem juntamente com as propriedades da atmosfera geográfica de um local (ROMERO, 2001). Os elementos climáticos principais são: temperatura do ar, umidade relativa do ar, precipitação e regime de vento.

2.1.1.1.1 Temperatura do ar

Segundo García (1978) e Ometto (1981) a temperatura do ar pode ser entendida como a energia contida e ligada ao balanço de energia, isso irá depender de fatores como umidade relativa do ar, por exemplo. A distribuição da temperatura do ar varia conforme a latitude, relevo, correntes advindas do oceano entre outros (ROMERO, 2001). A variação da temperatura do ar diária está ligada a quantidade de energia solar e como consequência o aquecimento do solo. Para mensurar a temperatura do ar na atmosfera são realizadas observações em altura de 1,25 a 2 metros acima do terreno e o instrumento usado para medir é o termômetro, existindo alguns tipos como o de máximo e mínimo. A unidade de medida utilizada pode ser Celsius, Fahrenheit e Kelvin (VAREJÃO, 2006; FRANCO, 2010).

2.1.1.1.2 Umidade relativa do ar

A quantidade de vapor de água presente no ar e a quantidade máxima de vapor de água que o ar pode conter em determinada temperatura é denominada

umidade relativa, sendo utilizada unidade de medida em porcentagem (FROTA; SCHIFFER, 1995). A umidade relativa do ar leva em conta a quantidade de água existente no ar e a quantidade máxima de água que poderia existir, existe uma relação direta com a temperatura do ar, e quando a umidade relativa chega a 100% há uma indicação de que o ar está com quantidade máxima de água, podendo resultar em condensação. Quando a temperatura diminui, a quantidade de vapor de água no ar não muda, mas a umidade relativa do ar aumenta (BUFERA, 1995). Segundo FRANCO (2010) a umidade relativa do ar pode ser mensurada por higrômetros ou psicrômetro. Atualmente o higrômetro elétrico é o mais utilizado.

2.1.1.1.3 Regime de vento

O vento é o resultado entre o atrito onde se encontra e se movimenta com a superfície da terra, sendo também resultante da pressão, que são zonas de alta para baixa pressão (BROWN; DEKAY, 2009). Na superfície terrestre ocorrem diversas modificações, que influenciam na modificação da direção do vento em áreas urbanas, tais modificações ocorrem por conta de prédios, ruas e tipos de arvores presentes no meio. O anemômetro é o instrumento utilizado para medir o vento (MENDONÇA; DANNI-OLIVEIRA, 2017).

2.1.1.1.4 Precipitação

Precipitação é uma referência a queda de água, seja no estado líquido (chuva), seja no estado sólido (neve e granizo), e nos dois casos exemplificados deriva da atmosfera. Devido a formação de ilhas de calor em áreas urbanas ocorre maior quantidade de precipitação, entre 5 a 10% quando compara-se com propriedades rurais (BARBIRATO et al., 2007). A forma tradicional para medir a precipitação é por meio de redes pluviométricas que contém pluviômetros e pluviógrafos. Porém, a precipitação também pode ser medida utilizando sensoriamento remoto (CONTI, 1998).

2.1.2 Anomalia térmica

Ilha de calor é um tipo de anomalia térmica, e também é o fenômeno mais importante das modificações que influenciam o armazenamento de calor durante o dia na cidade, isso ocorre pelas propriedades térmicas dos materiais de construções e sua devolução para a atmosfera (SANTOS; NOGUEIRA, 2012).

A ilha de calor atmosférica apresenta variabilidade temporal, é mais perceptível pela noite, e identifica alterações na troposfera, essas modificações ocor-

rem em decorrência da diferença dos tipos de uso da terra e cobertura entre as paisagens urbanas e rurais (FIALHO, 2012).

O desenvolvimento de ilhas de calor nas cidades é responsável pela diferença térmica entre o ambiente urbano e rural (ROCHA et al., 2011). Resultantes de propriedades físicas dos prédios e outras estruturas da área urbana, as ilhas de calor emitem calor por meio de atividades da sociedade (SOUZA et al., 2010). Quando compara-se a temperatura do ar em regiões urbanas com a rural é possível verificar que normalmente a temperatura é mais quente após o pôr do sol e mais frescas após o amanhecer, respectivamente (LOMBARDO, 1985). E com o aumento das áreas urbanas, essas ilhas de calor ficam maiores, porque ocorre a diminuição da vegetação e aumento do numero de prédio, casas e estruturas de concreto (GARTLAND, 2011). A ilha de calor urbano pode ser considerada como a diferença entre temperaturas medidas na área urbana e os valores medidos nas mediações deste espaço (COHEN et al., 2012). As ilhas de calor estão diretamente ligadas a temperatura e segundo Brandão (2003) e Rocha et al. (2011) é possível ter uma classificação de sua magnitude:

- Diferenças entre 0 a 2°C é considerada fraca.
- Intensidade entre 2 a 4°C pode se dizer moderada.
- Entre 4 a 6°C: forte.
- Intensidade maior que 6°C considera-se muito forte.

As ilhas de calor que possuem maior intensidade estão associadas com: geometria urbana, fator de visão de céu, redução da veloc idade de vento, poluição do ar, emissão de calor dos edifícios, tráfego e metabolismo dos organismos vivos, cobertura do solo e dos materiais de construção (ALCOFORADO; ANDRADE, 2006).

2.1.2.1 Classificação e influência do revestimento de solo

O tipo de uso do solo, seja residencial, comercial ou industrial, juntamente com o revestimento do solo influenciam como os componentes do ambiente sofrem modificações, em específico, na questão de radiação incidente, que tem relação com a altura das edificações, a distância entre elas e o material utilizado (PAULA, 2017).

O balanço de energia é específico em cada cidade, de modo que é composto por muitas variáveis, por exemplo: densidade da construção, densidade

populacional, tamanho, canalização de rios, entre outros (PAULA, 2017). Nessa linha, Givoni (1989) afirma que é na camada intraurbana que acontecem as trocas de energia que impactam os moradores da cidade, visto que as características da superfície e a forma das estruturas determinam essas trocas de energia.

Para Oke et al. (2004) as principais variáveis no estudo do clima urbano são: tipo de revestimento do solo, a estrutura urbana e tecido e metabolismo humano, isso porque esses elementos compõe a morfologia das áreas urbanas, pois utiliza-se a dimensão das escalas que impactam na atmosfera.

A caracterização do solo pode ser feita, basicamente, seguindo duas formas: conceitual e modelos do balanço de energia, que exige sólido embasamento teórico, grande quantidade de dados e equipamentos sofisticados (NASCIMENTO et al., 2013).

A classificação de cobertura e uso da terra pode ser feita a partir de imagens de satélite. Por meio dessas informações é possível gerar mapas com diferentes classificações, por exemplo, vegetação rasteira, vegetação arbórea e solo exposto, entre outros (PAULA, 2017).

Existem diferentes técnicas de classificação das imagens de satélite, que podem ser divididas, resumidamente, em supervisionada e não-supervisionada, sendo que a última não precisa de informações da classe. No caso da supervisionada é necessário identificar previamente as classes, essa opção é comumente usada quando já se tem conhecimento sobre as classes (PAULA, 2017).

2.1.2.2 Cobertura de solo

A utilização do solo está diretamente ligada ao uso da sociedade sobre a terra como construções, áreas de lazer, desmatamentos que modificam o solo e por consequência o meio ambiente (PRAKASAM, 2010). Neste contexto pode-se dizer que a cobertura do solo é composta por componentes que sobrepõem a superfície da terra, o que inclui vegetação, solos, neve ou água, que reflete o estado biofísico da terra (ASSIS, 2000; OKE, 2002). Os mais variados tipos de características de cobertura do solo trazem um aumento de temperatura, pois o ambiente absorve e emite uma maior quantidade de calor, pelo uso do solo e materiais que o sobrepõem (ALVES; VECCHIA, 2012).

2.2 Redes Neurais Artificiais

2.2.1 Inteligência Artificial

O campo de Inteligência Artificial (IA) começou após a Segunda Guerra Mundial, por volta do ano de 1950, e é um dos campos mais recentes em ciências e engenharia. Este campo de conhecimento vai ao encontro do anseio da sociedade que busca perceber, compreender e até prever o que ocorre no mundo. De modo que IA vai um passo além, não se restringe em compreender processos inteligentes, mas também construí-los, dessa forma, abre espaço para grandes descobertas (RUSSEL; NORVIG, 2013).

Inteligência Artificial envolve tarefas relacionadas com processamento simbólico, reconhecimento de imagens, e qualquer tipo de aprendizado, sua aplicação pode ir de jogos até prova de teoremas (ROSA, 2011). As tecnologias de IA têm potencial de modificar as formas de trabalho realizando tarefas que podem ser automatizadas, atividades administrativas podem ser realizadas em menor tempo, de forma melhor e com custo reduzido (KOLBJØRNSRUD et al., 2016). Ferramentas de Inteligência Artificial têm sido usadas em novas perspectivas, por exemplo, em teses de doutorado (CROSTA et al., 2018).

O uso de Inteligência Artificial envolve vários algoritmos de aprendizagem e generalização para simular capacidades humanas. Um exemplo é o estudo de como funciona o cérebro humano, como são recebidas e processadas as informações sensoriais (KONAR, 2018). Neste cenário os computadores podem usar a razão e criar seus próprios conceitos, podem resolver problemas complexos e até entender a linguagem humana (JACKSON, 2019).

Nesse sentido vários algoritmos de IA têm permitido que se explore grandes volumes de dados, por exemplo, gerando conhecimento a partir da análise de mídias sociais utilizando *Machine Learning* (ML), isto é, Aprendizado de Maquina (BECHMANN; BOWKER, 2019). ML pode ser usado na área ambiental tendo como base grandes volumes de dados, essas informações podem ser utilizadas para fazer predição, por exemplo, da temperatura (BROWNLEE, 2018b).

2.2.1.1 Machine Learning

Machine Learning é considerada uma área dentro de IA, seu conceito nasceu em 1980. Nas últimas décadas ocorreu grande crescimento no uso das técnicas ML, em diversas aplicações, é possível destacar tarefas que ajudam na tomada de decisão. Neste contexto, em que normalmente é necessário trabalhar com grande quantidade de dados. é possível citar, por exemplo, tarefas relaci-

onadas com diagnóstico laboratorial, processamento de imagem e mineração de dados (JUNIOR et al., 2018). O aprendizado pode ser ou não supervisionado de acordo com algumas características.

Algoritmos que usam aprendizado supervisionado têm a característica de trabalhar com pesos conhecidos, neste caso os valores, também chamados de classes, esperados na saída já são conhecidos. Um exemplo é o *backpropagation*. Por outro lado, algoritmos que usam aprendizado não-supervisionado não possuem conhecimento anterior da saída, isto é, não possuem classes ou rótulos, é possível citar como exemplo algoritmos de agrupamento (LIMA et al., 2016; SIMÕES; COSTA, 2007).

Para Russel e NORVIG (2013) um algoritmo está aprendendo quando é capaz de melhorar o seu desempenho em tarefas futuras após fazer uma determinada quantidade de observações necessárias. As melhorias dependem de quatro fatores principais: o componente que deve ser melhorado, o conhecimento prévio que já se tem, a representação usada para os dados e componente, e, por fim, o *feedback* disponível para aprendizagem. Cada componente pode ser aprendido por realimentação própria.

Em Lopes et al. (2017) o conceito de *Machine Learning* está relacionado com algoritmos capazes de aprender e fazer previsões tendo por base um conjunto de dados. Já em Dehaven (2014) ML relaciona-se com a capacidade de, a partir de uma base, fazer previsões e tomar decisões. Para Gracioso et al. (2018) a aprendizagem está relacionada com capacidade de detectar e prever padrões, por exemplo, de trânsito em uma rodovia, em que o aprendizado tem sucesso quando, por meio da análise de dados dos padrões anteriores, o algoritmo é capaz de melhorar suas previsões.

O conceito de ML existe já há algum tempo. As mudanças recentes são relacionadas com grandes quantidades de dados disponíveis e rotulados, um exemplo disto é o treinamento de *Machine Learning* tendo como base faces de pessoas. Outro fator importante é o aumento de poder computacional devido às unidades de processamento gráfico. Esses avanços proporcionaram que a mesma metodologia fosse usada, mas obtendo resultados muito melhores (MARX, 2019).

O processo de mineração em dados pode utilizar algoritmos de ML com objetivo de fazer regressão, classificação e agrupamento. Em Ciências Ambientais, no sensoriamento remoto urbano, o *Machine Learning* pode auxiliar na classificação das classes de cobertura do solo urbano alcançando elevados níveis de detalhamento. Estes algoritmos contribuem de forma significativa para melhorar a acurácia das classificações. (RUIZ et al., 2019).

Ainda na área ambiental, os algoritmos são capazes de auxiliar em diversas situações, por exemplo, contribuir com agências ambientais na tarefa de monitorar a qualidade do ar e fazer predição de poluentes, visto que a poluição atmosférica é nociva à saúde, em especial, para idosos e crianças. Vários sensores são capazes de captar dados sobre a concentração de poluição no ar. Essas informações coletadas podem ser consideradas séries temporais, pois representam uma sequência ordenada. Com o aumento do fluxo é necessário buscar algoritmos, por exemplo, Aprendizagem de Máquina, pela sua capacidade de aprender novos padrões ao longo do tempo (GRIM et al., 2019).

O uso intensivo de *Machine Learning* juntamente com uso de *Big Data*, isto é, grandes volumes de dados e bases, fomentam a necessidade de novas soluções que comportem este cenário. O uso de Aprendizado Profundo - *Deep Learning* - representa um ganho estatístico de performance, por exemplo, na classificação de imagens (PONTI; COSTA, 2018).

As Redes Neurais Artificiais (RNA) apresentam vantagens em comparação com outros modelos estatísticos (TUMBO et al., 2002; ZHANG et al., 2002), por exemplo, é possível citar sua capacidade de não precisar do conhecimento prévio, e, mesmo assim, ser capaz de encontrar a relação entre dados de entrada e saída (LIMA, 2019).

As Redes Neurais Artificiais tiveram sua origem por volta da década de 1940 por meio da analogia com células nervosas formulada por Warren McCulloch e Walter Pitts (HAYKIN, 2007).

No ano de 1949, Donald Hebb dá seguimento aos estudos de Warren McCulloch e Walter Pitts e cria um novo principio de aprendizado em sistemas nervosos complexos, que explica de forma quantitativa o funcionamento da sinapse.

Já no ano de 1951, surge o primeiro simulador de cadeia natural, denominado SNARC, foi desenvolvido por Marvin Minsky. Em 1956, Nathaniel Rochester apresentou modelo de Rede Neural Artificial que simulava centenas de neurônios.

Em 1959, Frank Rosenblatt criou o conceito de rede perceptron, uma rede com múltiplos neurônios. Nos anos posteriores Bernard Widrow criou modelo de processamento em Redes Neurais Artificiais com destaque para o aprendizado, modelo denominado como ADALINE - *ADAptive LINear Element*.

Na década de 1980, vários estudos na área de Neurocomputação ganharam destaque, é possível citar a Universidade de Stanford com os professor John Hopfield e David Rumelhart (HAYKIN, 2007). No ano de 1987, ocorreu a pri-

meira conferência de Redes neurais Artificias, momento em que a Neurobiologia ganhou destaque.

As RNAs procuram imitar o funcionamento do cérebro humano, isto é, existe uma inspiração na natureza (BRABAZON et al., 2015), pois baseiam-se no funcionamento dos neurônios, nas suas ligações e nas suas potencialidades. Dessa forma, pretende simular a fisiologia desse órgão humano, e isso é feito por meio da programação em linguagem computacional usando técnicas de modelagem. A Rede Neural funciona como um processados distribuído, composto por unidades de processamento simples, que são capazes de armazenar conhecimento experimental. (KOPILER et al., 2019)

Na área de Biologia existe a figura dos dentritos, que são responsáveis por conduzir estímulos captados, seja do ambiente ou mesmo de outras células. Existe ainda a figura do corpo celular e dos axônios, que transmite impulsos nervosos (AMANCIO et al., 2016; FERNEDA, 2006). Na Figura 1 é possível visualizar uma representação simples de um neurônio:

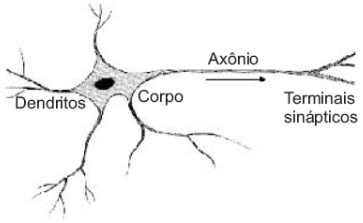


Figura 1: Neurônios com suas principais partes. Fonte: Ferneda (2006).

A semelhança das RNA com as redes neurais biológicas envolve questões elementares como as unidades básicas, nos dois casos representado pela figura do neurônio. O modelo mais simples de RNA é chamado de *perceptron*, que visualiza-se na Figura 2, em que existem apenas neurônios na camada de entrada *input cells* e neurônios de saída *output cells*.

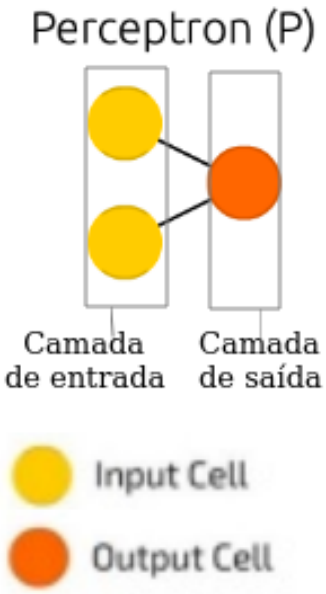


Figura 2: Neurônios dispostos em duas camadas.

Na perspectiva computacional o conjunto de vários neurônios gera camadas e, dessa forma, cria-se *Multilayer Perceptron* (MLP) (SILVA et al., 2010), isto é, uma Rede Neural com mais de uma camada, conforme Figura 3, que é composta por vários neurônios (SOUZA, 2019), que é uma RNA clássica (BROWNLEE, 2018b).

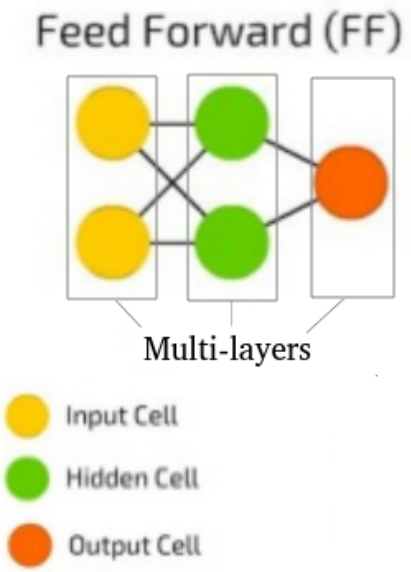


Figura 3: Neurônios dispostos em camadas de entrada, oculta e saída.

Na perspectiva fisiológica sinapse é o local de contato entre os neurônios e pode ser medida pela sua intensidade, o que pode ser simulado utilizando um

neurônio artificial. Em um neurônio o sinal percorre o corpo celular e por meio do axônio é transmitido para outros neurônios (CERQUEIRA et al., 2001). Dessa forma, o conhecimento sobre o modo de funcionamento da fisiologia do cérebro humano permite a realização de uma modelagem matemática, capaz de imitar o comportamento, por exemplo, dos neurônios e suas comunicações. Dessa forma, permite criar redes neurais artificiais (MIRANDA et al., 2015).

O funcionamento do neurônio artificial envolve o entendimento de entradas, pesos, e uso de funções como soma e degrau. Na Figura 4 é possível verificar círculos que representam neurônios, os pesos aplicados em cada neurônio e as linhas que interligam neurônios com as funções.

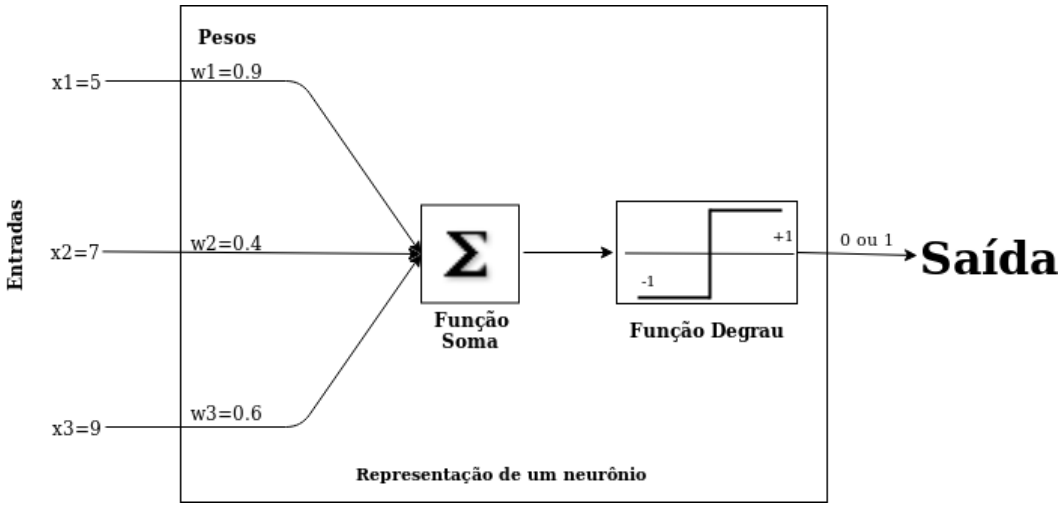


Figura 4: Funcionamento do neurônio artificial.

A equação abaixo representa os cálculos que envolvem este modelo simples de Rede Neural:

$$\text{soma} = \sum_{n=1}^{n} xi * wi$$

x = valor do neurônio
w = peso

Utilizando duas etapas diferentes e usando os valores e pesos da Figura 4 chega-se os aos seguintes resultados:

Etapa 1. Aplicar função Soma.
soma = (5 * 0.9) + (7 * 0.4) + (9 * 0.6)
soma = (4.5) + (2.8) + (5.4)

soma = 12.7

Etapa 2. Aplicar função Degrau. Essa função recebe o valor de saída da Função Soma, no caso exemplificado é 12.7, e considera que qualquer valor maior que zero fica então com valor 1, outras situações ficam com zero, ou seja, tem uma saída binária: zero ou um.

degrau: soma $> 0 = 1$
degrau: soma $< 0 = 0$
degrau: $12.7 > 0 = 1$
saída = 1

2.2.1.1.1 Funções de Ativação

As funções de ativação que também são chamadas de *transfer function* fazem transferência de valores de uma parte da rede neural para outra, que determina a saída da rede neural, pode ser a resposta final caso seja a camada de saída, pode ser também a transmissão de uma camada oculta para outra. O uso de funções de ativação permite que a saída não seja apenas linear (PONTI; COSTA, 2018). As funções de ativação aplicam funções matemáticas aos dados gerando uma resposta (BINOTI et al., 2014).

A função linear não tem nenhum tipo de ativação, é usada em saídas com problemas de regressão, envolve cálculo de valor contínuo, usa apenas o resultado das somas da rede neural, ou seja, a saída é linear, o que pode pode representar uma limitação. A equação da função Linear está representada abaixo:

f(x) = x

Na Figura 5 é possível verificar a representação gráfica da função Linear.

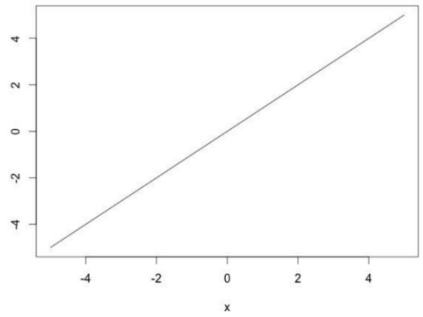


Figura 5: Função linear.

A função de ativação Sigmoid é muito utilizada em Redes Neurais Artificiais (LEAL et al., 2015; HAYKIN, 2007), o mesmo ocorre com a função de ativação Tangente Hiperbólica. A função Sigmoid aplica cálculo não linear, é usada em problemas linearmente não separáveis, em especial, saídas de classificação binária, isto é, zero (não ativação) ou um (ativação). Abaixo a equação da função Sigmoid:

$$sigmoid(z) = \frac{1}{1 + e^{-z}}$$

A função Sigmoid está representada na Figura 6:

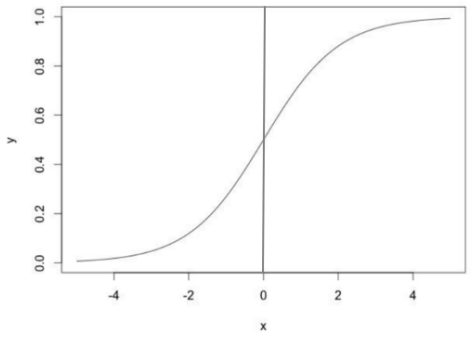


Figura 6: Função sigmoid.

