

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ENGENHARIA
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Pedro Henrique Macedo Nascimento

Aplicação de técnicas de aprendizagem de máquinas na previsão de
vertimento em usinas hidrelétricas

Juiz de Fora

2021

Pedro Henrique Macedo Nascimento

Aplicação de técnicas de aprendizagem de máquinas na previsão de
vertimento em usinas hidrelétricas

Dissertação apresentada ao Programa de Pós Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Sistemas de Energia.

Orientador: Prof. Dr. Ivo Chaves da Silva Junior

Coorientador: Prof. Dr. Leonardo Willer de Oliveira

Juiz de Fora

2021

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Sobrenome, Nome do autor.

Título : subtítulo / Autor. – Ano.

23 f. : il.

Orientador: Nome e sobrenome

Coorientador: Nome e sobrenome

Dissertação (Mestrado) – Universidade Federal de Juiz de Fora, Faculdade

X. Nome do Curso ou Programa, Ano.

1. Palavra-chave. 2. Palavra-chave. 3. Palavra-chave. I. Sobrenome,
Nome do orientador, orient. II. Título.

Pedro Henrique Macedo Nascimento

Aplicação de Técnicas de Aprendizagem de Máquinas na Previsão de
Vertimento em Usinas Hidrelétricas

Dissertação apresentada ao Programa de Pós Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: Sistemas de Energia Elétrica.

Aprovada em:

BANCA EXAMINADORA



Prof. Dr. Ivo Chaves da Silva Júnior - Orientador
Universidade Federal de Juiz de Fora



Prof. Dr. Leonardo Willer de Oliveira - Coorientador
Universidade Federal de Juiz de Fora



Profa. Dra. Patrícia Teixeira Leite Asano
Universidade Federal do ABC



Prof. Dr. Igor Delgado de Melo
Universidade Federal de Juiz de Fora

À minha família e amigos.

AGRADECIMENTOS

Agradeço aos meus pais Sebastião Heleno Farinazo Nascimento e Marinez Macedo Nascimento por todo carinho e amor que sempre tiveram por mim, o apoio dado durante toda a minha trajetória no mestrado, sempre me aconselhando e motivando a seguir em frente. Agradeço a minha irmã Danielle, seu marido Rondinelli e meus sobrinhos Arthur e Maria Clara, por todos os momentos felizes que passamos e pelo apoio incondicional.

Agradeço a minha namorada Patrícia Barros Santos por todo seu amor, carinho, amizade, companheirismo e dividir comigo momentos maravilhosos que passamos juntos e me dar forças para superar os momentos difíceis. Ter você na minha vida me torna cada dia melhor e mais feliz. Agradeço ao sogro Raymundo Nonato e minha sogra Mara Leonor por terem me acolhido como um filho e me ajudado nos momentos difíceis.

Ao meu orientador, Dr. Ivo Chaves da Silva Júnior, por todas as oportunidades e confiança, que me fizeram crescer como pessoa e profissionalmente. Agradeço Dr. Leonardo Willer de Oliveira pela orientação e suporte.

Aos meus amigos colegas de trabalho no mestrado Frederico Panoeiro, Ramon Abritta, Vinícius Albuquerque, Guilherme Marins, Othon Avila e Fábio Tavela com os quais dividi desafios e muitos momentos descontraídos que nos dão força para seguir trabalhando com prazer.

Agradeço à Universidade Federal de Juiz de Fora (UFJF) e ao Programa de Pós-Graduação em Engenharia Elétrica (PPEE), a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), ao Instituto Nacional de Energia Elétrica (INERGE), à EDP e à Agência Nacional de Energia Elétrica (ANEEL), os quais, que tornaram possíveis essa pesquisa, a qual está vinculada ao projeto sob código PD-00673-0052/2018.

“O começo de todas as ciências é o espanto de as coisas serem o que são - Aristóteles”

RESUMO

A operação de uma usina hidrelétrica é dependente de diversos fatores como a programação de geração de energia, do volume de água disponível no seu reservatório, as condições do rio a jusante e a segurança das barragens. Um grande desafio da operação é controlar o vertimento da água do reservatório. Embora a ação de vertimento represente a perda de recursos energéticos, esta ação também é uma estratégia poderosa para controlar o nível do reservatório, garantindo a segurança da barragem. A tomada de decisão quanto a essa operação é realizada com antecedência e geralmente se baseia em informações estimadas de nível e demanda. Neste contexto, este trabalho aplica técnicas de aprendizado supervisionado de máquina para prever, cinco horas a frente, a condição operativa de vertimento em uma usina hidrelétrica. Com o objetivo de ser utilizado em tempo real, este método visa auxiliar o operador, de modo que este consiga tomar decisões mais assertivas e seguras, preservando recursos energéticos e promovendo aumento da segurança das barragens e, conseqüentemente, dos trabalhadores e da população que reside às margens do rio a jusante da usina. Floresta Aleatória, Perceptron Multicamadas e a combinação destes algoritmos de aprendizado são utilizados e comparados neste trabalho. A metodologia proposta foi implementada e testada com uma usina hidrelétrica localizada no Rio Tocantins, Brasil, com capacidade de geração de 902,5MW. Os resultados da metodologia demonstraram que a ferramenta tem capacidade de ser um auxílio eficiente aos operadores de uma usina nas tomadas de decisão, visto que os modelos de previsão alcançaram patamares superiores à 99% de acerto nas previsões de vertimento.

Palavras-chave: Aprendizado de Máquina. Previsão. Energia Hidroelétrica. Telemetria. Gerenciamento de Recursos.

ABSTRACT

Operation at a hydroelectric plant is dependent on several factors such as the schedule of power generation, the volume of water available in its reservoir, the conditions of the river downstream, and the safety of the dams. A major challenge of the operation is to control the spillage from the reservoir. Although the spillage action represents the loss of energy resources, this action is also a powerful strategy to control the level of the reservoir, ensuring the safety of the dam. Decision-making regarding this operation is carried out in advance and is generally based on estimated level and demand information. In this context, this work applies supervised machine learning techniques to predict, for five hours to come, the operating condition of pouring in a hydroelectric plant. Intending to be used in real-time, this method aims to assist the operator, so that he can make more assertive and safer decisions, preserving energy resources and promoting increased safety of dams, and consequently, of workers and the population that resides river banks downstream of the plant. Random Forest, Multilayer Perceptron and the combination of these learning algorithms are adopted and compared in this work. The proposed methodology was implemented and tested with a hydroelectric plant located on the Tocantins River, Brazil, with a generation capacity of 902.5MW. The results of the methodology demonstrated that the tool has the capacity to be an efficient aid to the operators of a plant in decision making, since the forecasting models reached levels above 99% of correctness in the spillage forecasts.

Keywords: Machine Learning. Forecast. Hydroelectric Energy. Telemetry. Resource Management.

DIAGRAMAS

LISTA DE DIAGRAMAS

- Diagrama 1 – Metodologia proposta para desenvolvimento e avaliação de modelos de previsão de vertimento utilizando aprendizagem de máquina. . . 34

LISTA DE FIGURAS

Figura 1 – Dilema do operador.	15
Figura 2 – Tipos de aprendizado de máquina, com destaque ao método utilizado neste trabalho.	22
Figura 3 – Exemplo simples da estrutura de uma árvore decisão binária, que realiza a classificação da classe 0 e da classe 1, representadas respectivamente pelas cores preta e azul.	24
Figura 4 – Imagem da técnica <i>Bootstrap</i> aplicada para a criação de seis subconjuntos compostos por um número de dados igual à metade do total de dados que formam o conjunto inicial.	27
Figura 5 – Estrutura de um perceptron.	28
Figura 6 – Estrutura de camadas de um perceptron.	28
Figura 7 – Ilustrações de funções de ativação.	29
Figura 8 – Ilustração de uma divisão do banco de dados em grupos de treinamento, validação e teste.	30
Figura 9 – Representação da combinação de classificadores por maioria simples e por unanimidade.	31
Figura 10 – Matriz de confusão.	32
Figura 11 – Exemplo de validação cruzada com a divisão do banco de dados em oito subconjuntos.	33
Figura 12 – Ilustração do balanceamento de dados.	37
Figura 13 – Ilustração dos modelos preditores.	38
Figura 14 – Exemplo de configuração de dados de vertimento preditos para 5 horas futuras.	39
Figura 15 – Ilustração da validação cruzada adaptada para dados de dois anos. . .	41
Figura 16 – Localização das estações de telemetria.	43
Figura 17 – Estrutura de entradas e saídas do modelo de previsão do vertimento. .	46