A função Tangente Hiperbólica (TanH) é muito utilizada em camadas ocultas, e em problemas de classificação (SANTOS et al., 2005), possui valores 1 e -1, é mais rápida que a função Sigmoid ao utilizar dados com média zero. Possui desvantagem em comum com a Sigmoid: *varnishing gradients*, se o peso estiver em regiões com pouca inclinação, isto é, próxima de limite inferior ou superior, o

aprendizado se torna lento. Abaixo a equação da função TanH:

$$Y = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

Na Figura 7 visualiza-se o gráfico da função Tangente:

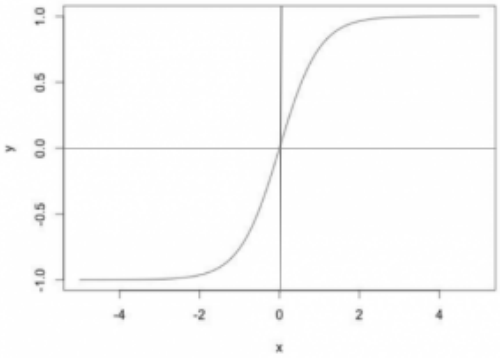


Figura 7: Função Tangente.

A função de ativação denominada *Rectied Linear Units* (ReLU) é muito usada em Redes Neurais Profundas, é uma função não linear, que retorna valor zero ou maior que zero, ou seja, cancela valores negativos. Nessa função a inclinação é constante e não possui valor máximo. O funcionamento da função ReLU envolve receber um valor que, sendo maior que zero o valor recebido retorna o próprio valor, e se o valor for menor que zero retorna o valor 0 (PONTI; COSTA, 2018). Na Figura 8 visualiza-se a representação da função ReLU:

f(x) = max(a,x)

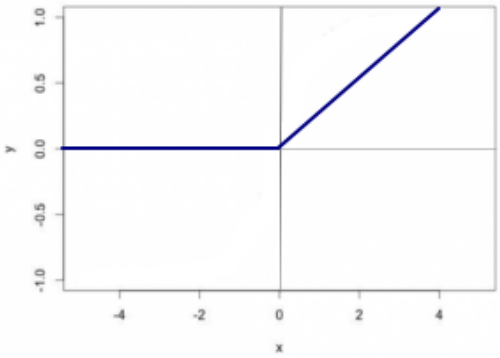


Figura 8: Função ReLU.

2.2.1.1.2 Otimizadores

Optimizers (otimizadores) funcionam como método de atualização dos pesos e buscam minimizar a função de erro. O otimizador mais simples é o que aplica diretamente o algoritmo *backpropagation*. Já outros otimizadores acrescentam termos à fórmula de atualização dos pesos, um exemplo é o Adam, que é um método de otimização estocástico (KINGMA; BA, 2014).

Learning rate (taxa de aprendizagem) é um parâmetro que pode ser fixado ao escolher o otimizador (DUCHI et al., 2011), usa-se valor zero ou superior, no caso do otimizador Adam um exemplo de valor que é comumente usado é "0.001".

Metrics (métricas) é uma função usada para avaliar a performance do modelo. É uma função similar a loss function, se diferencia por não utilizar os resultados da avaliação de uma métrica para treinar um modelo.

2.2.1.1.3 Loss Function

Loss function é a função de perda, que pretende minimizar o erro, e avalia um conjunto de pesos. Já a *loss* (perda) é o número da perda. O cálculo de erro envolve considerar a reposta correta e subtrair a resposta calculada. A comparação entre valores previstos e valores reais é feita usando uma medida (BROWNLEE, 2018a), nesse sentido, é possível citar três algoritmos: *Mean Absolute Error* (MAE), *Mean Squared Error* (MSE) e *Root Mean Squared Error* (RMSE).

MAE é um cálculo muito utilizado para avaliação de modelos. O MAE dá o mesmo peso para todos os erros, diferente do que ocorre com RMSE (CHAI; DRAXLER, 2014). MAE mede a diferença entre duas variáveis contínuas, é usado em estudos que envolvem regressão linear, e está representado abaixo na equação:

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|e_t|$$

O MSE é a média do erro ao quadrado, que é usado como função de perda para a regressão de mínimos quadrados, é a soma de todos os pontos de dados, do quadrado da diferença entre as variáveis de destino previstas e reais, dividida pelo número de pontos (DUA et al., 2017), e está representado abaixo:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2$$

fy = classe
yi = calculado

N = Quantidade de registros

Para Dua et al. (2017) o RMSE é a raiz quadrada do MSE e é medido nas mesmas unidades que a variável alvo e se diferencia do MSE pois aplica raiz quadrada, e quando eleva valor ao quadrado evita que tenha valor negativo. O *Root Mean Squared Error* está representado abaixo:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2}$$

fy = classe
yi = calculado
N = Quantidade de registros

As Redes Neurais Artificiais (RNAs) são um modelo matemático inspirado na neurobiologia (OLIVEIRA et al., 2015), e constituem um conjunto de unidades com funções matemáticas ligadas pelo valor do peso. São capazes de aprender conforme a variação do valor de pesos e a partir disso generalizar o aprendizado (GORGENS et al., 2009).

Equações são frequentemente usadas para criar modelos matemáticos que simulam de forma realista os mecanismos cerebrais e conseguem fazer simulações eficientes. Dessa forma, esses modelos são usados para simular o comportamento do sistema nervoso, assumindo que vários parâmetros do mecanismo alvo sejam conhecidos (FABRE, 2017; FERREIRA et al., 2019).

Desde que as RNAs foram criadas, por volta do ano de 1940, elas têm sido utilizadas para resolver vários problemas de aprendizado supervisionado, e isso despertou bastante interesse da comunidade científica, em particular, pelo novo arranjo proporcionado por sua arquitetura(FABRE, 2017).

AS RNAs funcionam como sistemas computacionais formados por unidades de processamento, também chamados de neurônios artificiais ou simplesmente nós. Os nós são modelos matemáticos que processam informações recebidas, que possuem pesos, e com isso fornecem uma saída (CAMPOS et al., 2017).

Para Oliveira et al. (2015) a principal característica das Redes Neurais Artificiais é a sua capacidade de aprender a partir do meio e melhorar o seu desempenho de forma ordenada, de modo que possa alcançar um objetivo estabelecido. Dessa forma, o processo de aprendizagem acontece de uma forma iterativa e tem por base os ajustes utilizados nos pesos, que podem ser comparados às intensidades dos sinais de sinapses dos neurônios biológicos.

Em Campos et al. (2017) e Bullinaria (2004) verifica-se que RNA tem

apresentado desempenho superior aos modelos de regressão. É possível elencar algumas características que contribuem para esse resultado: estrutura maciça e paralelamente distribuída, habilidade de aprender e generalizar, tolerância a falhas e ruídos; capacidade de modelar diversas variáveis e suas relações não-lineares.

As arquiteturas das Redes Neurais podem ser divididas em duas classificações principais: *Feedforward* e *Recurrent Neural Networks*, isto é, Redes Neurais Recorrentes (RNN) (FABRE, 2017). As RNs do tipo *Feedforward* não possuem realimentação, diferente das RNNs, os seus neurônios são organizados em uma ou mais camadas (PEREIRA et al., 2019).

2.2.1.1.4 Redes Neurais Recorrentes

As *Recurrent Neural Networks* (RNNs), Redes Neurais Recorrentes, têm característica própria, possui laços de realimentação, que criam uma memória intrínseca, algo que favorece o tratamento de problemas, por exemplo, de séries temporais, pois existe uma dependência temporal entre amostras (JAEGER, 2001; TADANO et al., 2017). Outra característica que é possível destacar é sua capacidade de processar dados sequenciais de comprimento variado, fator importante, por exemplo, processamento de idiomas (GAO et al., 2018).

As Redes Neurais Recorrentes podem ter camada única ou podem ter arquiteturas profundas, inclusive podem ser combinadas com outras arquiteturas alcançando resultados interessantes (SUN et al., 2019). As RNNs são uma classe das Redes Neurais Artificiais, e se diferencia pelo modo como seus neurônios são conectados formando um ciclo direto (MERITY et al., 2018).

As RNNs são capazes de tratar informações contextuais de longo prazo, conforme Figura 9, o que a diferencia das Redes Neurais do tipo *Feedforward*. Por outro lado, enfrenta questões delicadas como a dificuldade de treinamento com as funções de ativação mais comuns, como a tangente e a sigmoid, isso por conta do *vanishing gradient*, desaparecimento do gradiente, e *error blowing up problems*, explosão de erro. Para resolver essas questões propôs-se uma nova arquitetura chamada *Long-Short Term Memory* (LSTM) (DU et al., 2015). LSTM usa a combinação de *gates* para tratar o *vanishing gradient*.

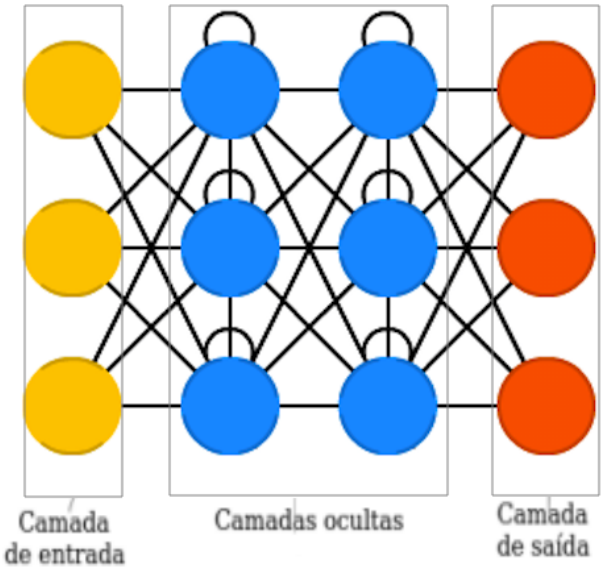


Figura 9: Exemplo de Rede Neural Recorrente.

LSTM resolve o problema de desaparecimento de gradiente criando uma conexão entre as ativações do *gate* e o gradiente, essa conexão cria um caminho para o fluxo de informações por meio do gate para informações que o LSTM não deve esquecer. Dessa forma, quando decidi-se que uma informação não deve ser esquecida ela tem valor próximo ao máximo, no caso 1.

Este tipo de Rede Neural trabalha de forma estruturada e é capaz de modelar sequências, o que a torna uma opção viável para diversas situações, por exemplo, modelar amostras de som digital (WYSE; HUZAIFAH, 2019), e modelagem de linguagem, em que umas das aplicações reconhecidas pela performance é tentar prever a próxima palavra (TSUJI et al., 2019; BANERJEE et al., 2019).

O aumento da complexidade em análises ao estudar sequências e modelar relacionamentos de longo prazo favoreceu a criação de novos tipos de RNAs. Ocorre que mesmo RNNs bidirecionais, uma variação com capacidade para aprender em dois sentidos, possuem limitada capacidade de reter vários passos para trás, algo que foi tratado por redes mais profundas (ARMENTEROS et al., 2019).

2.2.1.2 Deep Learning

Deep Learning (DL) é um campo dentro de *Machine Learning* (MOHSEN et al., 2018). O termo profundo (*deep*) é comumente usado como referência ao número de camadas utilizadas. Para Gracioso et al. (2018) e Jost (2015) e LeCun et al. (2015) o DL recorre à necessidade de múltiplas camadas. Para Kotu e Deshpande (2018) uma rede com três (3) ou mais camadas ocultas, isto é, entre

as camadas de entrada e saída, é considerada profunda. Em Wason (2018) existe um entendimento diferente: o número mínimo de camadas ocultas é quatro (4).

Ceci et al. (2015) define DL como o uso de técnicas *Machine Learning* para aprendizado de novos padrões. Já para Heaton (2018) a questão de profundidade está mais relacionada ao uso de um número maior de funções matemáticas do que os modelos tradicionais de aprendizagem.

Ponti e Costa (2018) colocam DL na perspectiva de descobrir modelos, padrões e parâmetros utilizando exemplos e um método para o aprendizado. Ao final do processo é possível obter uma função capaz de receber uma entrada de dados e fornecer uma saída com uma representação adequada para o problema analisado organizadas de forma hierárquica, na busca de padrões e características.

Técnicas *Deep Learning* podem ser usadas em várias situações, brevemente, é possível citar: auxílio na resolução de problemas em processamento de linguagem natural, reconhecimento automático de voz e indexação automática de imagens. A descrição e indexação de imagens na Web tem importante papel na organização do conhecimento, dessa forma, esses algoritmos se tornam uma ferramenta que contribui para realizar a representação documental de fotografias e documentos gráficos (GRACIOSO et al., 2018).

DL é uma ferramenta capaz de tratar problemas complexos, deste modo tem sido usado em análise de imagens médicas. Um exemplo é o uso dessa técnica para classificar tumores do cérebro com base em figuras de ressonância magnética. Em Mohsen et al. (2018) DL atingiu elevado nível de acurácia e obteve resultados superiores em comparação com outros classificadores como algoritmo KNN, *K-nearest neighbor*, LDA, *Linear Discriminant Analysis* , e os algoritmos *SMO, Sequential Minimal Optimization* (MOHSEN et al., 2017).

O uso de Aprendizado Profundo envolve questões como especificar camada de entrada, número de camadas ocultas, definir a camada de saída, quais propriedades devem ser mantidas ou excluídas em um banco ao fazer uma predição. Em Wang et al. (2018) esta técnica foi usada pra desenhar a estrutura da proteína e prever a possibilidade de vinte (20) aminoácidos naturais. Foram testadas diferentes combinações, dessa forma, o melhor ajuste não utilizou todas as propriedades do banco de dados. Os resultados apontaram performance superior ao SPIN utilizado em Li et al. (2014).

Redes Neurais de Aprendizado Profundo (DL) têm a capacidade de aprender automaticamente mapeamentos complexos de entrada para saída, pode suportar também múltiplas entradas e saídas, essas características tornam DL uma alternativa para previsão de séries temporais, em especial, para tratar de depen-

dências complexas e previsão em várias etapas (BROWNLEE, 2018b).

DL usa uma composição de funções, em que cada função normalmente é referência de uma camada, diferente de *Machine Learning* que usa apenas uma função para alcançar o resultado desejado. Assim, a profundidade de *Deep Learning* permite aprender uma sequência de funções que mapeiam e modificam vetores. Cada função atua sobre uma entrada e gera uma representação, que é passada adiante para a próxima função, dessa forma, buscando uma representação mais adequada na tarefa, de modo que pretende encontrar o escopo das relações nos dados originais. (PONTI; COSTA, 2018).

2.2.1.3 Long-Short Term Memory

LSTM é um tipo de arquitetura das Redes Neurais Recorrentes, e que tem a característica de preservar memória. As informações são passadas aos neurônios da próxima camada. É muito utilizada em estudos que envolvem séries temporais pela característica de dependência entre as variáveis com relação ao tempo (BROWNLEE, 2018b).

As Redes Neurais do tipo *Long-Short Term Memory* foram projetadas especificamente para resolver o problema de *vanishing gradient*, que é a dificuldade em encontrar a melhor combinação de pesos por meio do gradiente em que se alcance o menor erro possível. As RNNs clássicas possuíam limitada capacidade de acessar informações contextuais (GRAVES et al., 2008). As RNNs existentes, até então, produziam erros motivados por atrasos longos que eram inacessíveis por conta do problema de *vanishing gradient* (GRAVES; SCHMIDHUBER, 2005).

LSTM se diferencia das redes neurais classicas *Multilayer Perceptron*, pois não apenas propagam os dados de entrada. As *Long-Short Term Memory* possuem conexões recorrentes (Figura 10), de modo que consegue utilizar informações sobre ativações anteriores para formular uma saída (BROWNLEE, 2018a).

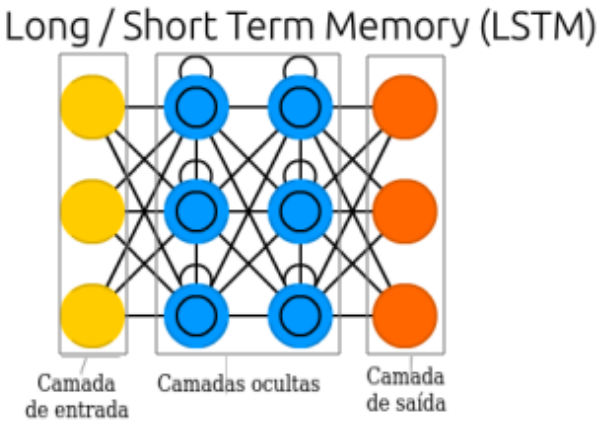


Figura 10: Representação do modelo Stacked LTSM.

Para Graves e Schmidhuber (2005) *Long-Short Term Memory* consistem em um conjunto de blocos conectados de forma recorrente, que possui uma ou mais células de memória conectadas também em forma recorrente, que possui três células multiplicativas: portas de entrada, de saída e esquecimento. Dessa forma, a rede interage com as células por meio das portas (GRAVES; SCHMIDHUBER, 2005).

Para (BROWNLEE, 2018b) de forma resumida, LSTMs possuem três benefícios em comparação com outras RNAs:

- LSTM é capaz de tratar problemas em desaparecimento e explosão de gradiente.
- Possui memória capaz de superar os problemas de dependência temporal de longo prazo com sequência de entrada.
- Permite dados de entrada e saída de comprimento variável, isto é, permite sequências de entrada de processo e sequências de saída, passo a passo, no tempo.

As *Long-Short Term Memory* possuem neurônios que recebem dados de entrada (Xt) que recebem uma série de processamentos por meio das funções de ativação. Existem três *gates* (portões): *forget gate* (x), *input gate* (+) e *output gate*. Dessa forma, o uso de funções sigma e tangente sobre dados de entrada contribui para decidir, na célula de memória, o que será esquecido, adicionado e enviado como saída do neurônio, esse processo está representado na Figura 11.

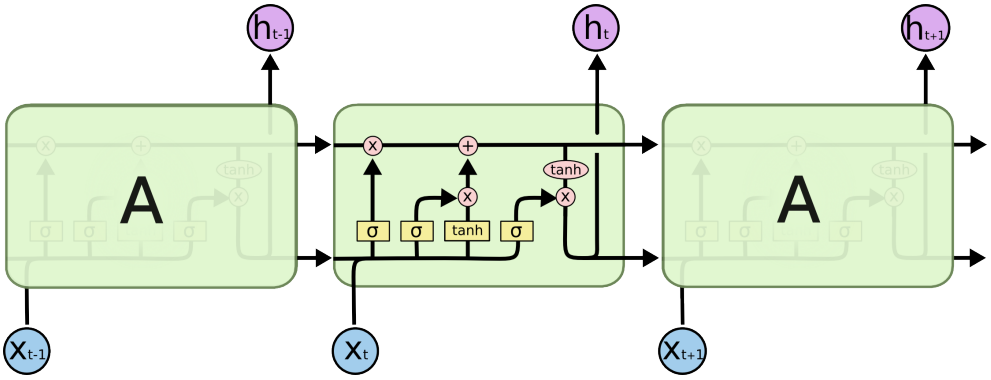


Figura 11: Representação dos três gates LTSM.

Uma das questões centrais da LSTM é a criação do estado de célula. Na Figura 12 visualiza-se uma linha horizontal que representa esse estado.

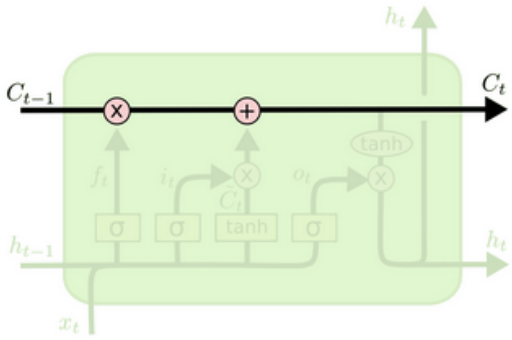


Figura 12: Linha com estado da célula.

Essa linha com a representação do estado da célula recebem as informações que foram calculadas em cada um dos *gates* LSTM.

A seguir, na Figura 13, visualiza-se o primeiro *gate* que é responsável por decidir o que será adicionado ao estado da célula.

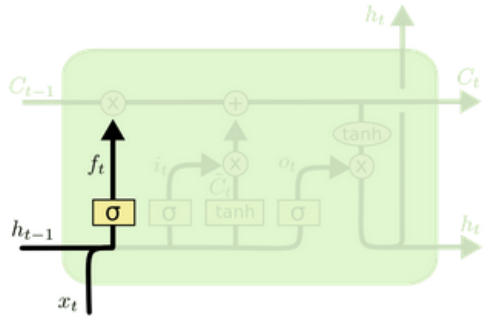


Figura 13: *Primeiro gate LSTM: forget gate.*

O *forget gate* recebe dados de entrada aplica função de ativação sigmóid e leva este resultado até a linha que guarda informações de estado da célula.

O próximo *gate*, representado na Figura 14, é o responsável por decidir o que será adicionado ao estado da célula.

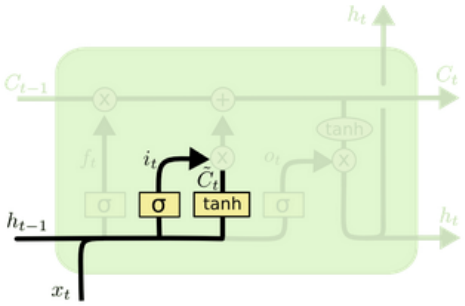


Figura 14: *Segundo gate LSTM: input gate.*

O *input gate* recebe dados de entrada aplica funções de ativação sigmóid e tangente, após estes cálculos levam as informações que devem ser adicionas na linha responsável pelo estado da célula.

Na sequência o último *gate*, representado na Figura 15, é o responsável por decidir qual será a saída do neurônio.

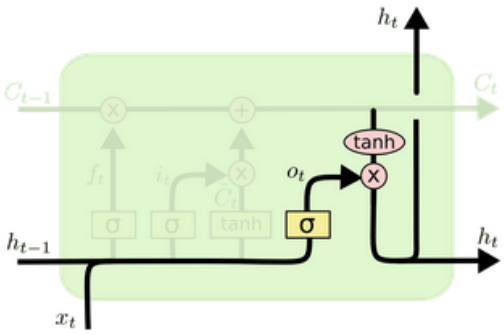


Figura 15: *Terceiro gate LSTM: output gate.*

O *output gate* por meio da função de ativação sigmoíd decide quais partes do estado da célula devem ser considerados, a partir disso aplica-se a função tangente colocando os valores entre -1 e 1. Ou seja, os cálculos serão efetuados apenas em informações que foram consideradas importantes. Por exemplo, apenas variáveis que possuem correlação com uma série que deseja-se prever.

Uma das aplicações das Redes Neurais LSTM é a interpretação de imagens, de modo que consegue descrever os respectivos conteúdos atingindo expressivos resultados (BROWNLEE, 2018b). No ano de 2014, foram criados vários algoritmos *deep learning* que foram capazes de detectar e descrever objetos em fotografias, de modo que foram criadas legendas coerentes com as imagens (KRAUSE et al., 2017).

Outra aplicação de LSTM é na tradução automática de textos. Por exemplo, é possível colocar na entrada sentenças em inglês e obter na saída sentenças em francês. As *Long-Short Term Memory* são capazes de aprender a tradução das palavras. (BROWNLEE, 2018b; SUTSKEVER et al., 2014).

Por outro lado, Redes Neurais LSTM possuem limitações como pequena janela de informação passada. Então, não possui capacidade de resolver todos os problemas em previsão de sequência, algo que já foi relatado na literatura e em testes de *benchmark* (GERS et al., 2002).

Segundo Brownlee (2018b) existem cinco tipos de LSTM: Vanilla, Stacked, Bidirectional,CNN-LSTM e ConvLSTM.

1. Vanilla LSTM é o tipo mais simples de *Long-Short Term Memory*, pois possui uma camada de entrada, apenas uma camada oculta com unidades LSTM, e uma camada de saída, que serve para fazer previsão, conforme Figura 16.

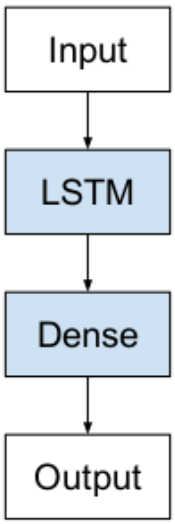


Figura 16: Modelo Vanilla LTSM.