LISTA DE GRÁFICOS

Gráfico 1	– Valores de correlação do nível medidos nas estações de telemetria com o nível a montante da UHE.	44
Gráfico 2	– Valores de correlação do nível medidos nas estações de telemetria com o vertimento da UHE.	45
Gráfico 3	– O gráfico boxplot com os valores de acurácia I para a validação cruzada feita com os modelos do Quadro 1.	53
Gráfico 4	– O gráfico boxplot com os valores de acurácia II para a validação cruzada feita com os modelos do Quadro 1.	54
Gráfico 5	– O gráfico boxplot com os valores de precisão para a validação cruzada feita com os modelos do Quadro 1.	54
Gráfico 6	– O gráfico boxplot com os valores de revocação para a validação cruzada feita com os modelos do Quadro 1.	55
Gráfico 7	– O gráfico boxplot com os valores de pontuação F1 para a validação cruzada feita com os modelos do Quadro 1.	56
Gráfico 8	– O gráfico boxplot com os valores de tempo de treinamento para a validação cruzada feita com os modelos do Quadro 1.	57
Gráfico 9	– O gráfico boxplot com os valores de acurácia I para a validação cruzada feita com os agrupamentos de modelos.	59
Gráfico 10	– O gráfico boxplot com os valores de acurácia II para a validação cruzada feita com os agrupamentos de modelos.	60
Gráfico 11	– O gráfico boxplot com os valores de precisão para a validação cruzada feita com os agrupamentos de modelos.	61
Gráfico 12	– O gráfico boxplot com os valores de revocação para a validação cruzada feita com os agrupamentos de modelos.	61
Gráfico 13	– O gráfico boxplot com os valores de pontuação F1 para a validação cruzada feita com os agrupamentos de modelos.	62

LISTA DE QUADROS

Quadro 1 – Descrição dos modelos e do tratamento realizado no banco de dados de treinamento.	52
Quadro 2 – Definição dos agrupamentos de modelos de aprendizagem.	57
Quadro 3 – Definição dos resultados para combinação de modelos com decisão por maioria simples.	58
Quadro 4 – Combinação dos resultados para combinação de modelos com decisão por unanimidade.	59
Quadro 5 – Quantidade de horas sem resposta de previsão para os modelos com decisão por unanimidade.	63

LISTA DE TABELAS

Tabela 1 – Localização das ETs em coordenadas geográficas.	17
Tabela 2 – Valores médios com cinco diferentes formações - primeira análise de sensibilidade RF.	47
Tabela 3 – Valores médios com cinco diferentes formações - segunda análise de sensibilidade RF.	48
Tabela 4 – Valores de acurácia I, II e Média para os cinco treinamentos/testes realizados na segunda análise de sensibilidade do modelo RF.	49
Tabela 5 – Média dos resultados de cinco treinamentos diferentes para cada configuração do modelo MLP.	50
Tabela 6 – Valores de acurácia I, II e Média para os cinco treinamentos/testes realizados na segunda análise de sensibilidade do modelo MLP.	52
Tabela 7 – Resultados dos treinamentos/testes com um grupo de validação cruzada com 71,41% de amostras de não vertimento, 27,14% de amostras de vertimento e 1,55% de amostras com mudança na condição de vertimento.	64
Tabela 8 – Resultados dos treinamentos/testes com um grupo de validação cruzada com 79,73% de amostras de não vertimento, 19,18% de amostras de vertimento e 1,09% de amostras com mudança na condição de vertimento.	64

LISTA DE ABREVIATURAS E SIGLAS

UFJF	Universidade Federal de Juiz de Fora
UHE	Usina Hidrelétrica
ONS	Operador Nacional do Sistema Elétrico
EPE	Empresa de Pesquisa Energética
SIN	Sistema Interligado Nacional
ANA	Agência Nacional de Águas
SNIRH	Sistema Nacional de Informações sobre Recursos Hídricos
ET	Estações de Telemetria
CH	Controle Hidráulico
ML	Aprendizado de Máquina (do inglês <i>Machine Learning</i>)
RF	Floresta Aleatória (do inglês <i>Random Forest</i>)
DT	Árvore de Decisão (do inglês <i>Decision Tree</i>)
MLP	Perceptron Multicamadas (do inglês <i>Perceptron Multilayer</i>)
OOB	Fora do Saco (do inglês <i>Out-Of-Bag</i>)
AI	Inteligência Artificial (do inglês <i>Artificial Intelligence</i>)
ANN	Redes Neurais Artificiais (do inglês <i>Artificial Neural Network</i>)
SVM	Máquina de Vetores de Suporte (do inglês <i>Support Vector Machine</i>)
CNN	Redes Neurais Convolucionais (do inglês <i>Convolutional Neural Network</i>)
ConvLSTM	Memória de Curto e Longo Prazo Convolucional (do inglês <i>Convolutional Long Short-Term Memory</i>)
CRA	Algoritmo de Busca do Cuco (do inglês <i>Cuckoo Research Algorithm</i>)
WOA	Algoritmo de Otimização da Baleias (do inglês <i>Whale Optimization Algorithm</i>)
ANFIS	Sistema de Inferência Neuro-Fuzzy Adaptativo (do inglês <i>Adaptive Neuro-Fuzzy Inference System</i>)
ReLU	Ativação Linear Retificada (do inglês <i>Rectified Linear Unit</i>)
ELU	Unidade Linear Exponencial (do inglês <i>Exponential Linear Unit</i>)
TN	Verdadeiro Negativo (do inglês <i>True Negative</i>)
TP	Verdadeiro Positivo (do inglês <i>True Positive</i>)
FP	Falso Positivo (do inglês <i>False Positive</i>)
FN	Falso Negativo (do inglês <i>False Negative</i>)
MSE	Erro Médio Quadrático (do inglês <i>Mean Squared Error</i>)
R ²	Coefficiente de correlação
CV	Validação Cruzada (do inglês <i>Cross-Validation</i>)

SUMÁRIO

1	INTRODUÇÃO	15
1.1	DESCRIÇÃO DO PROBLEMA	17
1.2	REVISÃO BIBLIOGRÁFICA	18
1.3	MOTIVAÇÕES	20
1.4	OBJETIVOS	20
1.5	PUBLICAÇÕES	20
1.5.1	Publicações Diretas	20
1.5.2	Publicações Correlatas	21
1.5.3	Publicações Paralelas	21
1.6	ORGANIZAÇÃO	21
2	APRENDIZADO DE MÁQUINA	22
2.1	TIPOS DE APRENDIZADO	22
2.2	MODELOS DE APRENDIZADO SUPERVISIONADO	23
2.2.1	Floresta Aleatória	23
2.2.2	Perceptron Multicamadas	27
2.2.3	Combinação de Classificadores	30
2.3	AVALIAÇÃO	31
3	METODOLOGIA	34
3.1	TRATAMENTO DOS DADOS	35
3.2	CORRELAÇÃO DE DADOS	35
3.3	AJUSTE DOS DADOS	36
3.4	TREINAMENTO E TESTE DOS MODELOS	37
3.5	ANÁLISE DE DESEMPENHO	40
4	RESULTADOS E DISCUSSÃO	42
4.1	ANÁLISE DE CORRELAÇÃO	42
4.2	ESTRUTURA DO MODELO DE PREVISÃO DE VERTIMENTO	45
4.3	SENSIBILIDADE DO MODELO <i>RANDOM FOREST</i>	46
4.4	SENSIBILIDADE DO MODELO <i>MULTILAYER PERCEPTRON</i>	48
4.5	SENSIBILIDADE À MODIFICAÇÃO NO BANCO DE TREINAMENTO	51
4.6	ANÁLISE DO AGRUPAMENTO DE MODELOS	57
5	CONCLUSÕES E TRABALHOS FUTUROS	65
5.1	CONCLUSÃO	65
5.2	TRABALHOS FUTUROS	66
	REFERÊNCIAS	67
	ANEXO A – Certificado	72