O termo "dense" pode ser observado na figura anterior e é uma referência de como os neurônios estão interligados, isto é, uma neurônio se conecta a todos os outros da camada seguinte, esse tipo de ligação é conhecida como "fully connect".

2. Stacked LSTM é o nome do modelo que faz empilhamento de camadas LSTM uma sobre a outra, que pode ser visualizado na Figura 17, em que é possível observar duas camadas LSTM após a camada de entrada. O tipo stacked necessita de uma entrada de dados tridimensional (informações de quantidade de linha, coluna e valores anteriores), e produz uma saída é bidimensional (informações de

linha e coluna) no final da sequência.

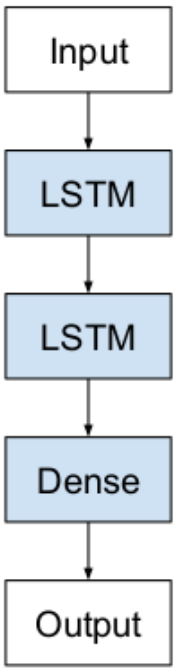


Figura 17: Modelo Stacked LTSM. Fonte: Elaborado pelo próprio autor.

3. Bidirectional LSTM permite que a LSTM aprenda a sequência em dois sentidos: entrada para frente e da frente para trás, isto é, esquerda para direita (So até Si) e direita para esquerda (S'o até S'i), conforme Figura 18. É utilizada em previsão de séries temporais univariadas, de modo que a primeira camada oculta trabalha de forma bidirecional.

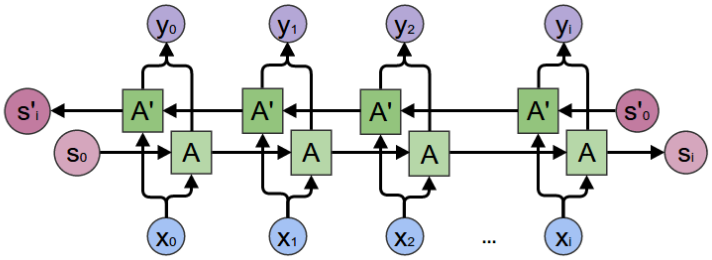


Figura 18: Bidirectional LTSM. Fonte: Elaborado pelo próprio autor.

Os modelos a seguir não foram utilizados nesta pesquisa, estão listado apenas para registro. O próximo modelo deriva de *Convolutional Neural Network* (CNN).

4. CNN LSTM é um tipo de Rede Neural que trabalha com dados de imagens bidimensionais. As CNN podem ser eficazes para aprender automaticamente

dados de sequência unidimensional. A CNN pode ser usada como um modelo hibrido, atuando de forma conjunta com LSTM no *backend*, modelo denominado CNN-LSTM.

5. modelo ConvLSTM é relacionado ao CNN-LSTM, de modo que a leitura convolucional da entrada é construída em cada unidade LSTM. ConvLSTM foi pensada para fazer leitura de dados espaço-temporais de duas dimensões, mas também permite adaptação para trabalhar com previsão univariada de séries temporais. A camada espera uma entrada que deve conter uma sequência de imagens bidimensionais. O formato dos dados da entrada deve ser: amostras, estados de tempo, linhas, colunas, recursos.

Redes Neurais Convolucionais são um tipo de RNA muito utilizada no reconhecimento de imagens. A saída é uma imagem, com as mesmas dimensões, que é comparada com a entrada. Este tipo de Rede Neural é considerada um grande avanço tecnológico, porém esbarra em dificuldades como a necessidade de grande bancos de dados rotulados (COSTA, 2017).

2.3 Estatística

Vários algoritmos já foram criados para realizar tarefa de predição em dados sequenciais (GUENICHE et al., 2013). Modelos matemáticos têm sido usados para auxiliar na tomada de decisão em diversas situações, por exemplo, na área de saúde, de modo que é possível prever casos de dengue (MIRANDA et al., 2015).

Modelos são representações da realidade em estudo, ao passo que se dá ênfase para características maiores e, muitas vezes, se despreza características que podem ser consideradas detalhes menores. O objetivo é descrever, simplificar e facilitar o entendimento sobre aquilo que se está estudando. (BARBETTA et al., 2010).

Modelos lineares tentam descrever fenômenos aleatórios buscando encontrar a normalidade procurada. A partir da década de 1970, o desenvolvimento computacional permitiu a criação de novos modelos com processos iterativos, por exemplo, modelo normal não linear e modelos lineares não generalizados. Nos anos posteriores surgiram outros modelos como quase-verossimilhança, dispersão e modelos multivariados (PAULA, 2004).

É possível exemplificar outros modelos, como o probabilístico ou estocástico, em que antes do experimento não é possível saber o resultado, mas somente a probabilidade. Um exemplo é o lançamento de uma moeda. Antes do lançamento

é possível atribuir 50% de chance para cara e 50% para coroa, na face superior após o lançamento (BARBETTA et al., 2010).

Em uma análise de dados é comum buscar similaridades, diferenças e correlação entre variáveis (LIRA; NETO, 2006). Para tanto, é necessário recorrer a área de Estatística (BUSSAB; MORETTIN, 2009). Uma das técnicas disponíveis para buscar comportamento próximo entre variáveis é a correlação linear de Pearson, que é capaz de medir o grau da relação linear entre duas variáveis (PRUNZEL et al., 2016).

Em Estatística, regressão linear é uma equação para estimar valor esperado de uma variável. Isto é, trata-se de uma busca por estimar um valor que não é possível inicialmente. Dessa forma, é um modelo matemático que estuda a relação, de forma linear, entre uma variável dependente e uma ou mais variáveis independentes. Mais que isso, em alguns estudos essa técnica tem sido usada para fazer predição de dados (SOUZA et al., 2018).

Outra técnica usada para análise de dados é o coeficiente de determinação (R^2), também conhecido como R quadrado, que possui valores entre 0 e 1, e é capaz de determinar a proximidade entre valores de duas variáveis. O R^2 é usado também como parâmetro de escolha em modelagens matemáticas (SOUZA et al., 2018), por exemplo, é utilizado para testar um conjunto de dados em diversos modelos e analisar qual atinge valor mais próximo de 1, dessa forma, apontando qual o modelo que atinge melhor resultado (SEVERINO et al., 2005).

O modelo R quadrado, então, é uma medida que calcula o grau de ajustamento da equação em relação aos dados amostrais. Isto é, mede a qualidade do ajuste de um modelo, de modo que é capaz de testar o grau de significância em um modelo de regressão linear, e pode ser traduzido, em formato mais didático, como um índice percentual (QUININO et al., 1991).

Existem outros testes e técnicas que podem se somar ao coeficiente de determinação em uma análise de dados. Um exemplo é o teste de Wilcoxon, um tipo de teste que usa dados amostrais em série e considera o tamanho das diferenças. É caraterizado como teste não paramétrico, ou seja, não exige que a população analisada tenha distribuição normal. É um tipo de teste que usa dados amostrais em série e considera o tamanho das diferenças

Com o Teste de Wilcoxon verifica-se a hipótese nula de que duas amostras pareadas foram selecionadas de duas populações idênticas ou de duas populações simétricas coma mesma média. Então, testa se a distribuição das diferenças dos escores na população é simétrica em relação ao valor zero (HOWELL, 2012).

Outra opção para análise de dados é o coeficiente de correlação de Ken-

dall, que pode variar entre -1 e 1. Quanto maior o coeficiente maior a associação das amostras, bem como valores negativos indicam associação negativa. É utilizado para testar hipóteses que verificam associação.

É um tipo de teste não paramétrico e verifica a semelhança entre ordens de dados. É usado, por exemplo, na comparação entre variáveis. Outro uso comum é para verificar se duas variáveis são dependentes. É também usado para checar a consistência (SILVA et al.,).

Para verificar correlações uma opção para analisar os dados é o coeficiente de correlação de Spearman, que é a mais antiga estatística baseada em postos (SIEGEL; JR, 1975). É uma medida que verifica um tipo de correlação não paramétrica. Não requer que o relacionamento entre as variáveis seja linear, e não requer que as variáveis sejam quantitativas. O coeficiente de Spearman não necessita que a população amostrada tenha distribuição normal, o que o diferencia do Coeficiente de Pearson (BAUER, 2007).

Assim como o coeficiente de correlação de Kendal o coeficiente de Spearman trabalha com o conceito de postos. É usado para medir a intensidade da relação entre as variáveis ordinais. Neste caso, mais importante que o valor é a ordem das observações. Dessa forma, pode ser utilizado mesmo com a presença de assimetrias e *outliers*.

Existem outras estatísticas que servem inclusive para trabalhar diretamente com predições. Um exemplo é o ARIMA.

Integrated Moving Average (ARIMA) é o nome de outro modelo utilizado em predição, em especial, ao trabalhar com séries temporais. Esse modelo ajuda no entendimento dos dados, e , algumas vezes, pode ser mais indicado que Redes Neurais (BROWNLEE, 2018b). Em outros casos, essas duas técnicas conseguem trabalhar juntas formando um modelo hibrido (CARDOSO; CRUZ, 2016). Ocorre que para suportar o comportamento de sazonalidade foi criado o modelo SARIMA (*Seasonal Autoregressive Integrated Moving Average*).

O modelo SARIMA é também usado para predição, capaz de capturar a sazonalidade em séries temporais, de modo que contém uma parte sazonal e outra não sazonal (ESPINOSA et al., 2010). É uma técnica que já foi usada em situações como: prever a temperatura (SILVA et al., 2008), o número de focos de calor (ESPINOSA et al., 2010), número de vendas de motocicletas (WALTER et al., 2013) e a demanda do setor calçadista (LIMA et al., 2019).

No estudo de predições um dos fatores que dificultam a realização de testes é o custo computacional. Dessa forma, é possível utilizar técnicas que contribuam para uma exigência menor de processamento computacional.

Capítulo 3

Materiais e Métodos

Neste capítulo apresenta-se os dados que foram utilizados. Dessa forma, ocorre uma apresentação de cada variável que foi usada como base de dados para treinamento da Rede Neural Artificial, bem como apresenta-se de que forma ocorreu o desenvolvimento da Rede Neural LSTM.

3.1 Dados Utilizados

Para esta pesquisa optou-se por utilizar dados coletados pelo Programa de Pós-Graduação em Física Ambiental (PPGFA). Essa escolha considerou a quantidade de varáveis microclimáticas coletadas, além de dados da cobertura do solo. Outro fator que contribuiu foi a continuidade da coleta de dados, que iniciou em 2011, depois 2012, depois 2016 e continua em 2019. Esses fatores contribuíram para formação de grandes volumes de dados, o que é pre-requisito para, por exemplo, aplicar técnicas de Redes Neurais.

A coleta de dados foi realizada utilizando a metodologia de transecto móvel noturno. Foi utilizado veículo automotor. A coleta foi feita às 20 horas da noite. A velocidade do veículo variando entre 30 e 40 quilômetros por hora (km/h), em dias com condições climáticas estáveis, com ventos leves e céu claro (OKE, 2004; AMORIM, 2005; SANTOS, 2012).

3.1.1 Sensores

Neste estudo utilizou-se Datalogger de temperatura e umidade do ar (HT-500, Instrutherm, São Paulo, São Paulo, Brasil), o equipamento foi protegido por abrigo, acoplado na lateral de um veículo a aproximadamente a 2 metros do solo, conforme Figura 19. Ao mesmo tempo, foi utilizado equipamento GPS/MAP da

marca Garmin, modelo 76Cx, obtendo-se as coordenadas geográficas angular e UTM (Universal Tranversa de Mercator), Zona 21, para precisão dos locais de medição dos pontos (SANTOS, 2012).

Figura 19: Sensor utilizado no transecto. Adaptado de Zavitoski, 2018.

O sensor possui um abrigo que foi desenvolvido por pesquisadores do PPGFA. O abrigo é constituído por cano de material PVC (ZAVITOSKI, 2018).

3.1.2 Localização

Devido a extensão da área de estudo em Cuiabá-MT, optou-se por dividir em dois transectos, sendo o transecto 1 - Leste/Oeste com 19,76km e o transecto 2 - Sul/Norte, com 11,6km, totalizando 91 pontos, Figura 20. Esta metodologia consiste em um percurso previamente determinado, registrando as medições das variáveis climáticas, temperatura do ar e umidade relativa do ar, nas quatro estações do ano, durante os anos de estudo.

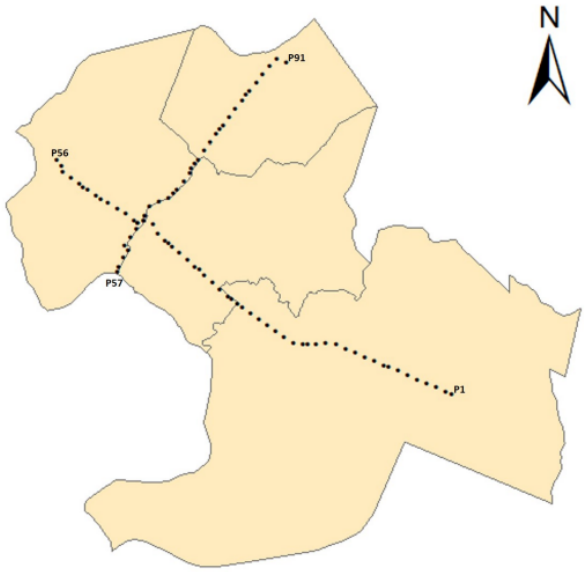


Figura 20: Área de estudo com 91 pontos em Cuiabá-MT.

3.1.3 Período

Os dados utilizados nesta pesquisa, se referem ao período de 2011-2012 e 2016. Para as coletas realizadas em 2011 e 2012 (Tabela 2).

Tabela 2: Período de medição por transectos móveis em 2011/2012.

Mês	Dia	Estação
Abril/2011	12/19/26	Outono
Maio/2011	09/17/26	Outono
Junho/2011	14/16/21/28	Outono/Inverno
Julho/2011	07/12/21/26	Inverno
Agosto/2011	08/16/29/30	Inverno
Setembro/2011	06/19/25/28	Inverno/Primavera
Outubro/2011	24	Primavera
Novembro/2011	09/23	Primavera
Fevereiro/2012	28	Verão
Março/2012	18/19	Verão

Para a coleta de dados realizada em 2016 (Tabela 3), utilizou-se Data-logger com GPS, modelo GKV02. Obtendo as coordenadas geográficas angular e posteriormente convertendo em UTM (Universal Tranversa de Mercator) utilizando Datum WGS84 Zona 21S, para precisão dos locais de medições dos pontos. Também se utilizou o RTL-10709 sensor de temperatura do ar e umidade relativa do ar, modelo AM2302, protegido por abrigo, acoplado na lateral de um veículo, fixado a aproximadamente 2 metros do solo.

Tabela 3: Período de medição por transectos móveis em 2016.

Mês	Dia	Estação
Março/2016	04/06/18/20	Verão
Maio/2016	11	Outono
Junho/2016	06/11/12	Outono
Julho/2016	12	Inverno
Agosto/2016	06	Inverno
Setembro/2016	10/15	Inverno
Outubro/2016	28/29	Primavera
Dezembro/2016	11/20	Primavera

Os dados de 2011 foram coletados de 10 em 10 segundos. Os dados de 2016 foram coletados de 1 em 1 segundo. Cada linha com o conjunto de informações das variáveis é chamada de *timestamp*.

Na sequência deste estudo os dados de 2011-2012 serão referenciados apenas como dados 2011.

3.1.4 Classificação da cobertura de solo

O conjunto de variáveis foi escolhido considerando o potencial de interferência na temperatura do ar e consequentemente na formação de ilhas de calor (SANTOS; NOGUEIRA, 2012).

Para a classificação do solo dos 91 pontos, utilizou-se a metodologia de OKE (2004), que defini a área de influência do instrumento, sendo a relação de a cada 2m de distância que o instrumento esta do solo, a área de influência é 200 metros de raio. Com isto, foram gerados os mapas temáticos empregando o método de classificação supervisionada, por meio da técnica MAXVER (máxima semelhança), utilizando imagens do software Google Earth. Obtendo as porcentagens referentes a cada classe de interesse, conforme Kastzschner (1997), isto é, cobertura com vegetação rasteira, vegetação arbórea, solo exposto, corpos d'água, área pavimentada e área edificada, nos anos de estudo em 2011-2012 e 2016.

3.2 Desenvolvimento do Modelo Deep Learning

3.2.1 Ambiente de Desenvolvimento

IDE (*Integrated Development Environment*) é um *software* que funciona como um ambiente integrado para o desenvolvimento de códigos, neste estudo

optou-se pelo Spyder[1]. E a linguagem computacional escolhida foi o python[2], na sua versão 3.6. Essa linguagem foi escolhida pela quantidade de bibliotecas disponíveis e pela documentação disponível na *web*.

O banco de dados utilizado foi o PostgreSQL[3], na sua versão 10.3, a escolha foi feita considerando a documentação disponível, o desempenho e sua flexibilidade(CARVALHO, 2017).

O processamento foi realizado em servidores da UFMT. Outra opção seria utilizar nuvens como da Google[4] e nuvem Microsoft[5].

O equipamento utilizado para processar os dados foi um servidor IBM, modelo x3650 M4, com 64 GigaByte *Random Access Memory* (GB RAM), disco do tipo *solid-state drive* (SSD), virtualizador VmWare Sphere[6] 6.5. É um equipamento do parque computacional da UFMT. Foram utilizadas 5 máquinas virtuais com a mesma configuração e processando os dados de forma simultânea, uma para cada estação do ano. A especificação segue abaixo:
-Sistema Operacional: Ubuntu 18.04
-virtualCPU: 10 unidades
-Memória: 10 GB RAM
-Disco: SSD 30 GB

Para atividades mais rápidas (poucas horas), como testes e validação de código, optou-se por utilizar o ambiente de Computação em Nuvem chamado COLAB[7], da empresa Google. O serviço é gratuito e seu uso é limitado para até 12 horas de *runtime* (uma rotina de execução).

3.2.2 Escolha da Rede Neural

A escolha da Rede Neural LSTM considerou o levantamento bibliográfico realizado no Capítulo 2. As bibliografias consultadas indicaram que a *Long-Short Term Memory* tem alcançado ótimos resultados em predição de séries de dados por ter características como a capacidade de persistência da memória e reconhecimento de padrões em sequência de dados. O fator que mais colaborou para escolha foi a capacidade de persistência da memória, visto que a LSTM é capaz de guardar informações a respeito de neurônios que estão vários passos atrás no desenrolar da Rede Neural.

[1] https://www.spyder-ide.org
[2] https://www.python.org
[3] https://www.postgresql.org
[4] https://cloud.google.com/Google/ComputeEngine
[5] https://azure.microsoft.com/pt-br/education
[6] https://www.vmware.com/vsphere
[7] https://colab.research.google.com

3.2.3 Seleção das Variáveis

A seleção de variáveis ambientais foi feita considerando todas as opções que foram captadas e tinham influência na temperatura do ar (SANTOS; NOGUEIRA, 2012), e consequentemente na formação de ilhas de calor no período noturno.

Segue abaixo a lista de variáveis coletadas nos transectos:

- Temperatura do ar
- Data
- Longitude
- Latitude
- Hora
- Umidade Relativa do ar
- Altitude
- Cobpais-pct (cobertura paisagística)
- Cobarb-pct (cobertura arbórea)
- Soloexp-pct (solo exposto)
- Apav-pct (área pavimentada)
- Aedf (área edificada)
- Aguapct (corpos d água)

Neste estudo optou-se por não utilizar as variáveis Data, Hora, Latitude e Longitude.

3.2.4 Preparação dos Dados

A preparação dos dados envolveu organiza-los para servir de entrada para Rede Neural LSTM. Neste estudo descartou-se linhas com dados incompletos em alguma coluna. A variável "velocidade" foi retirada por motivo de ser uma referência ao veículo que transportava o Datalogger.

Os dados que estavam em arquivos de planilha eletrônica foram transferidos para o PostgreSQL. Desse modo, foi possível fazer manipulações utilizando

código *Structured Query Language* (SQL), linguagem de consulta estruturada. Esse fluxo de atividades foi passado para o formato de um diagrama.

O modelo de diagrama denominado *Business Process Modeling Notation* (BPMN) é muito utilizado para modelar processos e fluxos de atividades, e nesse estudo foi utilizado para explicar como foi feita a modelagem do banco de dados, conforme pode ser observado na Figura 21.

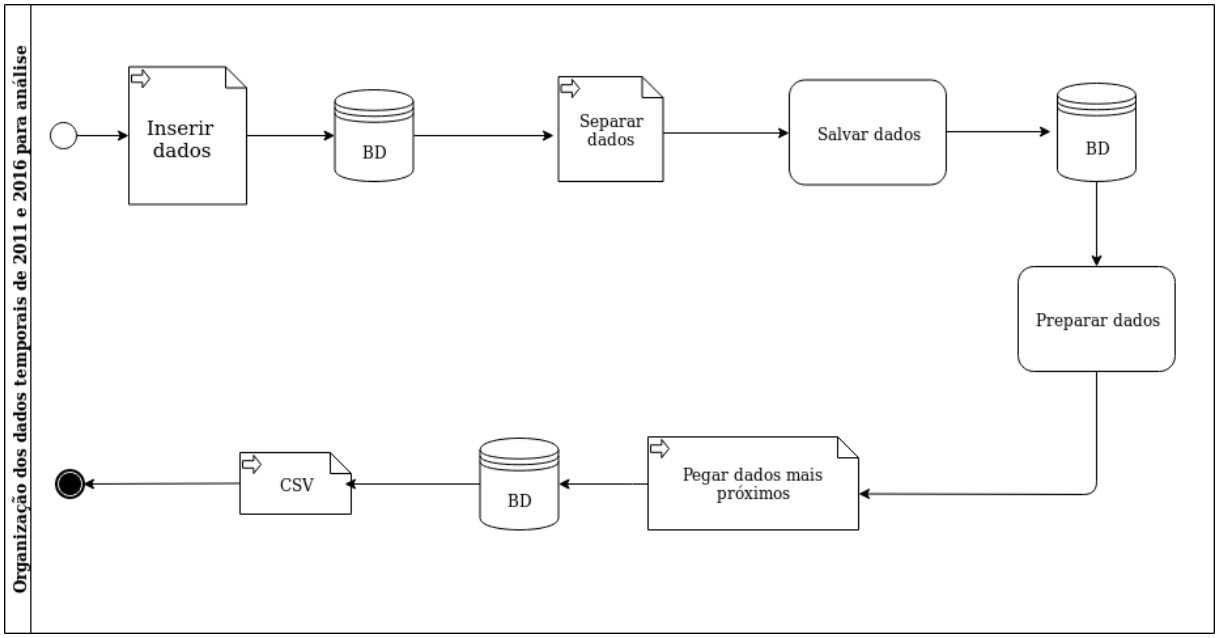


Figura 21: Conjunto de etapas na manipulação do banco de dados.

Abaixo estão as descrições com maiores detalhes das etapas que estão representadas na Figura 15:

1. Na primeira etapa os dados do *datalogger* foram inseridos no banco de dados PostgreSQL.

2. A seguir foi utilizado um *script* para separar e salvar os dados por estações: inverno, outono, primavera e verão.

3. Na etapa seguinte utilizou-se código SQL para gerar um novo banco de dados com os pontos mais próximos de um determinado ponto, por exemplo, um ponto de referência no ano de 2016 e os três pontos mais próximos de 2011. Também foram criados arquivos com pontos de referência 2016 juntamente com os três pontos mais próximos de 2016.

4. Na sequência os bancos criados foram exportados para arquivo CSV (*Comma Separated Value*).

5. Nesta etapa os arquivos foram tratados para servir de entrada para RN, então, foram deletadas as colunas: Data, Hora, Latitude e Longitude.

Depois de todas as manipulações no banco os dados foram organizados em grupos de variáveis com informações sobre cada *timestamp*, por exemplo, no primeiro caso (utilizar dados de 2011 para prever 2016) o primeiro grupo de variáveis continha informações sobre determinado ponto de 2016, conforme pode ser observado na Tabela 4:

Tabela 4: Grupo de variáveis de um determinado ponto.

Grupo 1								
v1	v2	v3	v4	v5	v6	v7	v8	v9
32.3	42.2	197	19	8	15	23	33	0
32.2	42.2	197	19	8	15	23	33	0
32.2	42.3	197	19	8	15	23	33	0

Abaixo a relação do nome das variáveis abreviadas com seus respectivos nomes:

- v1: Temperatura
- v2: Umidade
- v3: Altitude
- v4: Cobpais-pct
- v5: Cobarb-pct
- v6: Soloexp-pct
- v7: Apav-pct
- v8: Aedf
- v9: Aguapct

O conjunto v1 a v9 compõe a seleção de variáveis utilizadas. As três linhas representam um resumo de um banco de dados inteiro. O conjunto completo de dados é composto por arquivos com até vinte (20) mil linhas.