1 INTRODUÇÃO

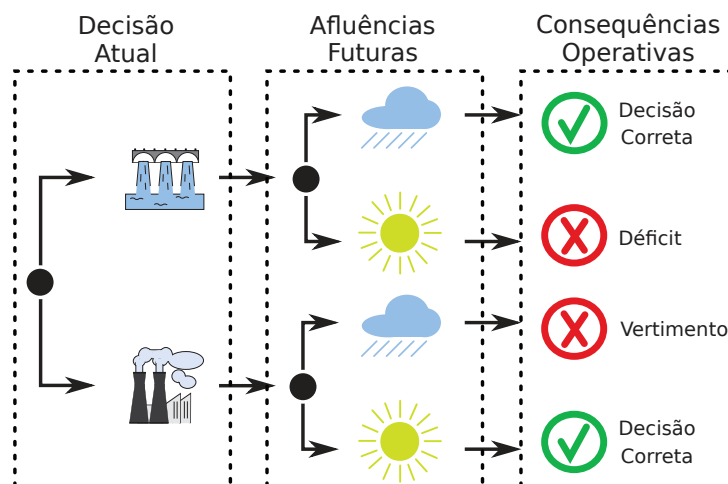
A geração de energia elétrica por meio de Usinas Hidrelétricas (UHE) tem grande representatividade no mundo. Em países com alto potencial hídrico, as usinas hidrelétricas costumam ser a principal fonte de geração de energia. Isso se deve ao baixo custo de produção de produção de energia com esta fonte, principalmente quando comparada às gerações termelétricas (IEA, 2020).

A matriz elétrica brasileira, até o final de 2021, contará com 63,9% da sua capacidade de geração instalada proveniente de fontes hídricas e 21,4% de geração por outras fontes renováveis. Em 2019, 74,3% da energia consumida no Brasil era proveniente de fontes hídricas e no ano de 2020 esse valor passou para 75% (ONS, 2021). Esses dados mostram a importância das UHEs para o cenário energético brasileiro, que possui o planejamento da sua expansão sob responsabilidade da Empresa de Pesquisa Energética (EPE).

No Brasil, o Operador Nacional do Sistema Elétrico (ONS) administra o Sistema Interligado Nacional (SIN), garantindo a coordenação, o controle e a otimização da operação na geração e transmissão de energia.

Neste cenário com predomínio de hidrelétricas, uma decisão, no presente, por manter as hidrelétricas turbinando água do reservatório, será uma decisão correta se houver uma alto índice de precipitação no futuro, caso este índice seja baixo, a decisão será equivocada, visto que seria necessário a complementação da geração energética com fontes de maior custo de geração, como as termelétricas. Assim como uma decisão por pouco turbinamento de água, no presente, será uma boa decisão se o índice de precipitação for baixo no futuro, e será uma decisão ruim se houver um elevado índice de precipitação (CASTRO, 2012). Estas decisões por de turbinamento presente tendo em vista as consequências futuras é conhecido como o dilema do operador, e é retratado na Figura 1.

Figura 1 – Dilema do operador.



Fonte: Elaborado pelo autor (2021).

O planejamento envolvendo a operação das UHEs é uma tarefa complexa que envolve diversos fatores, como minimizar custos de produção da energia, reduzir o risco de não atendimento à demanda, segurança, preservação do volume de água no reservatório, questões ambientais e outros.

Dentre as equações contempladas no planejamento da operação, tem-se a equação do balanço hídrico, apresentada na Equação 1.1. Esta equação descreve o comportamento do reservatório de uma usina (WESTIN *et al.*, 2021; RAMOS, 2015).

$$v_{t+1}^k = v_t^k + C(\text{Vinc}_t^k + \sum_{m \in \Theta_m} (q_t^m + s_t^m) - q_t^k - s_t^k) \quad \forall k \in K \quad (1.1)$$

Onde,

K é o conjunto de todas as hidroelétricas no sistema

C é uma contante para conversão de m^3/s para hm^3

v_{t+1}^k é o volume da UHE k no final do período t

v_t^k é o volume da UHE k no início do período t

Vinc_t^k é o fluxo incremental de vazão afluyente da UHE k no período t

Θ_m é o número de UHEs à montante da UHE k

q_t^m é a vazão turbinada pela UHE m no período t

s_t^m é a vazão vertida pela UHE m no período t

q_t^k é a vazão turbinada pela UHE k no período t

s_t^k é a vazão vertida pela UHE k no período t

O vertimento em uma UHE, pode ser usado como uma variável de controle para o atendimento à equação de balanço hídrico. E também tem um papel muito importante no que diz respeito à segurança e eficiência na produção de energia elétrica.

Neste trabalho, para a previsão da necessidade de vertimento em uma UHE, são aplicadas duas técnicas de aprendizagem de máquina, Perceptron Multicamadas (*Multilayer Perceptron* - MLP) e Floresta Aleatória (*Random Forest* - RF), almejando com este estudo, uma análise das previsões e uma análise comparativa entre os métodos.

1.1 DESCRIÇÃO DO PROBLEMA

A necessidade iniciar o vertimento em uma UHE pode ser proveniente de diferentes condições, como o baixo volume de água turbinada, um alto índice de precipitação e de volume de água que chega ao reservatório.

A condição de operação do vertimento ocorre através da abertura das comportas da usina hidrelétrica. Assim, a definição da ocorrência desta condição operativa é de extrema importância, por questões operacionais e/ou de segurança. Quando o operador não abre as comportas, a usina pode atingir níveis críticos em seu reservatório, levando à necessidade de verter um volume muito grande de água em um curto espaço de tempo, o que pode causar inundações nas comunidades ribeirinhas a jusante da barragem. Por outro lado, quando o operador abre as comportas, sem ser realmente necessário, há desperdício de recursos hídricos para a produção de energia.

A Agência Nacional das Águas e Saneamento Básico (ANA) é responsável pelo monitoramento e eventos críticos, segurança hídrica, gestão integrada de recursos hídricos entre outras funções. A ANA por meio do Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH) disponibiliza informações coletadas de Estações de Telemetria (ET) associadas às usinas em território nacional. Essas ETs comumente apresentam informações de nível, vazão e precipitação em certo ponto do rio onde estão localizadas. Uma UHE pode ter sob sua responsabilidade ETs a montante e também a jusante de sua barragem.

Neste cenário, o presente trabalho propõe a aplicação de um conjunto de dados de ET e CH referentes à UHE Luis Eduardo Magalhães (UHE Lajeado) localizada no Rio Tocantins, Brasil, na implementação de um modelo de aprendizado de máquina para previsão de vertimento.

A UHE Lajeado apresenta dez estações de telemetria sob sua administração, sendo seis localizadas a montante, uma localizada na barragem e outras três localizadas a jusante, como apresentadas na Tabela 1.

Tabela 1 – Localização das ETs em coordenadas geográficas.

Estação de Telemetria	Latitude	Longitude
Jacinto	-11.9817	-48.6569
Jerônimo	-11.7578	-47.8358
Ipueiras	-11.2467	-48.4586
Jurupary	-11.1519	-48.5158
Areias	-10.8950	-48.3483
Mangues	-10.3483	-48.6369
Barramento	-9.7603	-48.3697
Jusante	-9.7442	-48.3625
Lucena	-9.6917	-48.3689
Tocantínia	-9.5669	-48.3792

1.2 REVISÃO BIBLIOGRÁFICA

A precisa gestão dos recursos hídricos disponíveis e também a segurança dos equipamentos, funcionários e população residente na região a jusante da usina é de fundamental importância para a operação. Neste contexto, é possível identificar diversos estudos sobre o funcionamento de UHEs na literatura especializada. Em (SINGH; SINGAL, 2017) é realizado uma revisão quanto a questões operacionais de usinas hidrelétricas, que contribuem na minimização do custo de operação e minimizando efeitos no causados ao ambiente. Uma revisão de métodos heurísticos utilizados para otimização de custos de geração para programação da operação de curto prazo é apresentada (NAZARI-HERIS; MOHAMMADI-IVATLOO; GHAREHPETIAN, 2017).

Em (FINARDI; TAKIGAWA; BRITO, 2016), são realizados estudos de comprometimento de unidades hidrelétricas. Em (FREDO; FINARDI; MATOS, 2019), o foco está no problema de escalonamento de geração de longo prazo. (ABRITTA *et al.*, 2020) realizou a otimização de curto prazo da UHE considerando a indicação de vertimento pelo otimizador. A meta-heurística de otimização do enxame de partículas e fluxo da rede são usados em (ARAGÃO; ASANO; RABÊLO, 2020) para obter a solução ótima para as regras de operação do reservatório, através da transferência de água entre bacias.