Os três grupos seguintes (após grupo de 2016) eram dados de 2011 com informações sobre os três pontos mais próximos. Essa busca espacial no transecto foi feita usando um cálculo de proximidade, que considera a distância espacial juntamente com a temporal, sendo que a primeira tem prioridade maior na classificação dos dados. Abaixo a representação desse cálculo de distância:

$$D_i = \sqrt{(L1 - L2)^2 + (l1 - l2)^2} + |t1 - t2| \quad (1)$$

L = longitude
l = latitude
t = tempos em segundos

Por meio dessa equação foi possível verificar a distancia geométrica dos pontos, por exemplo, com informações de 2016, conforme Tabela 5:

Tabela 5: Quatro conjuntos de variáveis que servem para dar entrada na Rede Neural. Exemplo de um *timestamp*.

Grupo 1: Dados de 2016								
v1								
30.2								
Grupo 2: Dados do ponto mais próximo 2011								
v1	v2	v3	v4	v5	v6	v7	v8	v9
36.9	208	21	6.3	10.2	35.2	26.5	0	
Grupo 3: Dados do segundo ponto mais próximo 2011								
v1	v2	v3	v4	v5	v6	v7	v8	v9
32.2	37.0	207	21	6.3	10.2	35.2	26.5	0
Grupo 4: Dados do terceiro ponto mais próximo 2011								
v1	v2	v3	v4	v5	v6	v7	v8	v9
32.2	37.1	207	21	6.3	10.2	35.2	26.5	011

Depois de todos os tratamentos dos dados formou-se, então, os quatro grupos de variáveis que compõem um total de colunas, de modo que cria-se com isso o formato que a Rede Neural precisa como entrada para fazer os cálculos e realizar as predições.

3.2.5 Amostragem

O conjunto de dados utilizado foi dividido para realização dos testes de predição. Os dados foram mantidos na sequência original.

Amostragem - 90% para treinamento e 10% de teste.

A parte de teste utilizada representa o pedaço final dos *timestamps*, isto é, do conjunto inteiro de dados analisado em cada caso.

Neste estudo optou-se por não utilizar técnica de validação cruzada (*cross-validation)* como k-fold, pois isto elevaria de forma exponencial o tempo de processamento dos dados.

3.2.6 Open Science

Este estudo orientou-se por praticas de ciência aberta (*open science*) disponibilizando as informações e dados utilizados. O sistema operacional utilizado (Linux Ubuntu), a linguagem de computação, as bibliotecas e ferramentas utilizadas seguem a linha *open source*.

3.2.7 Bibliotecas Python

Biblioteca é o nome de um módulo que possui conjuntos de *scripts* que facilitam a realização de atividades específicas, por exemplo, manipular gráficos. O modelo de Rede Neural LSTM construído utilizou as seguintes bibliotecas (pacotes) da linguagem python:

- Tensorflow[8] - permite trabalhar com *Machine Learning*.
- Keras[9] - facilita o trabalho com redes neurais.
- Pandas[10] - possibilita realizar analise de dados.
- Numpy[11] - permite manipular matrizes e vetores.
- MatplotLib[12] - facilita a criação e visualização de gráficos.
- Sklearn[13] - faz mineração e análise de dados.
- Math[14] - possibilita uso de funções matemáticas.

Todas as bibliotecas utilizadas estavam em versão compatível com python 3.6. O modelo criado recebeu 28 variáveis na entrada da rede neural, já a saída possuí apenas um neurônio, mas isto também pode ser alterado, conforme o que se pretende prever.

[8] https://www.tensorflow.org
[9] https://keras.io
[10] https://pandas.pydata.org
[11] https://numpy.org/
[12] https://matplotlib.org/
[13] https://scikit-learn.org
[14] https://docs.python.org/3/library/math.html

3.2.8 Criação do Modelo LSTM

Para criação do modelo foram testadas diferentes combinações, dessa forma, foi possível salvar o melhor modelo, ou seja, a combinação dos parâmetros que atingiu maior valor de Coeficiente de Determinação (também conhecido por R^2), isto é, valor mais próximo de 1, quanto mais próximo do valor 1 mais próximo é o conjunto de dados previstos em comparação com dados reais. O melhor modelo foi salvo em arquivo .h5, que guardou os pesos utilizados na Rede Neural.

As camadas ocultas são grupos de neurônios que ficam entre a camada de entrada e saída. Os neurônios são células de memória. O batchSize define o número de amostras utilizadas antes da atualização dos parâmetros internos do modelo. Época é a quantidade de passagens em ciclos completos por meio do conjunto de dados para treinamento.

A quantidade de "valores anteriores" em uma previsão é uma variável que representa o número de registros exatamente anteriores a um dado que pertence a uma série, por exemplo, se o valor dessa variável é 300, só é possível prever usando LSTM a partir do registro (*timestamp*) 301 da série.

O modelo de RN criado, tipo LSTM, considerou os seguintes parâmetros e valores para realização dos testes:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 32 e 64.
- Tipo de LSTM: 1 (Stack) e 2 (Bidirectional).
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

Os dois tipos de LSTM testados são os que se encaixam nas definições de *Deep Learning* já apresentadas.

Dessa forma, existiam quatro (4) combinações diferentes para cada estação. É importante ressaltar que ao elevar as combinações, o número de camadas, neurônios, épocas e "valores anteriores" também eleva-se o custo computacional em vários sentidos, isto é, vai ter uma necessidade maior de memória, CPU (*Central Process Unit*), espaço de armazenamento e tempo para processamento. Para cada combinação de parâmetros salvou-se as informações estatísticas: RMSE, coeficiente de determinação, correlação de Spearman, correlação de Kendall e teste de Wilcoxon.

Na Tabela 6 é possível verificar um exemplo de combinação dos principais parâmetros utilizados ao criar uma Rede Neural LSTM:

Tabela 6: Exemplo de combinação dos parâmetros de treinamento da Rede Neural

Parâmetro	Valor
Épocas	100
Neuronios	100
Camadas Ocultas	4
Batch Size	32
Timestamps anteriores	300
Tipo de LSTM	1 (Stack)

Na Figura 22 é possível visualizar, ainda que de forma resumida utilizando diagrama, o fluxo de etapas na elaboração do algoritmo.

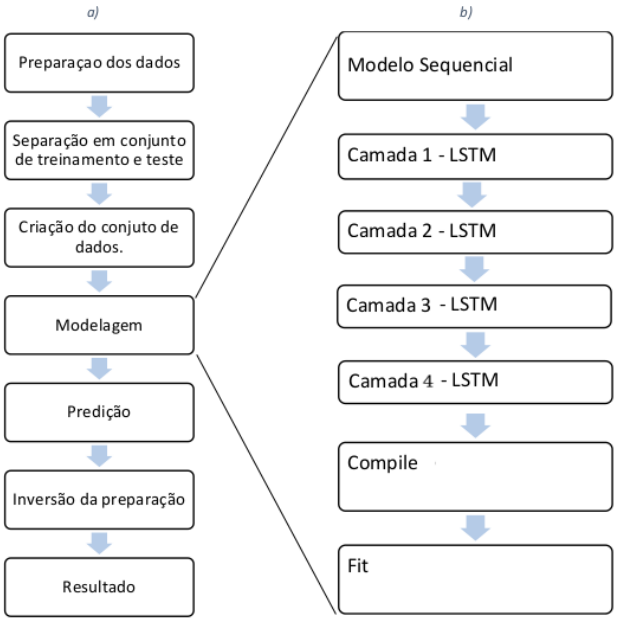


Figura 22: Fluxo de etapas. Fonte: Adaptado de Ferrari et al. 2019.

A etapa denominada "modelagem" envolve etapas menores como a definição do tipo de modelagem e a configuração da quantidade de camadas ocultas. O tipo de modelo usado foi o "Sequential", que é constituído por uma pilha linear de camadas, no caso camadas LSTM. A seguir a etapa de compilação em que ocorre aplicação da função de erro. Depois a etapa "fit" em que ocorrem as iterações definidas pelo número de épocas.

Existe também um diagrama que pode ser gerado pela própria Rede Neural. Na Figura 23 visualiza-se um diagrama simplificado da etapa específica

de modelagem da Rede Neural, tipo LSTM. O diagrama foi baseado no sumário criado (Anexo) pelo pacote Keras.

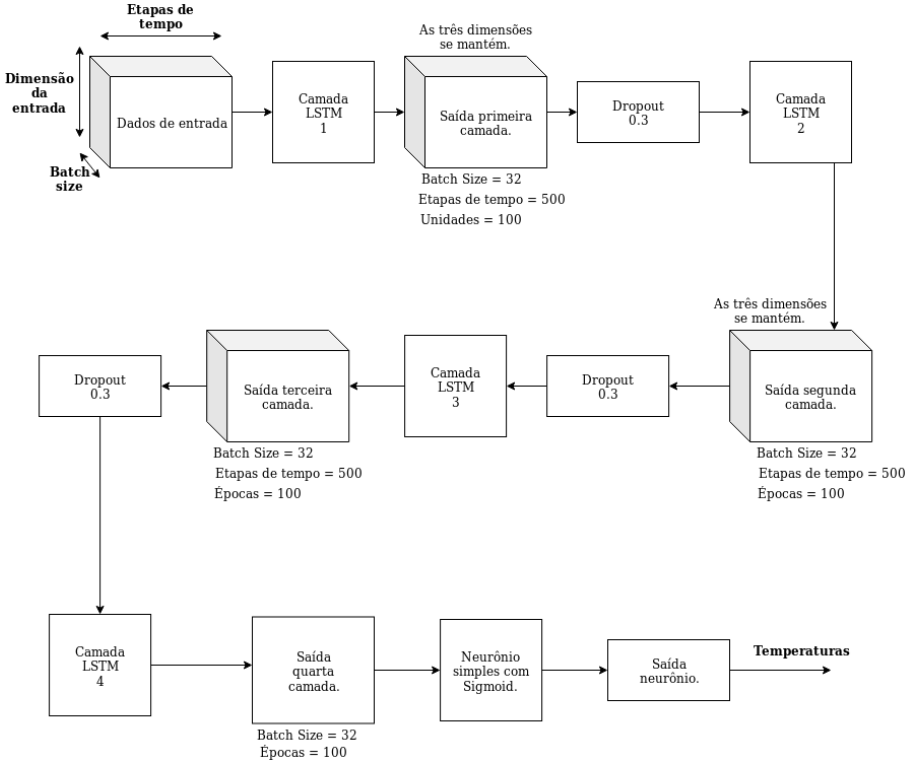


Figura 23: Diagrama com formato do modelo de entrada do Keras.

Na imagem gerada pelo pacote Keras com o sumário da Rede Neural LSTM é possível verificar as etapas da rede, informações sobre parâmetros e o uso do *dropout*. O *dropout* é um recurso de regulação em que parte dos dados são descartados de forma randômica para evitar *overfitting*, que é um tipo de ajustamento forçado da Rede Neural aos dados utilizados. Então, nesta pesquisa optou-se por utilizar um *dropout* de 30%.

3.2.9 Modelos criados

Em cada conjunto de testes foi salvo o melhor modelo, isto é, aquele que atingiu maior R^2. Criou-se modelos com os respectivos melhores resultados de coeficiente de determinação para as estações: inverno, outono, primavera e verão. E cada modelo foi testado nas outras estações. Os dados reais foram comparados com os previstos utilizando o Coeficiente de Determinação.

Os modelos criados guardaram informações de pesos da Rede Neural, o que é necessário para utilizar essa modelagem em outro conjunto de dados, em alguns casos como na previsão da primavera de 2016 o arquivo .h5 guardou infor-

mações sobre mais de oitenta mil (80000) parâmetros. Esse tipo de informação é armazenada pelo pacote Keras do Python.

Na simulação de temperatura para o ano de 2021, utilizou-se modelos criados com dados de 2011 e aplicados em dados de 2016. As escolha do ano de 2021 foi feita para seguir o mesmo intervalo de 5 anos. Nessa simulação foram usados modelos e dados de referência da mesma estação, por exemplo, no modelo gerado a partir de dados da primavera 2011 foram aplicados dados de primavera 2016 para simular a temperatura da primavera em 2021.

Capítulo 4

Apresentação e Análise dos Resultados

O capítulo de resultados foi dividido em quatro partes, a primeira parte apresenta-se o cenário com dados de 2011 para prever 2016. Na segunda etapa visualiza-se dados de uma estação de 2016 para prever as demais estações do mesmo ano. Na parte seguinte apresenta-se a simulação para o ano de 2021. Por fim,

4.1 2011 prever 2016

Na criação de modelos com dados de 2011 para prever 2016 utilizou-se dados de cada estação: inverno, outono, primavera e verão. Dessa forma, os dados foram separados por estação. Neste sentido, foram testadas as configurações informadas anteriormente e a combinação que obteve melhor resultado considerando o Coeficiente de Determinação serviu para criar o modelo de cada estação.

4.1.1 Amostragem 90% de treino e 10% de teste

A primeira amostragem testada considerou os primeiros 90% de dados para treinamento e os 10% restantes para teste, isto é, comparação entre dados reais e previstos.

4.1.1.1 Modelo Inverno

Na criação do modelo de inverno a melhor combinação obteve R^2 0,9725, MSE 0,0067 e RMSE 0,082 nos dados de teste, e teve a seguinte configuração:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 32.
- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

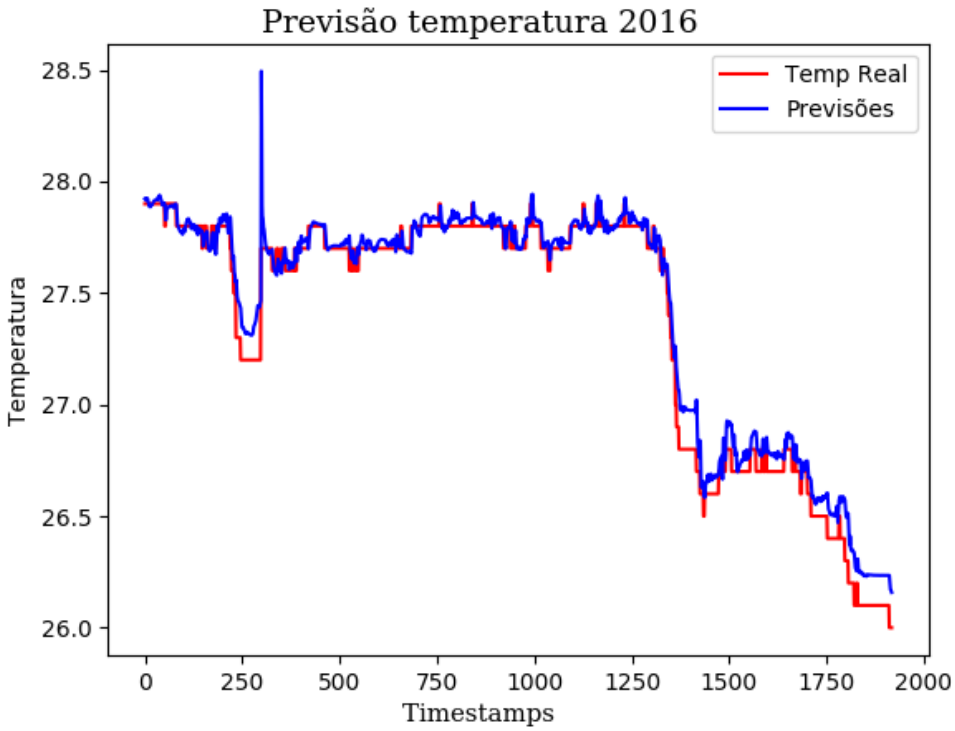
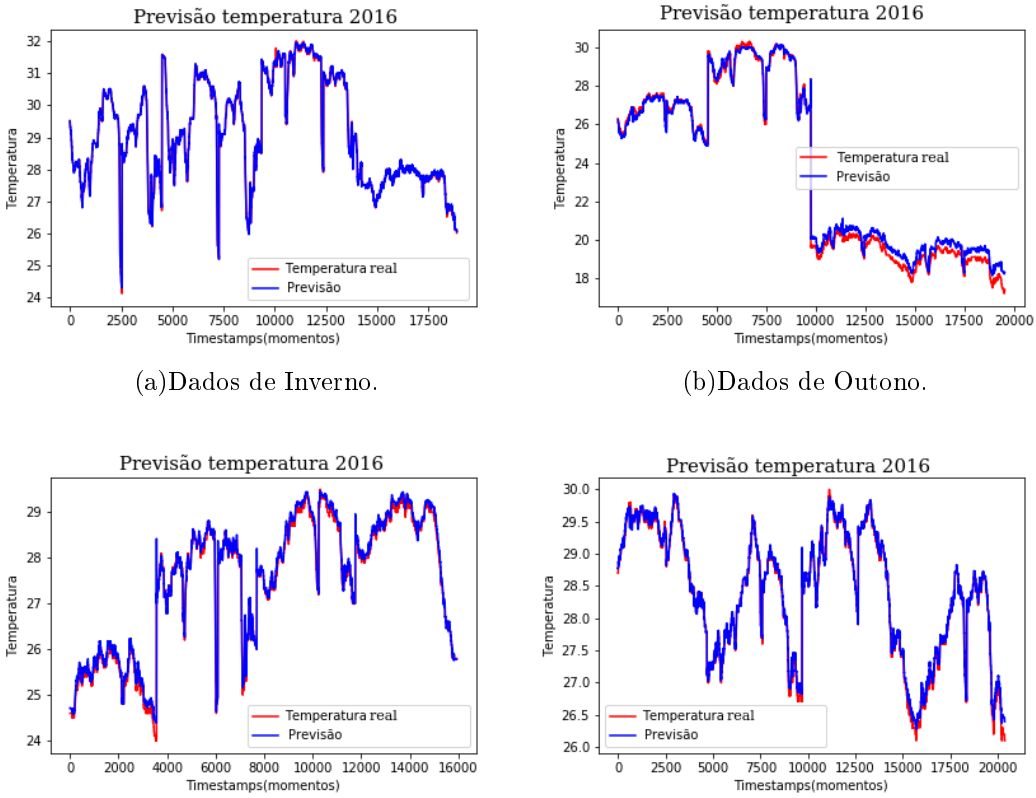


Figura 24: Gráfico com melhor modelo de inverno 2011 para prever inverno 2016.

É possível verificar na Figura 24 que a predição, linha azul, na maior parte do gráfico acompanhou de perto a linha vermelha, que são os dados reais de temperatura. É possível verificar que no decorrer do gráfico a temperatura cai de aproximadamente 28 graus Celsius para aproximadamente 26. Dessa forma, verifica-se que a predição acompanhou as variações da temperatura real na predição do inverno de 2016.

4.1.1.2 Aplicação do melhor modelo de Inverno



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 25: Aplicação do melhor modelo de inverno 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão.

O modelo criado com dados de inverno atingiu resultados expressivos no teste em conjuntos de dados com de estações: outono (R^2 0,99), primavera (R^2 0,99) e verão (R^2 0,993). No teste com dados de verão, em que atingiu melhor resultado ao se considerar o Coeficiente de Determinação, o MAE foi 0,28, MSE 0,13 e RMSE foi de 0,36, ou seja, o modelo foi capaz de capturar as oscilações de temperatura com grande taxa de acerto, conforme Figura 25d. O valor de p Wilcoxon foi zero, o que indica que as séries (dados reais e dados previstos) são de diferentes distribuições. Os valores de pvalue Kendall e Spearman p também foram zero, o que indica forte correlação entre dados reais e previstos. A média dos dados previstos foi de 23,75 e dados reais de 23,55.

Em Yang et al. (2017), os autores também consideram informações espaço-temporais para fazer a predição da temperatura na superfície do oceano. Os autores criaram um modelo combinado de LSTM com CNN e compararam os resultados alcançados com outros modelos: SVM (*Support Vector Machines*), SVR

(*Support Vector Regression*) e LSTM. Na comparação usaram RMSE como métrica. O modelo criado pelos autores atingiu melhor resultado com RMSE 0,40, resultado inferior ao encontrado neste estudo na previsão da temperatura do ar para o inverno de 2016.

Karevan e Suykens (2018) criaram modelo LSTM, tipo Stack, para prever a temperatura máxima e mínima. Os autores testaram configurações LSTM com e sem dados espaço-temporais. Usaram todos os dados de entrada disponíveis, testaram configurações com 20, 40, 80, 160, 320 e 640 neurônios na primeira camada. Foi testado também duas opções de funções de ativação: sigmoid e tangente. Para analisar os resultados atingidos usaram MAE e MSE como métrica. pelos resultados obtidos conclui-se que ao considerar as informações espaço-temporais atingi-se resultados melhores considerando as métricas MAE e MSE. Na maioria dos testes os resultados ao utilizar função sigmoid foram melhores que ao utilizar tangente.

Em Gong et al. (2017) criou-se modelo LSTM também com informações espaço-temporais e analisou-se a relação entre a *loss* e o número de épocas utilizado. Concluiu-se que após 100 épocas, valor utilizado no presente estudo, não ocorre ganho na loss. Isso significa que no estudo citado não existe melhora ao elevar o número além de 100 épocas ao se considerar o valor de *loss*.

4.1.1.3 Modelo Outono

A criação do modelo de outono apresentou resultados inferiores aos encontrados no inverno. O maior R^2 encontrado foi de 0,88, já o MSE 0,040 foi e RMSE de 0,020. Esta configuração segue abaixo:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 32.
- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

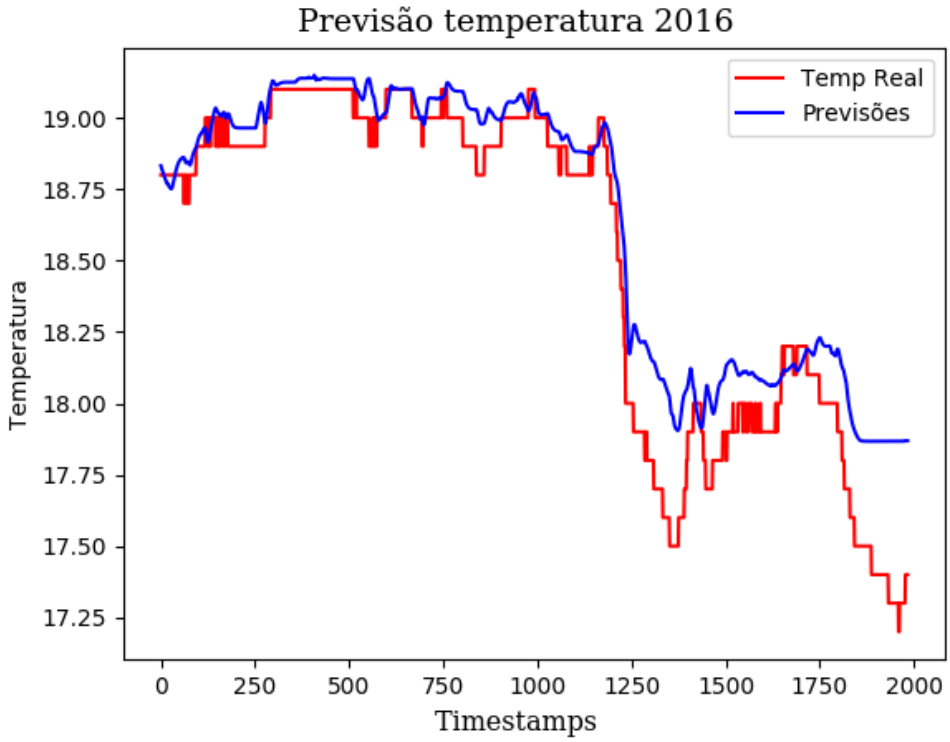
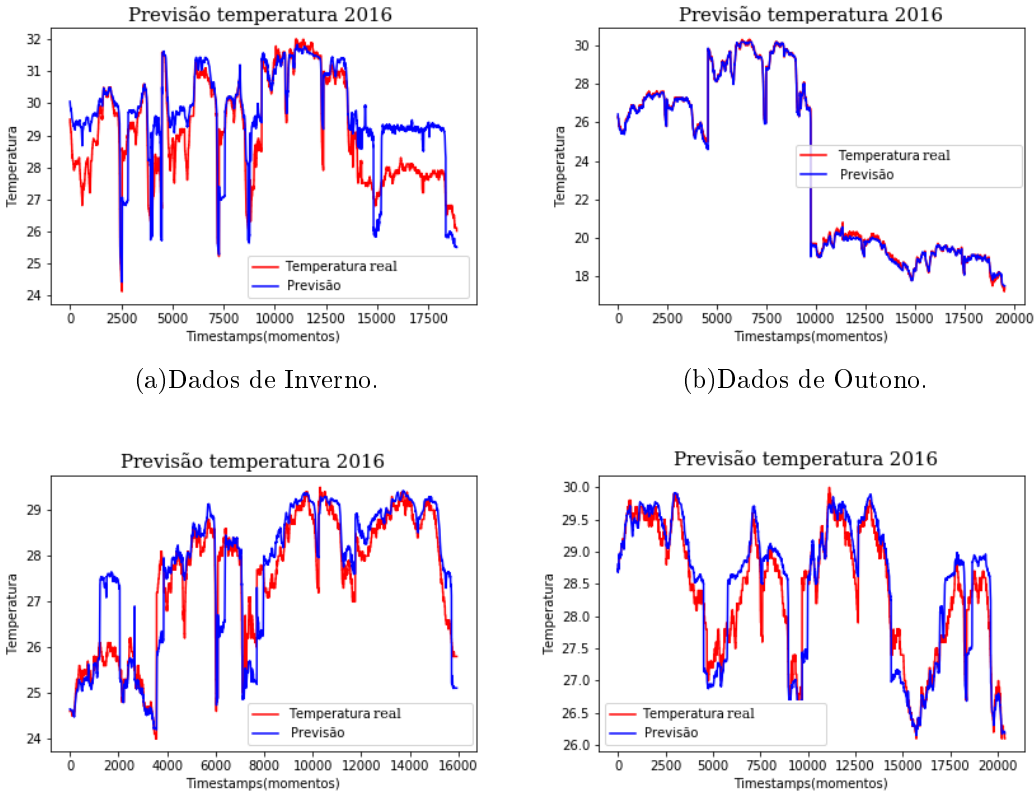


Figura 26: Gráfico com melhor modelo de outono.