Uma importante linha de pesquisa é a de previsão de eventos naturais que possam impactar diretamente na produção energética. O estudo apresentado em (BOSSANYI, 1985) utiliza filtro de Kalman na previsão de ventos, ainda sobre previsão de ventos, tem-se a utilização de Máquina de Vetores de Suporte (*Support Vector Machine* - SVM) em (MOHANDÉS *et al.*, 2004) e Redes Neurais Convolucionais (*Convolutional Neural Network* - CNN) aplicada por (HARBOLA; COORS, 2019). Outros trabalhos visam a previsão de chuvas, como em (ORTIZ-GARCÍA; SALCEDO-SANZ; CASANOVA-MATEO, 2014) que utiliza SVM e (KIM *et al.*, 2017) que aplica Memória de Curto e Longo Prazo Convolucional (*Convolutional Long Short-Term Memory* - ConvLSTM), para realizar as previsões.

Eventos naturais diretamente ligados à hidrografia e seus impactos na produção energética também são muito estudados. (RASOULI; HSIEH; CANNON, 2012) usou três modelos diferentes de aprendizado de máquina para prever o fluxo diário em uma bacia hidrográfica no Canadá. E os dados usados em sua pesquisa são dados de previsão do tempo, índices e dados meteorológicos. No mesmo campo de pesquisa, os autores de (BELVEDERESI *et al.*, 2020) usaram o Sistema de Inferência Neuro-Fuzzy Adaptativo (*Adaptive Neuro-Fuzzy Inference System* - ANFIS), como entrada simples e sequencial, para fazer previsões de fluxo de rios. Em (ZHOU *et al.*, 2020), é proposto um modelo de previsão de flutuação do nível de água usando MLP.

A operação de vertimento nas usinas é de grande importância para preservar as boas condições estruturais, garantir a segurança e ter eficiência na utilização de recursos. Os trabalhos de (LIMA, 2006; SILVA, 2019b) abordam o vertimento como característica

importante na operação das usinas. Os nível da água nos reservatórios está diretamente associado à segurança da barragem. Diversos estudos são realizados com o objetivo de firmar essa segurança. Em (SHANG *et al.*, 2019), aplica-se o nível previsto para a UHE e os autores relatam que este estudo contribui para garantir a segurança da barragem, devido à manutenção do nível em limites de segurança. Uma das formas de controlar o nível dos reservatórios é por meio do vertimento. (TALIB; HASAN, 2010) apresentou a implementação de Redes Neurais Artificiais (*Artificial Neural Network* - (ANN), para a previsão de eventos mensais de vertimento de barragens para uma UHE localizada no norte da Malásia.

Como afirmado em (GOODFELLOW; BENGIO; COURVILLE, 2016), as pessoas, em geral, esperam que a Inteligência Artificial (*Artificial Intelligence* - AI) realize automação de rotinas, seja capaz de compreender um áudio e até mesmo interagir, que faça diagnósticos de doenças e impulse as pesquisas científicas. A AI demonstrou sua eficácia na resolução de problemas que podem ser descritas por regras matemáticas. No entanto, situações que são fáceis de serem executadas por humanos, podem ser difíceis de descrever formalmente ou equacionar, como exemplo, o reconhecimento da voz de uma pessoa específica pode ser realmente desafiador. Muitos áreas de pesquisas tiveram seu estado da arte aprimorado por meio do aprendizado de máquina, como em problemas de série temporal (VOYANT *et al.*, 2017; YAO *et al.*, 2017; CHEN *et al.*, 2018), sensoriamento remoto (MAXWELL; WARNER; FANG, 2018) e demais aplicações (LIAKOS *et al.*, 2018; MOSAVI; OZTURK; CHAU, 2018).

Perceptron de multicamadas, é uma ANN *feed-forward* com uma ou mais camadas ocultas (HAYKIN, 1994). O MLP é aplicado em vários campos de estudo. Na medicina, por exemplo, é amplamente utilizado no reconhecimento de doenças por meio da análise de imagens (HOSSEINZADEH *et al.*, 2020; ABDAR; YEN; HUNG, 2018). Em engenharia química, pode ser usado para estimar os pesos moleculares de compostos químicos (HEMMATI-SARAPARDEH *et al.*, 2018). Esse modelo de aprendizado de máquina também tem sido utilizado em trabalhos e pesquisas envolvendo recursos hídricos. Em (ALI *et al.*, 2017), o MLP é usado para prever períodos de seca em uma região específica do Paquistão. A previsão das chuvas sazonais na bacia do rio Tarim, China, foi o objetivo do (HARTMANN *et al.*, 2016). O MLP pode ser utilizado de forma híbrida, com outras técnicas de otimização, visando uma melhor eficiência dos modelos em estudo. (PHITAKWINAI; AUEPHANWIRIYAKUL; THEERA-UMPON, 2016) combinou o MLP com o Algoritmo de Busca do Cuco (*Cuckoo Research Algorithm* - CRA), para prever o nível do rio Ping, Tailândia, com 7h de antecedência. Em (DIOP *et al.*, 2020), o algoritmo de otimização baseado no comportamento das baleias (*Whale Optimization Algorithm* - WOA) é usado em conjunto com o MLP para realizar a previsão de precipitação anual em uma determinada região do Senegal.