Na Figura 32 é possível checar o comportamento da predição em relação com os dados reais. É possível verificar mudança considerável na temperatura no decorrer de poucos *timestamps*, o que pode ter trazido dificuldade no processo de aprendizagem.

4.1.1.4 Aplicação do melhor modelo de Outono



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 27: Aplicação do melhor modelo de outono 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão.

O modelo criado para outono foi o que atingiu R^2 mais baixo. Dois fatores podem contribuir para este resultado, em primeiro lugar, nessa estação ocorre uma transição, existem períodos com e sem chuva. Em segundo lugar, na segunda metade dos *timestamps* coletados a temperatura diminuiu em aproximadamente 10 graus Celsius em comparação com a primeira metade dos dados coletados, o que representa uma mudança brusca. Dessa forma, o maior Coeficiente de Determinação alcançado no teste em outro conjunto de dados foi 0,85 e na estação verão, MAE 0,28, MSE 0,13 e RMSE 0,37, conforme Figura 33d. O valor de p Wilcoxon foi zero, assim como pvalue Kendall e Spearman p.

4.1.1.5 Modelo Primavera

No processo de criação do modelo da primavera o R^2 mais alto atingiu 0,94, MSE 0,003 e RMSE 0,058. O R^2 foi superior ao encontrado no outono, por

exemplo. A configuração que atingiu esse resultado segue abaixo:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 64.
- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

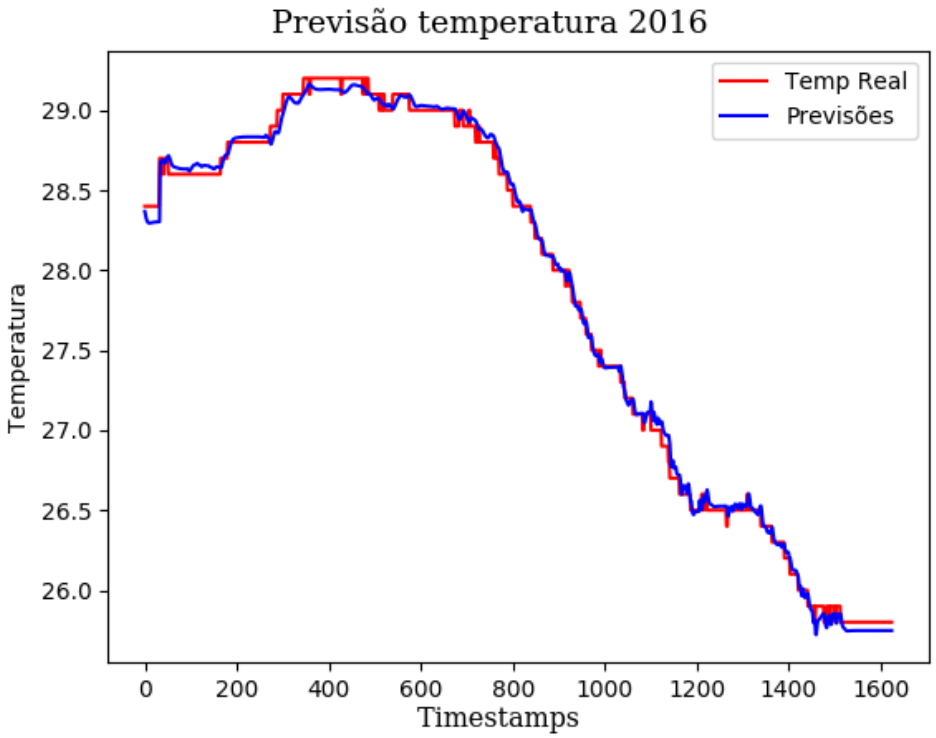
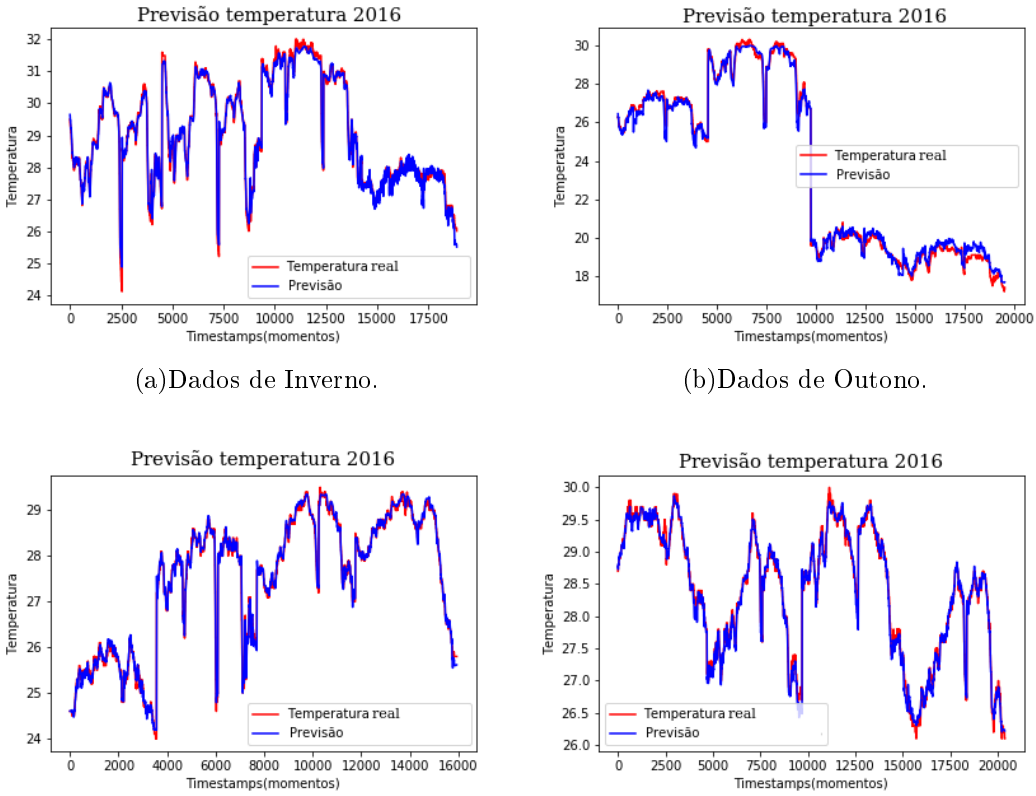


Figura 28: Gráfico com melhor modelo da primavera.

Na Figura 34 observa-se que na maior parte do eixo temporal a variável de predição acompanhou de perto as oscilações da temperatura real. Diferente de outono, neste caso ocorreram movimentações mais lentas no decorrer dos *timestamps*, é possível verificar que depois de um breve aumento da temperatura houve uma mudança de rota e a temperatura passou a cair praticamente de forma contínua, o que pode ter relação com o processo de dissipação de calor acumulado durante o dia.

4.1.1.6 Aplicação do melhor modelo Primavera



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 29: Aplicação do melhor modelo de primavera 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão.

O Modelo da primavera ao ser aplicado em outros conjuntos de dados atingiu os seguintes números R^2 0,98 para inverno, R^2 0,995 para outono e R^2 0,99 para o verão. Ou seja, o melhor resultado ocorreu em outono com MAE 0,23, MSE 0,099, RMSE 0,31, a média dos dados reais foi de 23,55 e dos dados previstos foi de 23,56, Kendall pvalue e Spearman p com valor zero indicando forte correlação. Observa-se a proximidade entre as séries na Figura 29b.

Em Luo et al. (2017) os autores criaram um modelo LSTM para prever a temperatura de eixos em locomotivas. Como métrica utilizaram MSE. Selecionaram diferentes pontos dos eixos para fazer análise. Nos resultados alcançados a diferença entre temperatura real e temperatura prevista ficou entre 1,2 e 1,9 grau Celsius, resultado inferior ao encontrado neste estudo ao prever a temperatura da primavera em 2016.

Wang et al. (2017) os autores propuseram modelo LSTM, tipo Stack, com informações espaço-temporais, combinando com uma RNN. O objetivo foi

compartilhar espaço de memória. Comparou-se os resultados obtidos com outros modelos: CDNA, DFN, VPN baseline. A métrica usada foi o MSE. O modelo criado pelos autores conseguiu melhores resultados em comparação com os outros modelos ao considerar valores de MSE.

4.1.1.7 Modelo Verão

O melhor modelo de verão alcançou R^2 de 0,99, MSE 0,007, RMSE 0,082, e foi criado com a seguinte configuração:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 64.
- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

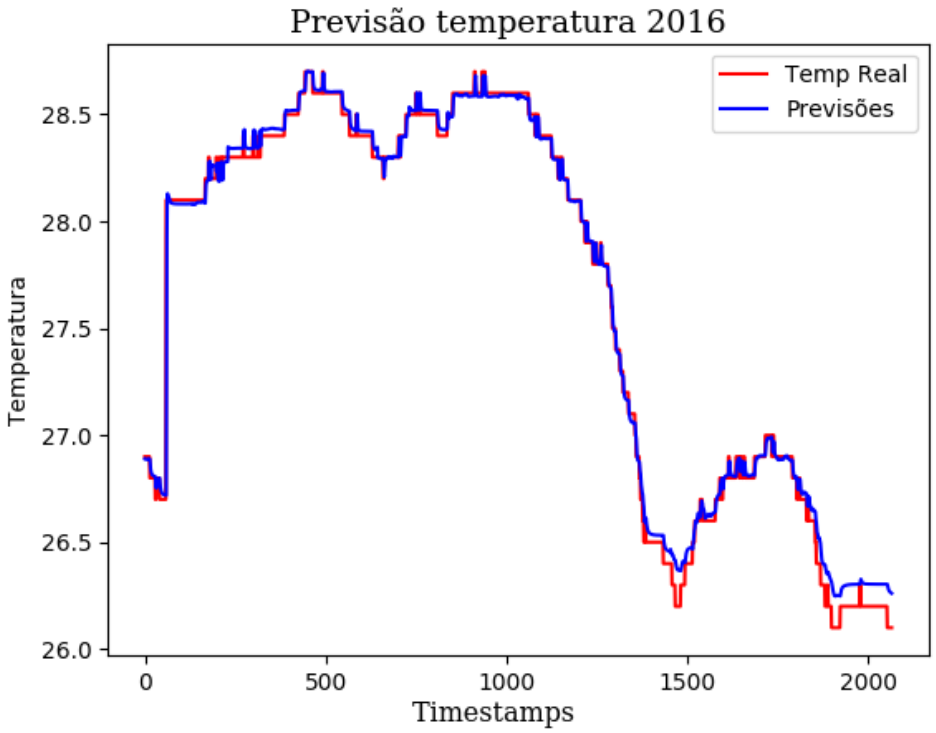
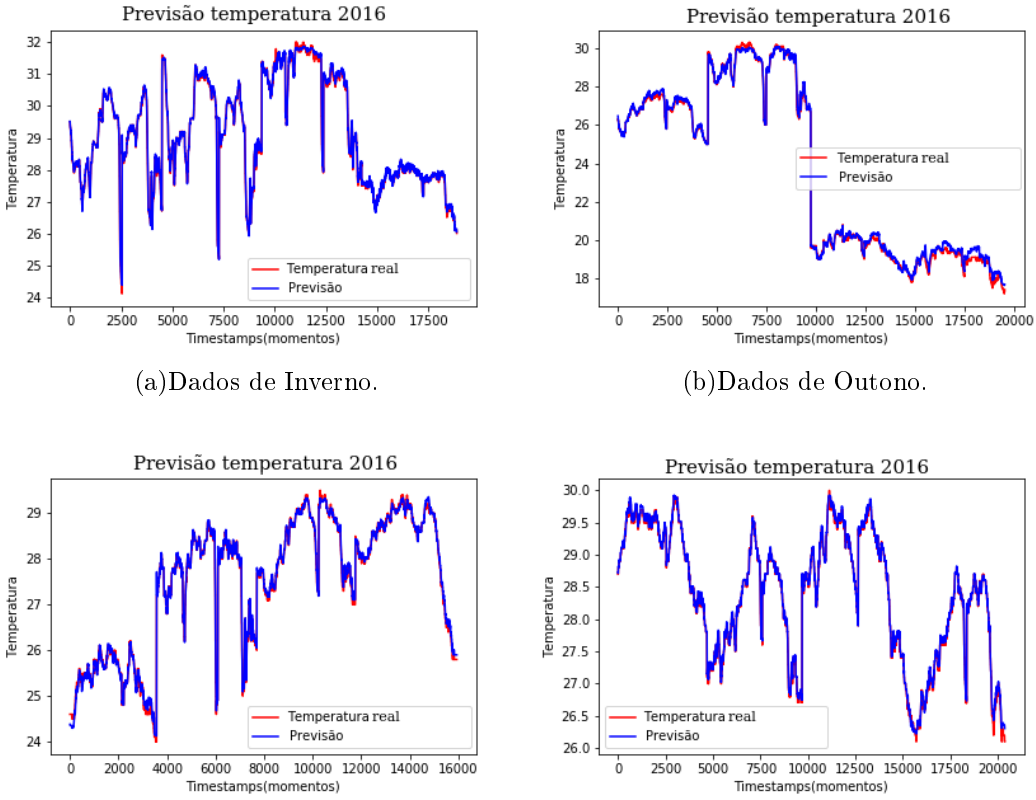


Figura 30: Gráfico com melhor modelo de verão.

Na Figura 36 verifica-se que a variável de predição consegue acompanhar de perto os dados reais na maior parte do eixo temporal. O maior distanciamento ocorre na parte final da série temporal em que ocorrem dois movimentos gráficos acentuados e a Rede Neural não conseguiu capturar essa tendência com a mesma precisão de antes, a variável de predição acaba invertendo a tendência de queda antes do movimento da variável de temperatura real.

4.1.1.8 Aplicação do melhor modelo Verão



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 31: Aplicação do melhor modelo de verão 2011 previsão 2016 aplicado no conjunto inteiro de dados de inverno, outono, primavera, verão.

O modelo criado com dados de verão obteve o melhor resultado ao ser aplicado nos dados de outono, com R^2 0,997, MAE 0,14, MSE 0,044 e RMSE 0,021, ou seja, mesmo com as mudanças abruptas na série de outono o modelo conseguiu se ajustar. A media dos dados reais ficou em Media Real = 23,55 e previstos com 23,65, Wilcoxon p, Kendal pvalue e Spearman p obtiveram valor zero. A proximidade entre as séries pode ser observado na Figura 31b. No caso do outono o R^2 foi de 0,99, mesmo valor de primavera e verão.

Em Liu et al. (2018) os autores criaram um modelo LSTM considerando a dependência temporal na previsão da temperatura marítima em diferentes profundidades. A métrica utilizada para avaliar o método foi o MSE. Os autores compararam os resultados atingidos com outros métodos: SVR, MLPR (*Multilayer Perceptron Regressor*). O modelo dos autores alcançou MSE 0,030, resultado próximo ao MSE encontrado neste estudo na previsão do verão de 2016.

Já em Salman et al. (2018) os autores criaram modelos de sequência temporal LSTM com 1 e com 4 camadas para prever o valor das variáveis temperatura, umidade e pressão. Como métrica utilizaram o RMSE. Em todos os testes resultados que usaram modelo LSTM com quatro camadas obtiveram resultados superiores aos modelos com apenas uma camada. Na previsão da temperatura obteve RMSE de 0,08 utilizando utilizando 100 épocas, resultado que é inferior ao alcançado no presente estudo na previsão do verão em 2016.

4.1.1.9 Tabela do Coeficiente de Determinação

O conjunto de R^2 dos modelos de inverno, outono, primavera e verão criados considerando a primeira amostragem foram aplicados nas diferentes estações, conforme Tabela 7.

Tabela 7: Valores de R^2 com dados de 2011 para prever 2016 com 90% dos dados para treinamento.

Valores de R^2				
-	Modelos			
Dados	Inverno	Outono	Primavera	Verão
Inverno	0,997	0,613	0,986	0,994
Outono	0,993	0,999	0,995	0,997
Primavera	0,964	0,793	0,994	0,995
Verão	0,974	0,857	0,990	0,996

Na Tabela 7 verifica-se um resumo dos valores do Coeficiente de Determinação alcançados nos testes.

4.2 2016 prever 2016

Nessa sessão apresenta-se resultados de modelos criados em cada estação de 2016 e testados nas demais estações também do ano de 2016.

4.2.1 Amostragem 90% de treino e 10% de teste

As mesmas amostragens testadas com dados de 2011 para prever 2016 foram usadas nessa sessão, qual seja 80% e 90% dos dados para treinamento.

4.2.1.1 Modelo Inverno

Na criação do modelo de inverno 2016 com 90% dos primeiros dados da série para treinamento e previsão dos últimos 10% a melhor combinação de parâmetros atingiu R^2 0,98, MSE 0,004 e RMSE 0,068. A configuração segue abaixo:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 64.
- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

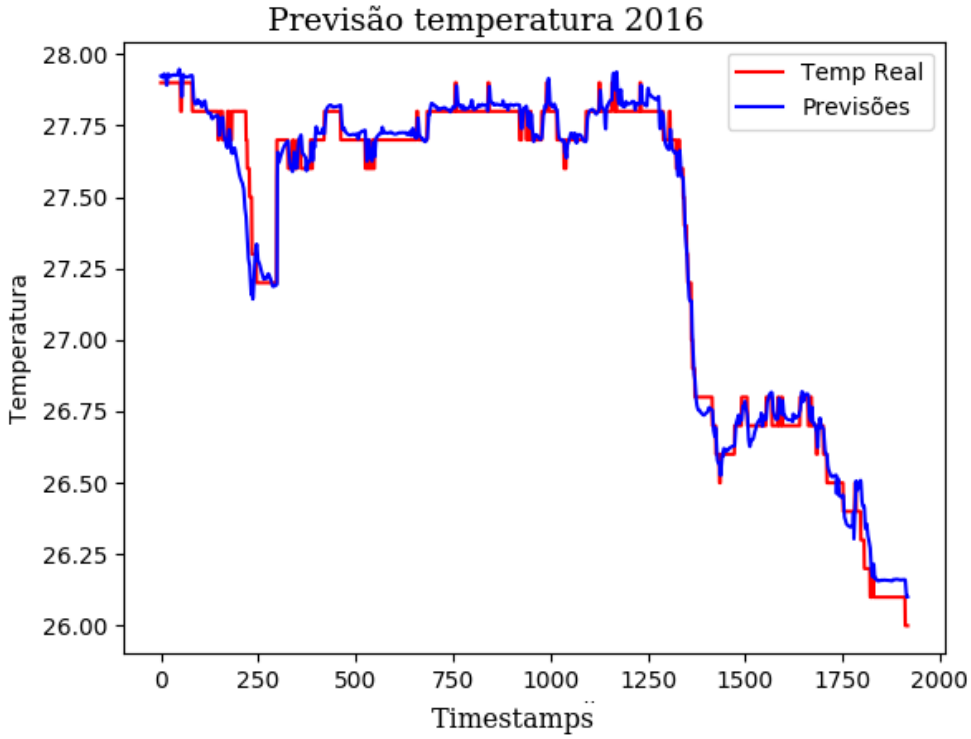
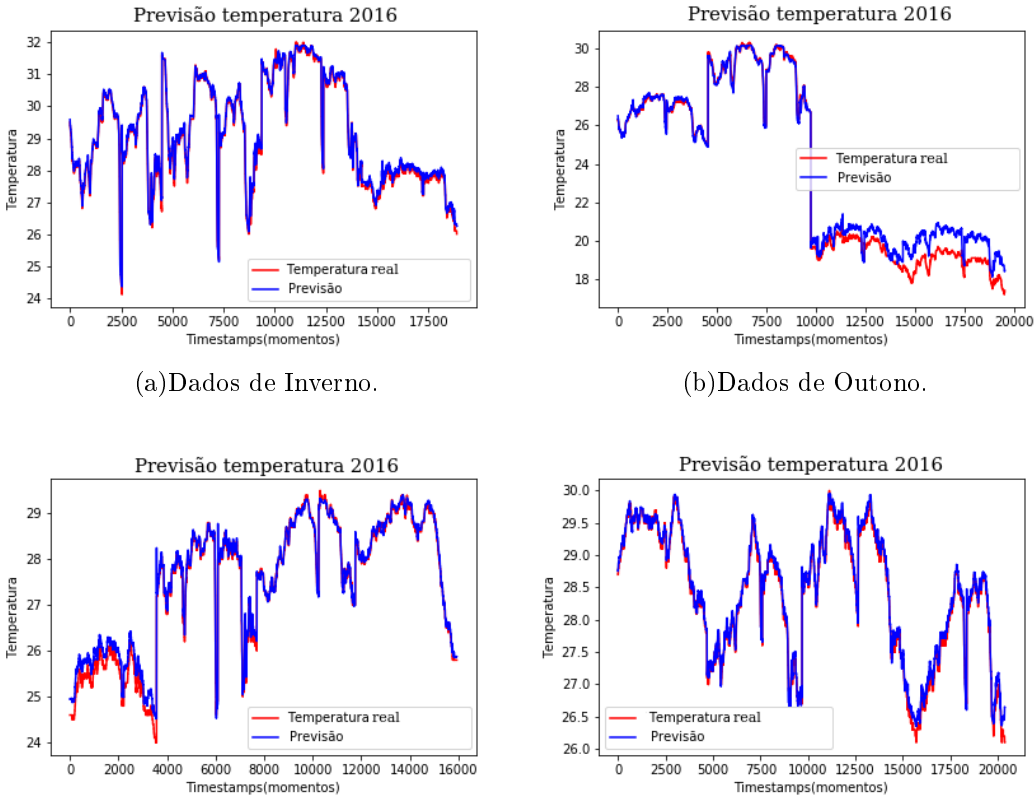


Figura 32: Modelo inverno 2016 aplicado em dados de inverno 2016.

Na Figura 32 visualiza-se que o a variável de predição acompanhou as movimentações da variável com dados reais, seja no eixo horizontal ou vertical, isto é, a predição alcançou valores próximos aos reais da temperatura.



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 33: Aplicação do melhor modelo de inverno 2016 na previsão das outras estações do mesmo ano.

Na aplicação do melhor modelo de inverno 2016 aplicado nas outras estações destaca-se os resultados atingidos nas outras três estações (Figuras 33b, 33c e 33d), visto que no outono obteve R^2 0,97, na primavera R^2 0,98,e, por fim o verão, que obteve o melhor resultado: R^2 0,99, com MAE 0,076, MSE 0,008, RMSE 0,094, a media dos dados reais foi 28,38 e valores previstos de 28,44. Wilcoxon p, Kendall pvalue e Spearman p com valor zero.

4.2.1.2 Modelo Outono

Na criação do melhor modelo de Outono a variável de predição encontrou dificuldades na segunda metade da série em acompanhar movimentações acentuadas nos dados reais, o modelo obteve o R^2 0,93, MSE 0,0228 e RMSE 0,1513. A configuração segue abaixo:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 32.
- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

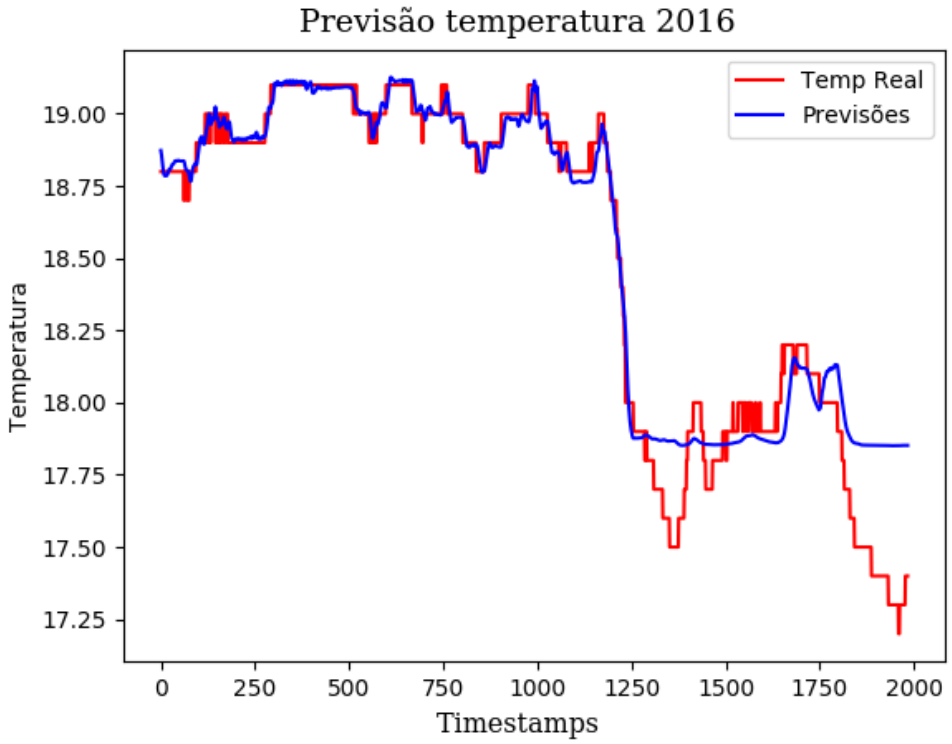
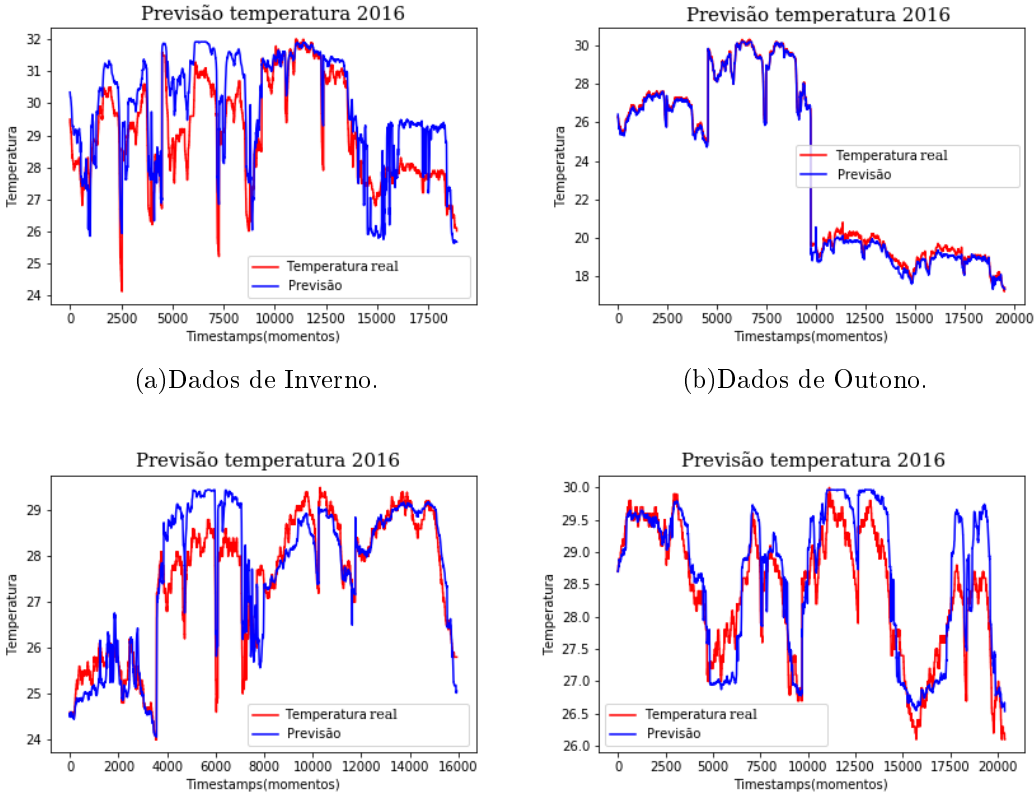


Figura 34: Modelo outono 2016.

Na Figura 34 é possível verificar que em vários *timestamps* ocorre diferença próxima a 1 grau Celsius, o modelo teve dificuldade em acompanhar a tendência e velocidade das alterações na temperatura real principalmente na parte final da série de dados.



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 35: Aplicação do melhor modelo de outono 2016 no conjunto inteiro de dados de inverno, outono, primavera, verão.

Na aplicação do modelo de outono em outras estações de 2016 é possível destacar o resultado encontrado nos dados de primavera com R^2 0,79, MAE 0,46, MSE 0,41 e RMSE 0,64, a media dos dados reais foi 27,48 e previstos 27,61.

4.2.1.3 Modelo Primavera

Na criação do modelo de primavera 2016 para prever as outras estações de 2016 o melhor ajuste de parâmetros obteve R^2 0,99, MSE 0,0060 e RMSE 0,0776. Segue a configuração do modelo abaixo:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 32.
- Tipo de LSTM: Stack.
- Número de épocas: 100.

- Quantidade de valores anteriores em uma previsão: 300.

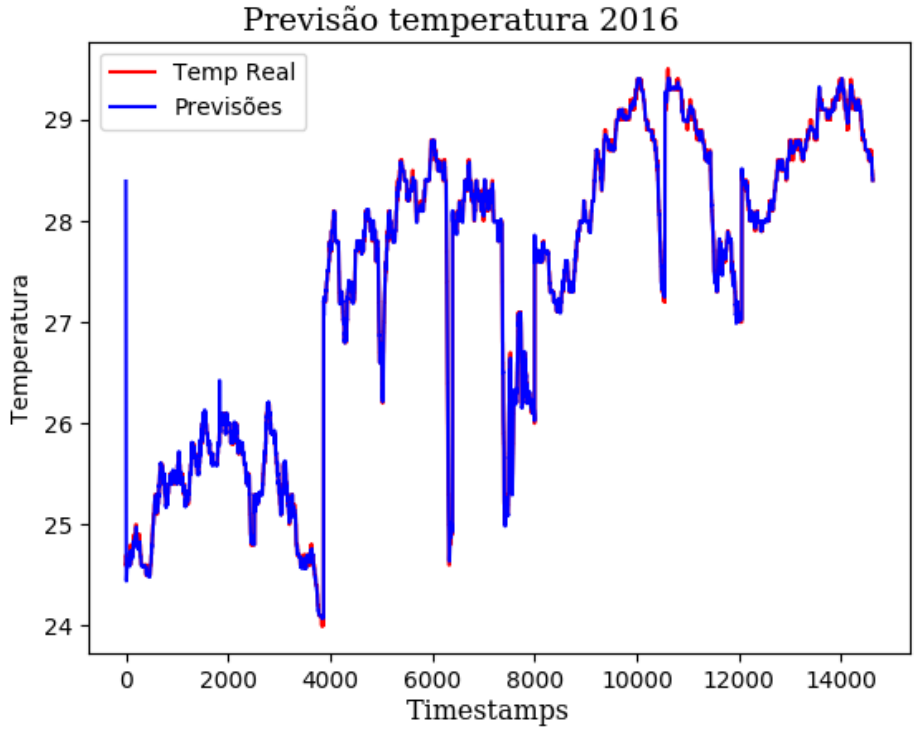
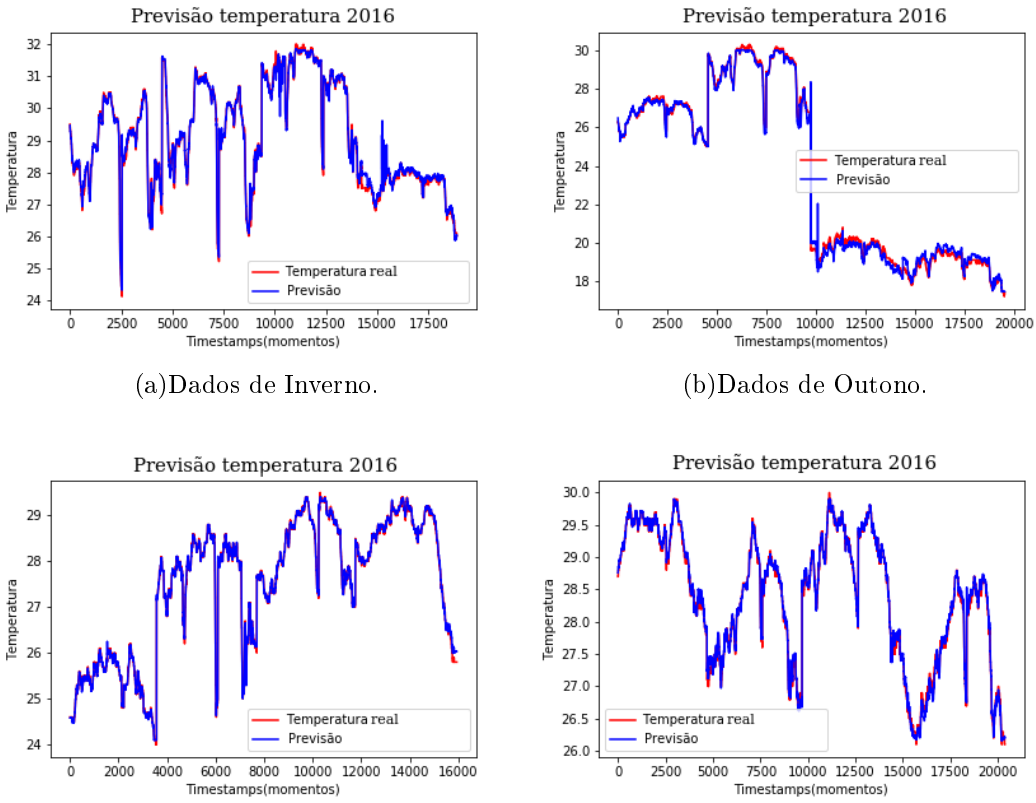


Figura 36: Modelo primavera 2016...

Na Figura 36 visualiza-se que ocorre movimento inicial de subida na temperatura e a seguir acontece um movimento de descida que vai até o final da série, e a variável de predição foi capaz de se manter próxima aos dados reais, o que justifica o valor de R^2 alcançado.



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 37: Aplicação do melhor modelo da primavera de 2016 nos dados das outras estações do mesmo ano.

Na aplicação do modelo de primavera de 2016 no conjunto de dados das outras estações destacou-se o resultado atingido com dados do outono, qual seja R^2 0,996, MAE 0,17, MSE 0,061, RMSE 0,24, os dados reais tiveram média de 23,55 e dados previstos 23,53, Figura 37b. Nos dados de inverno obteve R^2 0,99 assim como outono e verão.

4.2.1.4 Modelo Verão

Na criação do modelo de verão para testar os últimos 10% de dados o melhor modelo alcançou R^2 0,99, MSE 0,0026 e RMSE 0,0515. Segue a configuração abaixo:

- Número de camadas ocultas: 4.
- Quantidade de Neurônios: 100.
- Valor de BatchSize: 64.

- Tipo de LSTM: Stack.
- Número de épocas: 100.
- Quantidade de valores anteriores em uma previsão: 300.

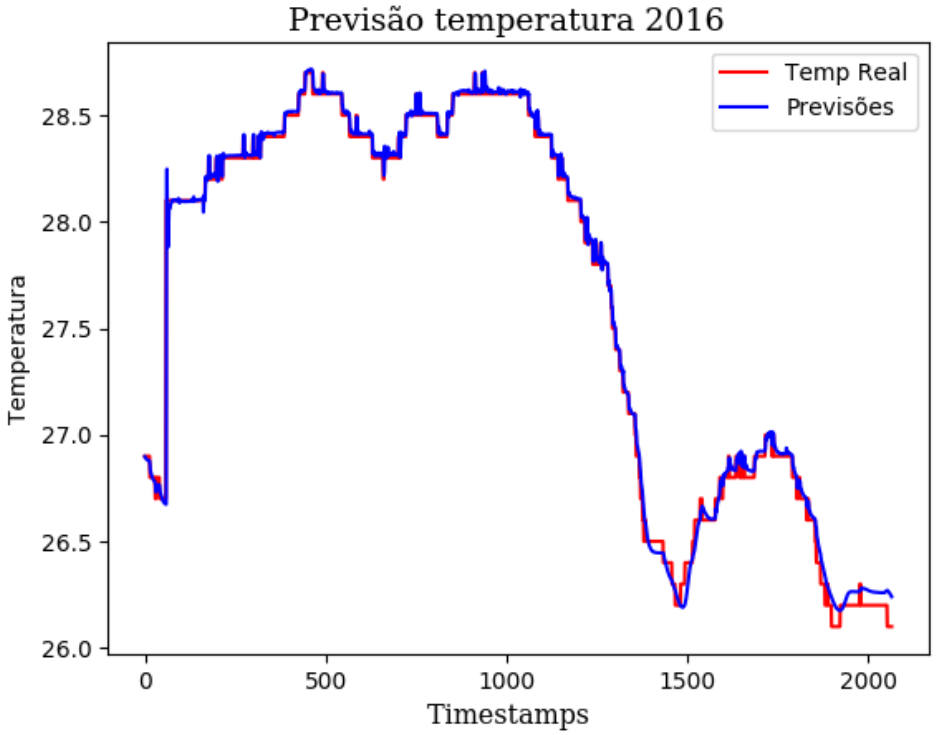
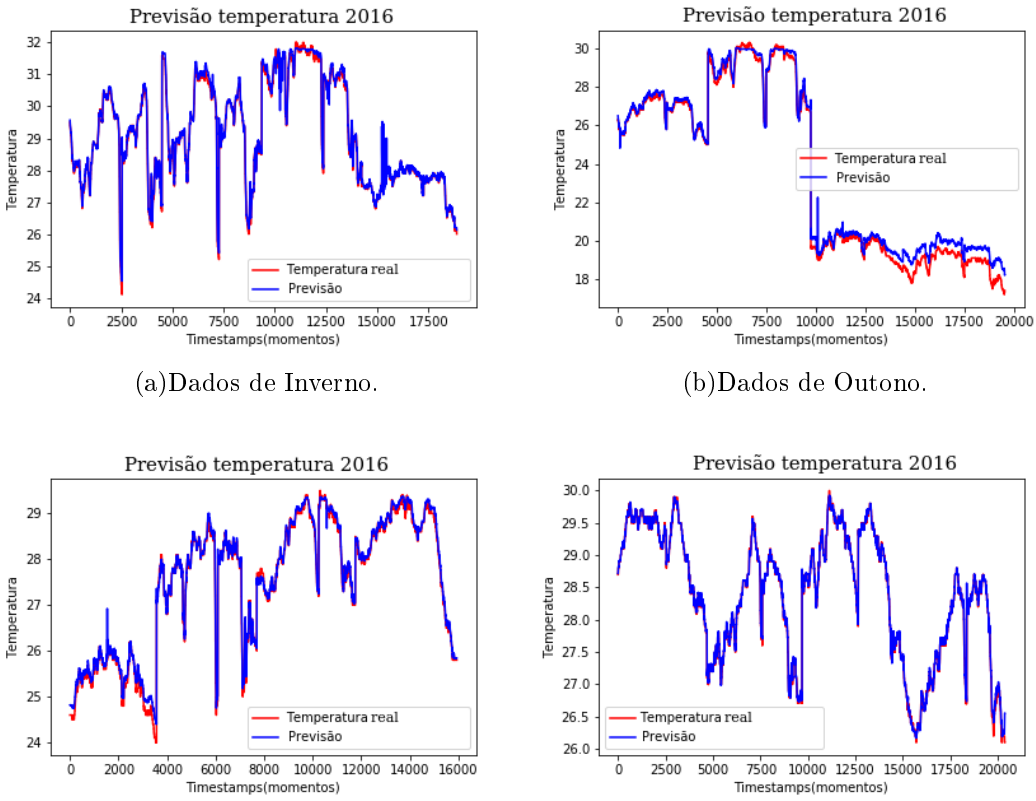


Figura 38: Modelo verão 2016.

Na Figura 38 observa-se que a variável de predição foi capaz de seguir as variações gráficas no comportamento da variável de temperatura com dados reais, o que justifica o valor de R^2 encontrado.



(a)Dados de Inverno.

(b)Dados de Outono.

(c)Dados de Primavera.

(d)Dados de Verão.

Figura 39: Aplicação do melhor modelo de verão 2016 no conjunto inteiro de dados das outras estações de 2016

Na aplicação do modelo de verão 2016 em outros conjuntos de dados também de 2016 destaca-se o resultado alcançado em dados na primavera, em que se alcançou R^2 0,993, MAE 0,076, MSE 0,01, RMSE0,11, Figura 39c. O valor de Wilcoxon p, Kendall pvalue e Spearman p foi zero. A media da série com dados reais foi de 27,48 e dados previstos 27,54. Nas estações de inverno e outono atingiu R^2 de 0,99.

4.2.1.5 Tabela do Coeficiente de Determinação

O conjunto de R^2 dos modelos de inverno, outono, primavera e verão criados considerando a primeira amostragem foram aplicados nas diferentes estações, conforme Tabela 9.

Tabela 8: Valores de R^2 com dados de 2016 para prever outras estações de 2016 com 90% dos dados para treinamento.

Valores de R^2				
-	Modelos			
Dados	Inverno	Outono	Primavera	Verão
Inverno	0,993	0,550	0,990	0,990
Outono	0,978	0,990	0,996	0,990
Primavera	0,985	0,798	0,998	0,993
Verão	0,990	0,742	0,995	0,998

Na Tabela 9 verifica-se um resumo dos valores do Coeficiente de Determinação alcançados nos testes.

4.3 Previsão 2021

Para simular a temperatura de utilizou-se o modelo gerado com dados de 2011 para prever 2016, e dessa vez, esse modelo teve como entrada os dados de 2016 para prever 2021.

4.3.1 Inverno 2021

Na simulação da temperatura do ar no inverno de 2021, Figura 40, é possível observar que a média ficou em torno de 25 graus Celsius, com pico inferior próximo a 24ºC, e pico superior próximo a 32ºC.

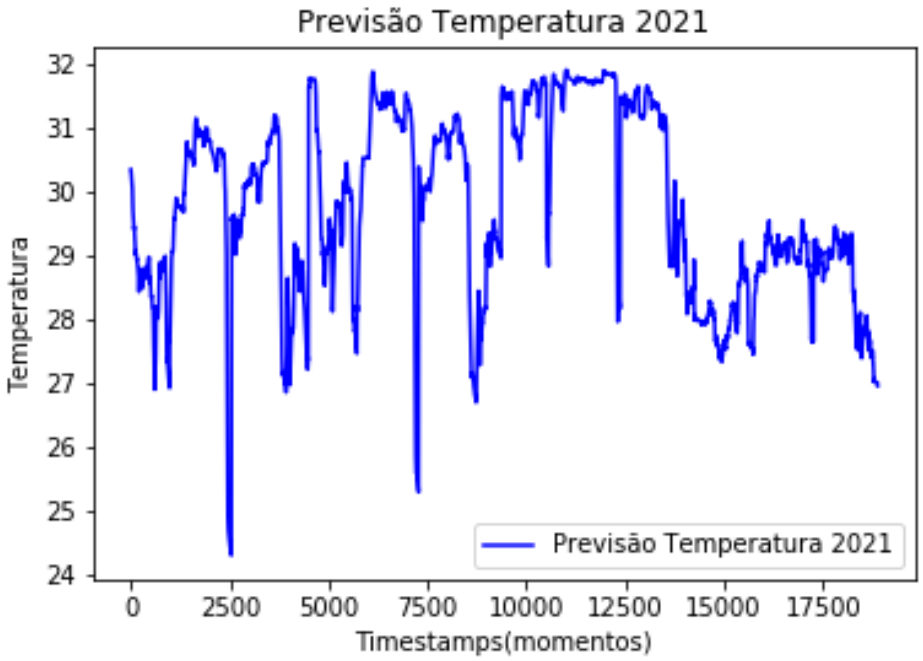


Figura 40: Previsão de temperatura para o ano de 2021.

4.3.2 Outono 2021

Na simulação da temperatura do ar em outono de 2021 é possível observar que a média ficar em torno de 24 graus Celsius, com pico inferior próximo a 18^{o}C, e pico superior próximo a 30^{o}C, conforme Figura 41.

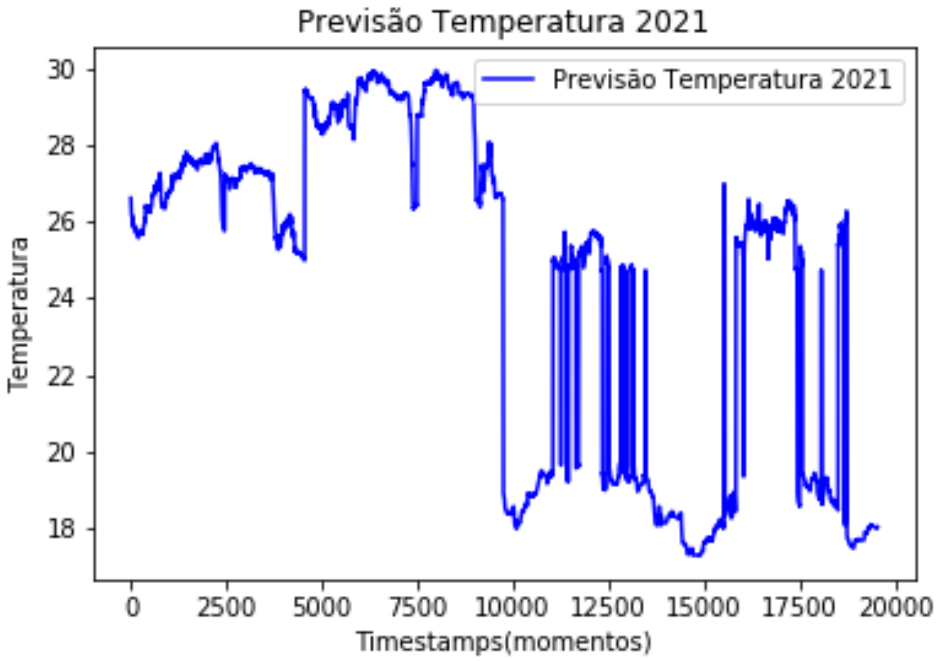


Figura 41: Previsão de temperatura para outono de 2021.

4.3.3 Primavera 2021

Na simulação da temperatura do ar na primavera de 2021, Figura 42, é possível observar que a média ficar em torno de 27 graus Celsius, com pico inferior próximo a 24^{o}C, e pico superior próximo a 29^{o}C.

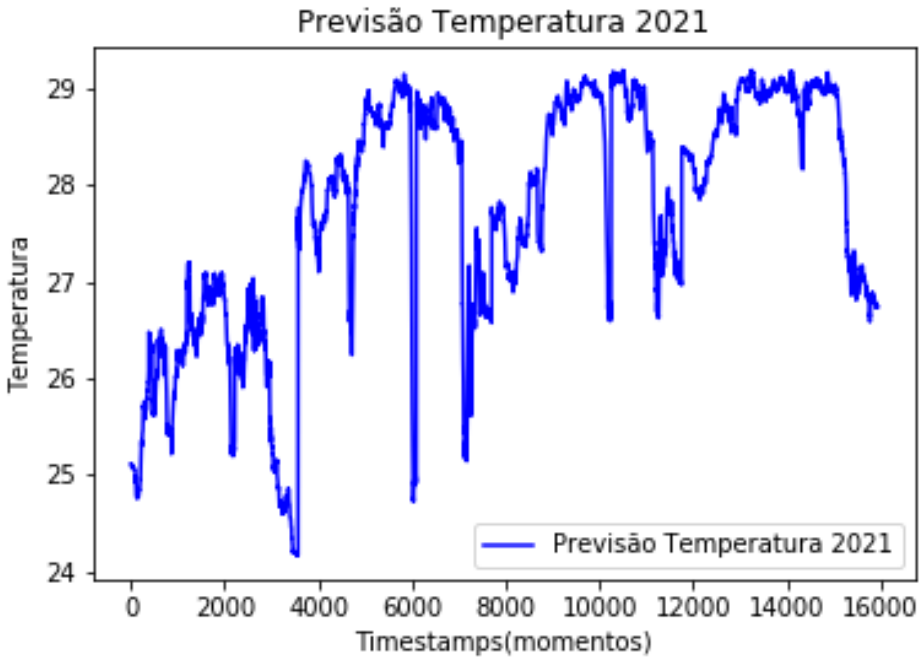


Figura 42: Previsão de temperatura para primavera de 2021.

4.3.4 Verão 2021

Na Figura 43 visualiza-se a simulação da temperatura do ar na primavera de 2021, em que é possível observar que a média ficar em torno de 27 graus Celsius, com pico inferior próximo a 24^{o}C, e pico superior próximo a 30^{o}C.

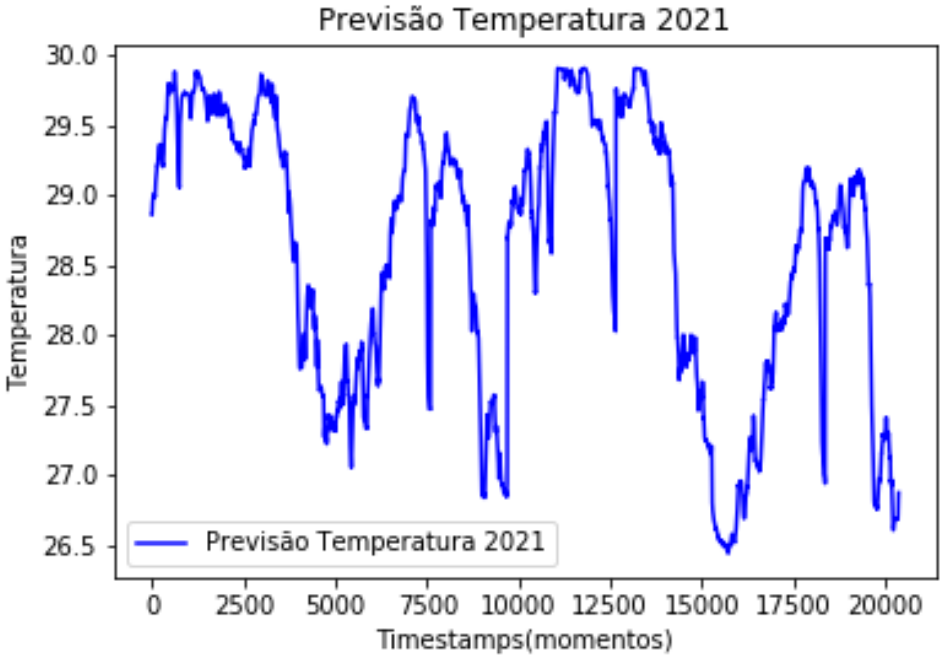


Figura 43: Previsão de temperatura para verão de 2021.

4.4 Dados de 2011, 2016 e previsão 2021

No comparativo entre dados de 2011, 2016 e a simulação de 2021 foi necessário usar médias para ajustar a dimensão dos gráficos, visto que existiam quantidades diferentes de *timestamps*. Dessa forma, foi possível projetar 2011, 2016 e 2021 no mesmo gráfico.

4.4.1 Comparativo Inverno 2011, 2016, 2021

Na Figura 44 é possível verificar o comportamento das três (3) séries: 2011, 2016 e a simulação de 2021, considerando a estação inverno. Nessa figura visualiza-se que a temperatura do ar em 2011 destoa das demais séries.

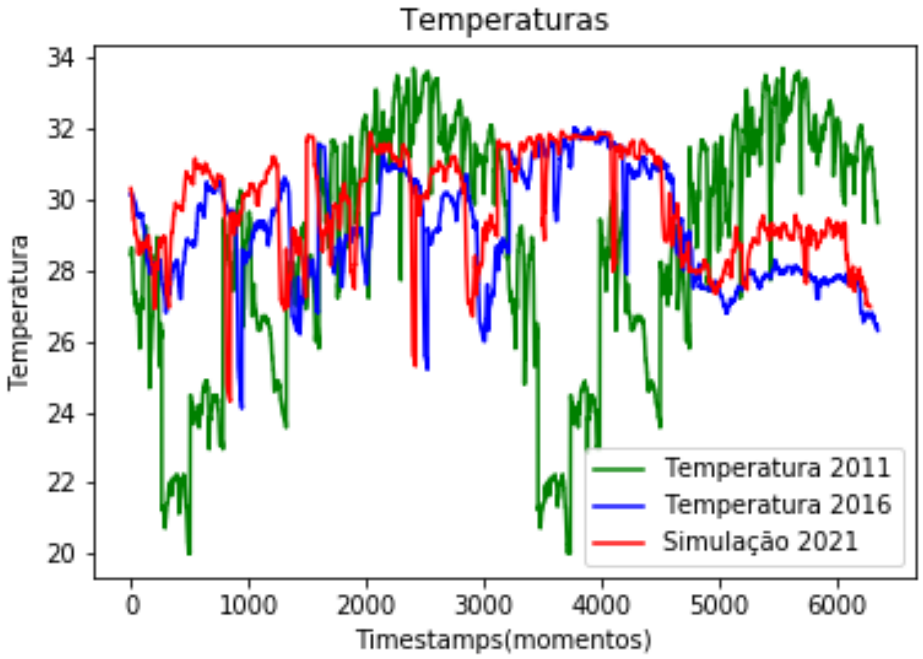


Figura 44: Comparativo da temperatura de inverno.

4.4.2 Comparativo Outono 2011, 2016, 2021

Na Figura 45 é possível verificar o comportamento das três (3) séries: 2011, 2016 e a simulação de 2021, considerando a estação outono. Nessa figura visualiza-se que a temperatura do ar em 2011 chega a ser oito graus Celsius acima da simulação de 2021.

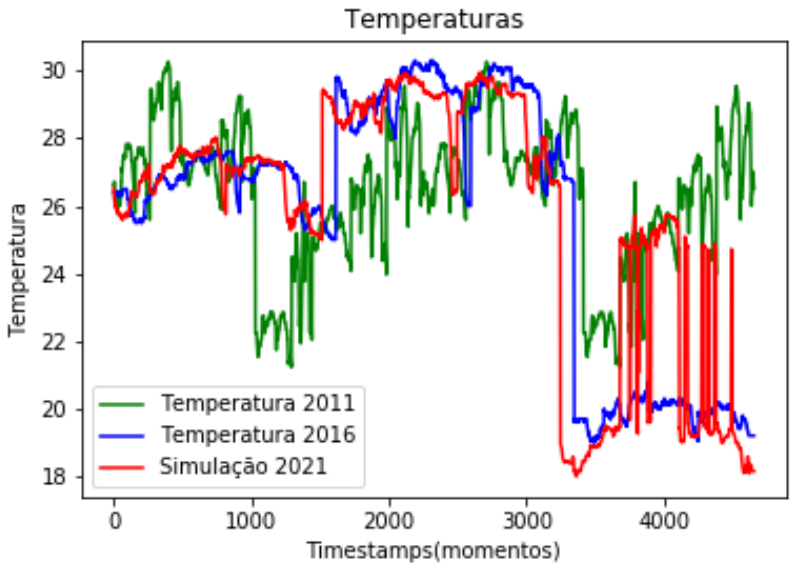


Figura 45: Comparativo da temperatura de outono.

4.4.3 Comparativo Primavera 2011, 2016, 2021

Na Figura 46 é possível verificar o comportamento das três (3) séries: 2011, 2016 e a simulação de 2021, considerando a estação da primavera. Nessa figura visualiza-se que a série de 2011 descola da série 2016 e 2021, já as séries 2016 e 2021 se mantém próximas.

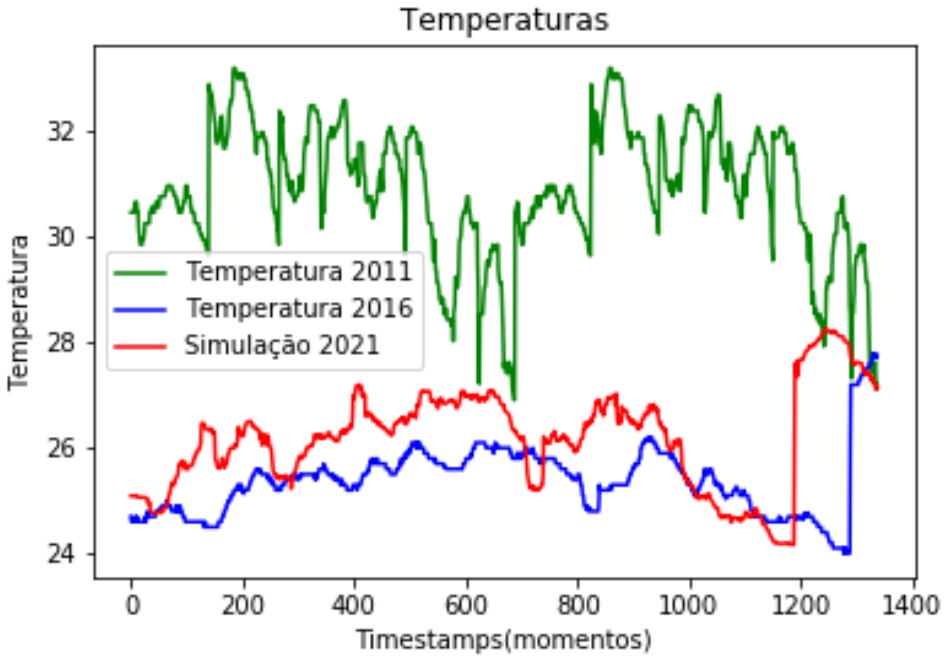


Figura 46: Comparativo da temperatura da primavera.

4.4.4 Comparativo Verão 2011, 2016, 2021

Na Figura 47 é possível verificar o comportamento das três séries considerando a estação do verão. Nessa figura visualiza-se que a série de 2011 possui na maioria dos *timestamps* temperaturas menores que as demais, as séries 2016 e 2021 se mantém próximas.

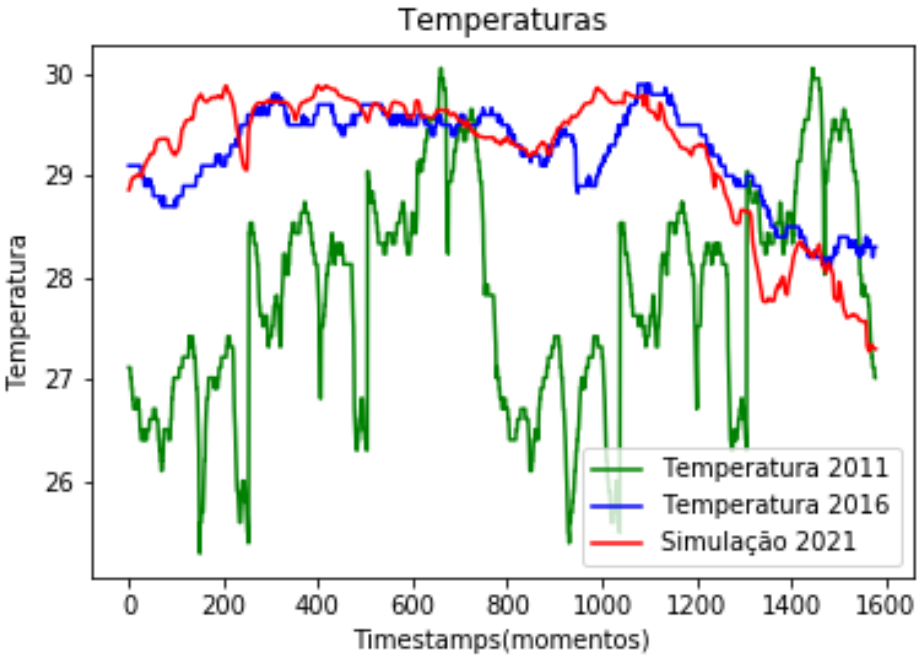


Figura 47: Comparativo da temperatura do verão.

Capítulo 5

Considerações Finais

O presente trabalho propõe uma abordagem que usa Redes Neurais Recorrentes do tipo *Long-Short Term Memory* para fazer a predição da temperatura do ar utilizando dados microclimáticos coletados usando a metodologia de transecto móvel noturno.

Essa abordagem utilizou conjuntos de dados ambientais e de cobertura do solo do ano de 2011 para prever 2016 e criou e validou modelos usando técnica computacional para teste dos dados e para simular a temperatura do ar em 2021.

Os dados coletados com variáveis ambientais e de cobertura do solo foram preparados com êxito e serviram como entrada para Rede Neural Artificial.

Por meio de método estatístico foi possível comparar dados reais com dados previstos, por exemplo, na previsão da temperatura do ar em 2016 foi possível encontrar valores de R^2 superiores a 0,9.

Durante esta pesquisa diversas dificuldades foram encontradas como falta de informações básicas nas publicações consultadas a respeito de parâmetros LSTM, e dificuldade de conseguir uma infraestrutura computacional que não demorasse meses para processar o volume de dados utilizado.

Esta pesquisa esteve relacionada a estudos em desenvolvimento no PPGFA. Os resultados extraídos da Rede Neural LSTM servirão de base para esses estudos verificarem ilhas de calor na área urbana de Cuiabá-MT, além de fazer simulações da temperatura do ar para anos posteriores considerando a questão espacial e temporal. Em 2019 está sendo realizada uma nova coleta de dados, após o final a rede vai ser treinada novamente considerando essas novas entradas.

Por meio dos resultados atingidos é possível constatar que o uso de LSTM utilizando como entrada dados ambientais e de cobertura do solo, organizados de forma espaço-temporal, se mostrou viável na predição da temperatura do ar.

5.1 Contribuições

As principais contribuições deste estudo são:

- Criar novo cálculo de distância que considera a questão espacial e temporal na previsão da temperatura do ar. Por meio deste cálculo foi possível selecionar os três pontos mais próximos de um ponto escolhido. As informações dos três pontos mais próximos foram usadas como entrada na Rede Neural LSTM para fazer predição da temperatura do ar.

- Listar e detalhar os elementos técnicos necessários para realizar predição utilizando Rede Neural Artificial LSTM.

- Compartilhar informações sobre predição da temperatura do ar contribuindo, por exemplo, com estudos em andamento do PPGFA a respeito das ilhas de calor em Cuiabá-MT.

- Disponibilizar na Internet todos os códigos criados neste estudo seguindo a linha de *open science*. Por meio dos códigos disponibilizados é possível variar parâmetros, implementar melhorias e testar outros otimizadores, outras funções.

- Desenvolver modelos preditivos precisos e confiáveis, por meio da influência dos parâmetros climáticos e físicos, fornecendo assim uma informação quantitativa sobre a dependência da temperatura do ar e a cobertura do solo, com os dados nos anos de 2011-2012, 2016, coletados dentro do perímetro urbano da cidade de Cuiabá-MT.

  Por meio dessas contribuições espera-se que os pesquisadores tenham mais elementos para aumentar a qualidade dos seus estudos, em especial, ao trabalhar com predição de dados ambientais usando Redes Neurais Artificiais LSTM.

5.2 Trabalhos Futuros

Como sugestão para trabalhos futuros sugere-se:

- Testar com combinação de diferentes Redes Neurais Artificiais. Por exemplo, é possível combinar LSTM com CNN e formar uma Rede Neural Artificial híbrida, capaz de utilizar as diferentes características de cada arquitetura.

- Implementar periodograma para definir exatamente o valor dos ciclos de informações. Por meio de técnicas como auto-sarima é possível identificar precisamente quantos valores anteriores são necessários para realizar uma previsão.

- Utilizar técnicas de redução de dimensionalidade como a PCA (*Principal Components Analysis*), em português, Análise de Componentes Principais, é uma ferramenta de análise multivariada de dados. Por meio da PCA é possível verificar as relações entre as variáveis medidas e até mesmo entre amostras. Por meio da Análise de Componentes Principais transforma-se o conjunto de variáveis originais em novos conjuntos de variáveis denominadas componentes principais, que possuem objetivo de concentrar o máximo de informação possível ao mesmo tempo que diminui a massa de dados.

REFERÊNCIAS

ALCOFORADO, M.-J.; ANDRADE, H. Nocturnal urban heat island in lisbon (portugal): main features and modelling attempts. *Theoretical and applied climatology*, Springer, v. 84, n. 1-3, p. 151–159, 2006. Citado na página 8.

ALVES, E. D. L.; VECCHIA, F. A. S. Influência de diferentes superfícies na temperatura e no fluxo de energia: um ensaio experimental influence of different surfaces on the temperature and energy flux: an experimental trial. *Ambiência*, v. 8, n. 1, p. 101–111, 2012. Citado na página 9.

AMANCIO, L. B.; UFPI, T.; NETO, S. A. D.; ANJOS, G. J. M. dos; UFPA, B. *Modelo de Previsão de Recalques em Estacas Hélice Contínua Utilizando Redes Neurais Artificiais*. [S.l.]: COBRAMSEG, 2016. Citado na página 13.

AMORIM, M. C. d. C. T. *O clima urbano de Presidente Prudente/SP*. Tese (Doutorado), 2000. Citado na página 5.

ARMENTEROS, J. A.; SALVATORE, M.; EMANUELSSON, O.; WINTHER, O.; HEIJNE, G. V.; ELOFSSON, A.; NIELSEN, H. Detecting novel sequence signals in targeting peptides using deep learning. *BioRxiv*, Cold Spring Harbor Laboratory, p. 639203, 2019. Citado na página 22.

ASSIS, E. S. d. Impactos da forma urbana na mudança climática: método para previsão do comportamento térmico e melhoria de desempenho do ambiente urbano. *São Paulo*, 2000. Citado na página 9.

BANERJEE, I.; LING, Y.; CHEN, M. C.; HASAN, S. A.; LANGLOTZ, C. P.; MORADZADEH, N.; CHAPMAN, B.; AMRHEIN, T.; MONG, D.; RUBIN, D. L. et al. Comparative effectiveness of convolutional neural network (cnn) and recurrent neural network (rnn) architectures for radiology text report classification. *Artificial intelligence in medicine*, Elsevier, v. 97, p. 79–88, 2019. Citado na página 22.

BARBETTA, P. A.; REIS, M. M.; BORNIA, A. C. *Estatística para cursos de engenharia e informática*. [S.l.]: Atlas São Paulo, 2010. v. 3. Citado 2 vezes nas páginas 30 e 31.

BARBIRATO, G. M.; SOUZA, L. C. L. de; TORRES, S. C. *Clima e cidade: a abordagem climática como subsídio para estudos urbanos*. [S.l.]: UFAL, 2007. Citado na página 7.

BAUER, L. Estimação do coeficiente de correlação de spearman ponderado. 2007. Citado na página 32.

BECHMANN, A.; BOWKER, G. C. Unsupervised by any other name: Hidden layers of knowledge production in artificial intelligence on social media. *Big Data & Society*, SAGE Publications Sage UK: London, England, v. 6, n. 1, p. 2053951718819569, 2019. Citado na página 10.

BINOTI, D. H. B.; BINOTI, M. L. M. da S.; LEITE, H. G. Configuração de redes neurais artificiais para estimação do volume de árvores. *Revista Ciência da Madeira (Brazilian Journal of Wood Science)*, v. 5, n. 1, p. 10–12953, 2014. Citado na página 16.

BRABAZON, A.; O'NEILL, M.; MCGARRAGHY, S. *Natural computing algorithms*. [S.l.]: Springer, 2015. Citado na página 13.

BRANDÃO, A. O clima urbano da cidade do rio de janeiro. *Clima urbano*, Contexto Editora São Paulo^ eSP SP, v. 2, p. 121–153, 2003. Citado na página 8.

BROWN, G.; DEKAY, M. *Sol, vento e luz: estratégias para o projeto de arquitetura*. [S.l.]: Bookman Editora, 2009. Citado na página 7.

BROWNLEE, J. *Deep Learning for Time Series Forecasting: Predict the Future with MPLS, CNN and LSTMs in Python*. [S.l.]: Jason Brownlee, 2018. Citado 2 vezes nas páginas 19 e 24.

BROWNLEE, J. *Long Short-Term Memory Networks with Python: Develop Sequence Prediction Models with Deep Learning*. [S.l.]: Jason Brownlee, 2018. Citado 7 vezes nas páginas 10, 14, 24, 25, 27, 28 e 32.

BUFERA, F. *Architettura e ambiente. Manuale per il controllo della qualità termica, luminosa e acustica degli edifici*. [S.l.]: Etas libri, Milano, 1995. Citado na página 7.

BULLINARIA, J. A. Introduction to neural networks. *University of Birmingham, UK*, 2004. Citado na página 20.

BUSSAB, W. d. O.; MORETTIN, P. A. Estatística básica. 2009. Citado na página 31.

CAMPOS, B. P. F.; SILVA, G. F. da; BINOTI, D. H. B.; MENDONÇA, A. R. de; LEITE, H. G. Descrição do perfil do tronco de árvores em plantios de diferentes espécies por meio de redes neurais artificiais. *Pesquisa Florestal Brasileira*, v. 37, n. 90, p. 99–107, 2017. Citado na página 20.

CARDOSO, C. A. V.; CRUZ, G. L. Forecasting natural gas consumption using arima models and artificial neural networks. *IEEE Latin America Transactions*, IEEE, v. 14, n. 5, p. 2233–2238, 2016. Citado na página 32.

CARVALHO, V. *PostgreSQL: Banco de dados para aplicações web modernas*. [S.l.]: Editora Casa do Código, 2017. Citado na página 37.

CECI, F. et al. Um modelo baseado em casos e ontologia para apoio à tarefa intensiva em conhecimento de classificação com foco na análise de sentimentos. 2015. Citado na página 23.

CERQUEIRA, E. O. d.; ANDRADE, J. C. d.; POPPI, R. J.; MELLO, C. et al. Redes neurais e suas aplicações em calibração multivariada. *Química Nova*, SciELO Brasil, 2001. Citado na página 15.

CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae)?–arguments against avoiding rmse in the literature. *Geoscientific model development*, Copernicus GmbH, v. 7, n. 3, p. 1247–1250, 2014. Citado na página 19.

COHEN, P.; POTCHTER, O.; MATZARAKIS, A. Daily and seasonal climatic conditions of green urban open spaces in the mediterranean climate and their impact on human comfort. *Building and Environment*, Elsevier, v. 51, p. 285–295, 2012. Citado na página 8.

CONTI, J. B. Clima e meio ambiente. 1998. Citado na página 7.

COSTA, M. G. F. Redes neurais convolucionais na saúde. *Journal of Health Informatics*, v. 9, n. 4, 2017. Citado na página 30.

CROSTA, L.; EDWARDS, A.; WANG, R.; REIS-JORGE, J.; MUDALIAR, M. How international online students from a professional doctorate in education are using social media and artificial intelligence tools into the thesis stage? In: *Proceedings of EduLearn18 Conference 2018*. [S.l.: s.n.], 2018. p. 1461–1469. Citado na página 10.

DEHAVEN, V. R. Machine learning: Future capabilities and their implications. The University of Arizona., 2014. Citado na página 11.

DU, Y.; WANG, W.; WANG, L. Hierarchical recurrent neural network for skeleton based action recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2015. p. 1110–1118. Citado na página 21.

DUA, R.; GHOTRA, M. S.; PENTREATH, N. *Machine Learning with Spark*. [S.l.]: Packt Publishing Ltd, 2017. Citado 2 vezes nas páginas 19 e 20.

DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, v. 12, n. Jul, p. 2121–2159, 2011. Citado na página 19.

ESPINOSA, M. M.; PRADO, S. M.; GHELLERE, M. Uso do modelo sarima na previsão do número de focos de calor para os meses de junho a outubro no estado de mato grosso. *Ciência e Natura*, v. 32, n. 2, p. 07–21, 2010. Citado na página 32.

FABRE, G. I. *Modeling neural connectivity of Caenorhabditis Elegans*. Tese (Doutorado) — Massachusetts Institute of Technology, 2017. Citado 2 vezes nas páginas 20 e 21.

FERNEDA, E. Redes neurais e sua aplicação em sistemas de recuperação de informação. *Ciência da Informação*, SciELO Brasil, v. 35, n. 1, 2006. Citado na página 13.

FERREIRA, R. P.; MARTINIANO, A.; SASSI, R. J. Reconhecimento de padrões em quadrados mágicos com rede neural artificial. *LINKSCIENCEPLACE-Interdisciplinary Scientific Journal*, v. 6, n. 1, 2019. Citado na página 20.

FIALHO, E. S. Ilha de calor: Reflexões acerca de um conceito (heat island: reflections on a concept). *Acta Geográfica*, p. 61–76, 2012. Citado na página 8.

FRANCO, F. *Configuração urbana e sua Interferência no microclima local: estudo de caso no bairro do Porto em Cuiabá-MT. Cuiabá. 137p.* Tese (Doutorado) — Dissertação (Mestrado em Física e Meio Ambiente)-Universidade de Mato Grosso, 2010. Citado 2 vezes nas páginas 6 e 7.

FROTA, A. B.; SCHIFFER, S. R. *Manual de conforto térmico*. [S.l.]: Studio Nobel, 1995. Citado na página 7.

GAO, P.; YU, L.; WU, Y.; LI, J. Low latency rnn inference with cellular batching. In: ACM. *Proceedings of the Thirteenth EuroSys Conference*. [S.l.], 2018. p. 31. Citado na página 21.

GARCÍA, P. F. M. *Características climáticas de los invernaderos de plástico*. [S.l.], 1978. Citado na página 6.

GARTLAND, L. *Ilhas de calor: como mitigar zonas de calor em áreas urbanas*. [S.l.]: Oficina de Textos, 2011. Citado na página 8.

GERS, F. A.; ECK, D.; SCHMIDHUBER, J. Applying lstm to time series predictable through time-window approaches. In: *Neural Nets WIRN Vietri-01*. [S.l.]: Springer, 2002. p. 193–200. Citado na página 28.

GIVONI, B. Urban design in different climates. World meteorological organization, 1989. Citado na página 9.

GONG, C.; QI, L.; HEMING, L.; KARIMIAN, H.; YUQIN, M. Spatio-temporal simulation and analysis of regional ecological security based on lstm. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus GmbH, v. 4, p. 153, 2017. Citado na página 50.

GORGENS, E. B.; LEITE, H. G.; SANTOS, H. do N.; GLERIANI, J. M. Estimação do volume de árvores utilizando redes neurais artificiais. *Revista Árvore*, Universidade Federal de Viçosa, v. 33, n. 6, p. 1141–1147, 2009. Citado na página 20.

GRACIOSO, L. d. S.; SIMIONATO, A. C.; MACHADO, L. M. O.; SIMÕES, M. d. G. M. Indexação automática de imagens na web: tendências e desafios no contexto deep learning. *Revista Ibero-americana de Ciencia de la Información (RICI)*, Universidade de Brasília-UnB, v. 11, n. 2, p. 541–561, 2018. Citado 3 vezes nas páginas 11, 22 e 23.

GRAVES, A.; LIWICKI, M.; FERNÁNDEZ, S.; BERTOLAMI, R.; BUNKE, H.; SCHMIDHUBER, J. A novel connectionist system for unconstrained handwriting recognition. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 31, n. 5, p. 855–868, 2008. Citado na página 24.

GRAVES, A.; SCHMIDHUBER, J. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, Elsevier, v. 18, n. 5-6, p. 602–610, 2005. Citado 2 vezes nas páginas 24 e 25.

GRIM, L. F. L.; BARAJAS, J. A. B.; GRADVOHL, A. L. S. Implementações paralelas para o algoritmo online sequential extreme learning machine aplicado à previsão de material particulado. *Revista Brasileira de Computação Aplicada*, v. 11, n. 2, p. 13–21, 2019. Citado na página 12.

GUENICHE, T.; FOURNIER-VIGER, P.; TSENG, V. S. Compact prediction tree: A lossless model for accurate sequence prediction. In: SPRINGER. *International Conference on Advanced Data Mining and Applications*. [S.l.], 2013. p. 177–188. Citado na página 30.

HAYKIN, S. *Redes neurais: princípios e prática*. [S.l.]: Bookman Editora, 2007. Citado 2 vezes nas páginas 12 e 17.

HEATON, J. *Ian goodfellow, yoshua bengio, and aaron courville: Deep learning*. [S.l.]: Springer, 2018. Citado na página 23.

HOWELL, D. C. *Statistical methods for psychology*. [S.l.]: Cengage Learning, 2012. Citado na página 31.

JACKSON, P. C. *Introduction to artificial intelligence*. [S.l.]: Courier Dover Publications, 2019. Citado na página 10.

JAEGER, H. The "echo state" approach to analysing and training recurrent neural networks-with an erratum note. *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, v. 148, n. 34, p. 13, 2001. Citado na página 21.

JOST, I. Aplicação de deep learning em dados refinados para mineração de opiniões. Universidade do Vale do Rio dos Sinos, 2015. Citado na página 22.

JUNIOR, L. A. P. et al. On the training algorithms for restricted boltzmann machine-based models. Universidade Federal de São Carlos, 2018. Citado na página 11.

KAREVAN, Z.; SUYKENS, J. A. Spatio-temporal stacked lstm for temperature prediction in weather forecasting. *arXiv preprint arXiv:1811.06341*, 2018. Citado na página 50.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. Citado na página 19.

KOLBJØRNSRUD, V.; AMICO, R.; THOMAS, R. J. How artificial intelligence will redefine management. *Harvard Business Review*, v. 2, 2016. Citado na página 10.

KONAR, A. *Artificial intelligence and soft computing: behavioral and cognitive modeling of the human brain*. [S.l.]: CRC press, 2018. Citado na página 10.

KOPILER, A. A.; SILVA, V. N. A. L.; SILVA, L. R. A. d. A.; FONSECA, B. L. d. C. Redes neurais artificiais e suas aplicações no setor elétrico. *Revista de Engenharia da Faculdade Salesiana*, 2019. Citado na página 13.

KOTU, V.; DESHPANDE, B. *Data Science: Concepts and Practice*. [S.l.]: Morgan Kaufmann, 2018. Citado na página 22.

KRAUSE, J.; JOHNSON, J.; KRISHNA, R.; FEI-FEI, L. A hierarchical approach for generating descriptive image paragraphs. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2017. p. 317–325. Citado na página 27.

LEAL, F. A.; MIGUEL, E. P.; MATRICARDI, E. A. T.; PEREIRA, R. S. Redes neurais artificiais na estimativa de volume em um plantio de eucalipto em função de fotografias hemisféricas e número de árvores. *Revista Brasileira de Biometria*, v. 33, n. 2, p. 233–249, 2015. Citado na página 17.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436, 2015. Citado na página 22.

LI, Z.; YANG, Y.; FARAGGI, E.; ZHAN, J.; ZHOU, Y. Direct prediction of profiles of sequences compatible with a protein structure by neural networks with fragment-based local and energy-based nonlocal profiles. *Proteins: Structure, Function, and Bioinformatics*, Wiley Online Library, v. 82, n. 10, p. 2565–2573, 2014. Citado na página 23.

LIMA, I.; PINHEIRO, C. A.; SANTOS, F. A. O. *Inteligência artificial*. [S.l.]: Elsevier Brasil, 2016. v. 1. Citado na página 11.

LIMA, J. d. J. A. d. *Fusão de sensores para obtenção de dados de produtividade em colhedora de cana-de-açúcar*. Tese (Doutorado) — Universidade de São Paulo, 2019. Citado na página 12.

LIMA, J. E. C.; CASTRO, L. F. de; CARTAXO, G. A. A. Aplicação do modelo sarima na previsão de demanda no setor calçadista/application of the sarima model in the forecast for demand in the footwear sector. *ID on line REVISTA DE PSICOLOGIA*, v. 13, n. 46, p. 892–913, 2019. Citado na página 32.

LIRA, S. A.; NETO, A. C. Coeficientes de correlação para variáveis ordinais e dicotômicas derivados do coeficiente linear de pearson. *Ciência & Engenharia*, v. 15, n. 1/2, p. 45–53, 2006. Citado na página 31.

LIU, J.; ZHANG, T.; HAN, G.; GOU, Y. Td-lstm: temporal dependence-based lstm networks for marine temperature prediction. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 18, n. 11, p. 3797, 2018. Citado na página 57.

LOMBARDO, M. *Análise das mudanças climáticas nas metrópoles o exemplo de São Paulo e Lisboa*. [S.l.]: Da produção ao consumo: impactos socioambientais no espaço urbano [online ..., 2009. Citado na página 5.

LOMBARDO, M. A. *Ilha de calor nas metrópoles: o exemplo de São Paulo*. [S.l.]: Editora Hucitec com apoio de Lalekla SA Comércio e Indústria, 1985. Citado na página 8.

LOPES, A. T.; AGUIAR, E. de; SOUZA, A. F. D.; OLIVEIRA-SANTOS, T. Facial expression recognition with convolutional neural networks: coping with few data and the training sample order. *Pattern Recognition*, Elsevier, v. 61, p. 610–628, 2017. Citado na página 11.

LUO, C.; YANG, D.; HUANG, J.; DENG, Y.-D. Lstm-based temperature prediction for hot-axles of locomotives. In: EDP SCIENCES. *ITM Web of Conferences*. [S.l.], 2017. v. 12, p. 01013. Citado na página 54.

MAITELLI, G.; ZAMPARONI, C.; LOMBARDO, M. Ilha de calor em cuiabá-mt: Uma abordagem de clima urbano. *Encontro Nacional de Estudos sobre Meio Ambiente*, v. 3, p. 561–571, 1991. Citado na página 4.

MARX, V. Machine learning, practically speaking. *Nature methods*, Nature Publishing Group, v. 16, n. 6, p. 463, 2019. Citado na página 11.

MENDONÇA, F.; DANNI-OLIVEIRA, I. M. *Climatologia: noções básicas e climas do Brasil*. [S.l.]: Oficina de textos, 2017. Citado 2 vezes nas páginas 5 e 7.

MERITY, S. J.; SOCHER, R.; BRADBURY, J.; XIONG, C. *Domain specific language for generation of recurrent neural network architectures*. [S.l.]: Google Patents, 2018. US Patent App. 15/953,265. Citado na página 21.

MIRANDA, A. P.; CERETTA, P. S.; LOPES, L. F. D. Estratégias de mercado acionário utilizando previsão de redes neurais em comparação com modelos autorregressivos. *Revista de Administração da Universidade Federal de Santa Maria*, Universidade Federal de Santa Maria, v. 8, n. 1, p. 42–59, 2015. Citado 2 vezes nas páginas 15 e 30.

MOHSEN, H.; EL-DAHSHAN, E.; EL-HORBATY, E.; SALEM, A. Brain tumor type classification based on support vector machine in magnetic resonance images. *Annals Of "Dunarea De Jos" University Of Galati, Mathematics, Physics, Theoretical mechanics, Fascicle II, Year IX (XL)*, n. 1, 2017. Citado na página 23.

MOHSEN, H.; EL-DAHSHAN, E.-S. A.; EL-HORBATY, E.-S. M.; SALEM, A.-B. M. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, Elsevier, v. 3, n. 1, p. 68–71, 2018. Citado 2 vezes nas páginas 22 e 23.

NASCIMENTO, A. d.; RUBIM, I. B.; PEREIRA, E. G. S.; BARROS, R.; RICHTER, M. Classificação da cobertura da terra, utilizando os programas livres: Interimage, weka e quantumgis. *SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO*, v. 16, p. 2258–2265, 2013. Citado na página 9.

NÓBREGA, R. S.; VERÇOSA, T. O microclima e o (des) conforto térmico em ambientes abertos na cidade do recife. *Revista de Geografia (Recife)*, v. 28, n. 1, p. 93–109, 2011. Citado na página 5.

NOGUEIRA, M. C. d. J. A.; DUARTE, L. C.; NOGUEIRA, J. de S. Conforto térmico na escola pública em cuiabá-mt: Estudo de caso. *REMEA-Revista Eletrônica do Mestrado em Educação Ambiental*, v. 14, 2005. Citado na página 4.

NOGUEIRA, M. C. d. J. A.; SANTOS, F. M. d. M. Microclima urbano: Comparação antes e pós intervenções urbanisticas em cuiabá-mt. 2012. Citado na página 5.

OKE, T. Boundary layer climate, methuen & co. *Ltd. New York*, 1978. Citado na página 4.

OKE, T. R. *Boundary layer climates*. [S.l.]: Routledge, 2002. Citado na página 9.

OKE, T. R. et al. Initial guidance to obtain representative meteorological observations at urban sites. Citeseer, 2004. Citado na página 9.

OLIVEIRA, A. A.; SIQUEIRA, P. H.; NISGOSKI, S.; MUNIZ, G. B. D.; FERREIRA, J. H. Identificação de madeiras utilizando a espectrometria no infravermelho próximo e redes neurais artificiais. *TEMA (São Carlos)*, SciELO Brasil, v. 16, n. 2, p. 81–95, 2015. Citado na página 20.

OLIVEIRA, L. A.; COSTA, A. d. Influência da arborização na temperatura e umidade do ar no campus da ufpa. In: *XIV Congresso Brasileiro de Agrometeorologia*. [S.l.: s.n.], 2005. p. 26–27. Citado na página 6.

OMETTO, J. C. Bioclimatologia vegetal. AGRONOMICA CERES,, 1981. Citado na página 6.

PAULA, D. C. d. *Análise termohigrométrica pós intervenções urbanas em Cuiabá-MT. 2017. 104 f.* Tese (Doutorado) — Dissertação Mestrado, 2017. Citado 4 vezes nas páginas 5, 6, 8 e 9.

PAULA, G. A. *Modelos de regressão: com apoio computacional*. [S.l.]: IME-USP São Paulo, 2004. Citado na página 30.

PEREIRA, C.; ARUQUIPA, W.; OLIVEIRA, E. C.; SILVA, C. A. M. Proposta de utilização de redes neurais feedforward multicamadas para a otimização de padrões de recarga do combustível em um reator pwr. *Brazilian Journal of Radiation Sciences*, v. 7, n. 3A, 2019. Citado na página 21.

PONTI, M. A.; COSTA, G. B. P. da. Como funciona o deep learning. *arXiv preprint arXiv:1806.07908*, 2018. Citado 5 vezes nas páginas 12, 16, 18, 23 e 24.

PRAKASAM, C. Land use and land cover change detection through remote sensing approach: A case study of kodaikanal taluk, tamil nadu. *International journal of Geomatics and Geosciences*, Integrated Publishing Association, v. 1, n. 2, p. 150, 2010. Citado na página 9.

PRUNZEL, J.; TOEBE, M.; LOPES, A. B.; MOREIRA, V. S. Modelos de regressão linear múltipla aplicados à avaliação de terrenos urbanos-caso do município de itaqui-rs. *Boletim de Ciências Geodésicas*, v. 22, n. 4, 2016. Citado na página 31.

QUININO, R. C.; REIS, E. A.; BESSEGATO, L. F. O coeficiente de determinação r 2 como instrumento didático para avaliar a utilidade de um modelo de regressão linear múltipla. *Belo Horizonte: UFMG*, 1991. Citado na página 31.

ROCHA, L. M. V.; SOUZA, L. C. L.; CASTILHO, F. J. V. Ocupação do solo e ilha de calor noturna em avenidas marginais a um córrego urbano. *Ambiente Construído*, SciELO Brasil, v. 11, n. 3, p. 161–175, 2011. Citado na página 8.

ROMERO, M. A. B. *Princípios bioclimáticos para o desenho urbano*. [S.l.]: Pro-Editores Associados, 2001. Citado na página 6.

ROSA, J. L. G. Fundamentos da inteligência artificial. 2011. Citado na página 10.

RUIZ, P. R. D. S.; ALMEIDA, C. M. D.; ANJOS, C. S. D. Classificação da cobertura do solo urbano usando árvores de decisão a partir de uma cena worldview-2 para diferentes níveis de legenda. *Geociências (São Paulo)*, v. 37, n. 3, p. 597–609, 2019. Citado na página 11.

RUSSEL, S.; NORVIG, P. Inteligência artificial. 3a edição. *Editora Campus*, 2013. Citado 2 vezes nas páginas 10 e 11.

SALMAN, A. G.; HERYADI, Y.; ABDURAHMAN, E.; SUPARTA, W. Single layer & multi-layer long short-term memory (lstm) model with intermediate variables for weather forecasting. *Procedia Computer Science*, Elsevier, v. 135, p. 89–98, 2018. Citado na página 57.

SANTOS, A. M. d.; SEIXAS, J. M. d.; PEREIRA, B. d. B.; MEDRONHO, R. d. A. Usando redes neurais artificiais e regressão logística na predição da hepatite a. *Revista Brasileira de Epidemiologia*, SciELO Public Health, v. 8, p. 117–126, 2005. Citado na página 17.

SANTOS, F. M. de M.; NOGUEIRA, M. C. d. J. A. Análise da influência da ocupação do solo na variação termo-higrométrica por meio de transectos noturnos em cuiabá-mt. *Caminhos de Geografia*, v. 13, n. 41, 2012. Citado 3 vezes nas páginas 7, 36 e 38.

SENATORE, G. M. *Metodologia para determinação de áreas de iluminação e sombreamento através de um sistema de informações geográficas*. Tese (Doutorado) — Universidade de São Paulo, 2006. Citado na página 5.

SEVERINO, L.; CARDOSO, G. D.; SANTOS, J. dos et al. Método para determinação da área foliar da mamoneira. *Embrapa Algodão-Boletim de Pesquisa e Desenvolvimento (INFOTECA-E)*, Campina Grande: Embrapa Algodão, 2005., 2005. Citado na página 31.

SIEGEL, S.; JR, N. J. C. *Estatística não-paramétrica para ciências do comportamento*. [S.l.]: Artmed Editora, 1975. Citado na página 32.

SILVA, I. d.; SPATTI, D. H.; FLAUZINO, R. A. Redes neurais artificiais para engenharia e ciências aplicadas. *São Paulo: Artliber*, v. 23, n. 5, p. 33–111, 2010. Citado na página 14.

SILVA, J. G. R.; XAVIER, C. R.; VIEIRA, V. da F.; CARVALHO, I. A. de. Estudo comparativo de métricas de ranqueamento em redes complexas utilizando coeficientes de correlaçao. Citado na página 32.

SILVA, M. I.; GUIMARÃES, E. C.; TAVARES, M. Previsão da temperatura média mensal de uberlândia, mg, com modelos de séries temporais. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 12, n. 5, p. 480–485, 2008. Citado na página 32.

SILVEIRA, A. L. R. C. d. Parâmetros bioclimáticos para avaliação de conjuntos habitacionais na região tropical subúmida do brasil. 2007. Citado na página 5.

SIMÕES, A. d. S.; COSTA, A. H. R. Aprendizado não-supervisionado em redes neurais pulsadas de base radial: um estudo da capacidade de agrupamento para a classificação de pixels. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, SciELO Brasil, v. 18, n. 2, p. 251–264, 2007. Citado na página 11.

SOUZA, A. de; IKEFUTI, P. V.; GARCIA, A. P.; SANTOS, D. A.; OLIVEIRA, S. Análise da relação entre o3, no e no2 usando técnicas de regressão linear múltipla. *GEOgraphia*, v. 20, n. 43, p. 124–135, 2018. Citado na página 31.

SOUZA, A. S. d. *Estudos em quimioinformática baseados na estrutura do ligante e do receptor de candidatos a agentes antichagásicos*. Tese (Doutorado) — Universidade de São Paulo, 2019. Citado na página 14.

SOUZA, L. C. L. d.; TENTE, C. M.; GIUNTA, M. B.; NAKATA, C. M. Fator de visão do céu e intensidade de ilhas de calor na escala do pedestre. *Ambiente construído*, Associação Nacional de Tecnologia do Ambiente Construído-ANTAC, p. 155–167, 2010. Citado na página 8.

SUN, Y.; AHMED, A. B.; AMANO, H. Acceleration of deep recurrent neural networks with an fpga cluster. In: ACM. *Proceedings of the 10th International Symposium on Highly-Efficient Accelerators and Reconfigurable Technologies*. [S.l.], 2019. p. 18. Citado na página 21.

SUTSKEVER, I.; VINYALS, O.; LE, Q. V. Sequence to sequence learning with neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2014. p. 3104–3112. Citado na página 28.

TADANO, Y. S.; ALVES, T. A.; SILVA, N. S.; VALADARES, H. S. Impacto da poluição atmosférica e das alterações climáticas na saúde populacional utilizando redes neurais artificiais atmospheric pollution and climate changes impact on population's health using artificial neural. *Revista da Associação Portuguesa de Análise Experimental de Tensões ISSN*, v. 1646, p. 7078, 2017. Citado na página 21.

TSUJI, M.; ISOKAWA, T.; YUMOTO, T.; MATSUI, N.; KAMIURA, N. Heterogeneous recurrent neural networks for natural language model. *Artificial Life and Robotics*, Springer, v. 24, n. 2, p. 245–249, 2019. Citado na página 22.

TUMBO, S. D.; WAGNER, D. G.; HEINEMANN, P. H. Hyperspectral–based neural network for predicting chlorophyll status in corn. *Transactions of the ASAE*, American Society of Agricultural and Biological Engineers, v. 45, n. 3, p. 825, 2002. Citado na página 12.

VAREJÃO, M. Meteorologia e climatologia. *Versão digital*, v. 2, p. 463, 2006. Citado na página 6.

WALTER, O. M. F. C.; HENNING, E.; MORO, G.; SAMOHYL, R. W. Aplicação de um modelo sarima na previsão de vendas de motocicletas. *Exacta*, Universidade Nove de Julho, v. 11, n. 1, p. 77–88, 2013. Citado na página 32.

WANG, J.; CAO, H.; ZHANG, J. Z.; QI, Y. Computational protein design with deep learning neural networks. *Scientific reports*, Nature Publishing Group, v. 8, n. 1, p. 6349, 2018. Citado na página 23.

WANG, Y.; LONG, M.; WANG, J.; GAO, Z.; PHILIP, S. Y. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. In: *Advances in Neural Information Processing Systems*. [S.l.: s.n.], 2017. p. 879–888. Citado na página 54.

WASON, R. Deep learning: Evolution and expansion. *Cognitive Systems Research*, Elsevier, v. 52, p. 701–708, 2018. Citado na página 23.

WYSE, L.; HUZAIFAH, M. Conditioning a recurrent neural network to synthesize musical instrument transients. *arXiv preprint arXiv:1903.10703*, 2019. Citado na página 22.

YANG, Y.; DONG, J.; SUN, X.; LIMA, E.; MU, Q.; WANG, X. A cfcc-lstm model for sea surface temperature prediction. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 15, n. 2, p. 207–211, 2017. Citado na página 49.

ZAVITOSKI, É. V. d. S. *Análise termohigrométrica nos eixos viários centrais da malha urbana original de SINOP-MT. 2018. 67 f.* Tese (Doutorado) — Dissertação Mestrado, 2018. Citado na página 34.

ZHANG, Q.; YANG, S. X.; MITTAL, G. S.; YI, S. Ae—automation and emerging technologies: Prediction of performance indices and optimal parameters of rough rice drying using neural networks. *Biosystems Engineering*, Elsevier, v. 83, n. 3, p. 281–290, 2002. Citado na página 12.

Anexo A: Material complementar

DOCUMENTAÇÃO

Todos os códigos python (comentados) e todos os dados utilizados neste estudo estão disponíveis em:

www.github.com/pgfa

Sumário automático gerado pelo pacote Keras.

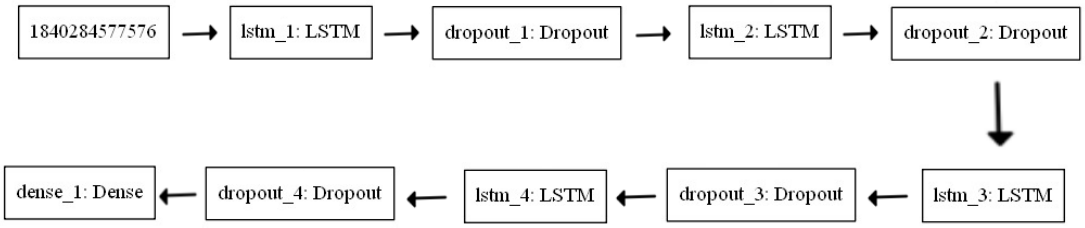


Figura 48: Etapas da criação do modelo criado de Rede Neural do tipo LSTM.