O método RF, é derivado da Árvore de Decisão (*Decision Tree* - DT), que se

baseia na hierarquia e na importância de seus ramos. O RF agrega um grande número de resultados das árvores de decisão treinadas com subconjuntos de treinamento e variáveis aleatórias, resultando em uma combinação dos resultados individuais das árvores treinadas (BREIMAN, 2001). Esta metodologia tem capacidade para resolver problemas com diferentes objetivos, como problemas de classificação, agrupamento e regressão. Em (SARICA; CERASA; QUATTRONE, 2017), a técnica de RF é utilizada na predição da doença de Alzheimer por meio de análises de neuroimagem. Na linha de análise de imagens, (BELGIU; DRĂGUȚ, 2016) apresentou uma revisão da aplicação desta metodologia em sensoriamento remoto. Em (IANNACE; CIABURRO; TREMATERRA, 2019), é apresentada a previsão de ruído de turbinas eólicas. Em (NOSHAD *et al.*, 2019), é apresentada a detecção de falhas em redes de sensores sem fio com classificação RF.

1.3 MOTIVAÇÕES

A motivação para este trabalho surgiu como uma demanda de um projeto de P&D e da constante necessidade de aprimorar o uso de recursos hídricos e aumentar a segurança. Os momentos em que a operação da usina precisa abrir as comportas e colocar a usina em condição de vertimento e os momentos que precisam fechar as comportas são de grande importância e de dificuldade, pois são momentos que ocorrem poucas vezes. Neste contexto busca-se conter desperdício de recursos energéticos e a melhoria da segurança na estrutura da barragem, para os funcionários que trabalham na usina e para a comunidade ribeirinha que margeiam o rio a jusante da barragem.

1.4 OBJETIVOS

O objetivo principal deste estudo é a elaboração de um modelo de previsão da necessidade ou não de vertimento em uma usina, com base em aprendizagem de máquina. De modo que a ferramenta seja capaz de auxiliar o operador da UHE na tomada de decisão quanto ao vertimento.

1.5 PUBLICAÇÕES

1.5.1 Publicações Diretas

- NASCIMENTO, P. H. M. et al. “Spillage Forecast Models in Hydroelectric Power Plants Using Information from Telemetry Stations and Hydraulic Control.” *Energies* 14.1 (2021): 184.
- NASCIMENTO, P. H. M. et al. “Spillage Forecast in Hydroelectric Power Plants via Machine Learning.” VIII Simpósio Brasileiro de Sistemas Elétricos (SBSE), 2020.

Melhor trabalho apresentado na seção *SBSE 3.4 - Inteligência Computacional Aplicada a Sistemas Elétricos*, vide Anexo A.

1.5.2 Publicações Correlatas

- COSTA, M. A. R. et al. “Previsão do Nível a Montante de uma Usina Hidrelétrica Baseado em Dados de Telemetria.” LII Simpósio Brasileiro de Pesquisa Operacional (SBPO), 2020.

1.5.3 Publicações Paralelas

- NASCIMENTO, P. H. M. et al. “Estimativa de Perdas Técnicas em Sistemas de Distribuição de Energia via Aprendizagem de Máquina.” LII Simpósio Brasileiro de Pesquisa Operacional (SBPO), 2020.
- NASCIMENTO, P. H. M. et al. “Machine Learning Applied to Forecasting Technical Losses in Distribution Systems.” XIII Latin-American Congress on Electricity Generation and Transmission (CLAGTEE), Santiago - Chile, 2019.
- NASCIMENTO, P. H. M. et al. “Impact of Distributed Generation Penetration on Distribution Network Technical Losses.” 16th International Conference on The European Energy Market (EEM), Ljubljana - Eslovênia, 2019.

1.6 ORGANIZAÇÃO

O trabalho é organizado em cinco capítulos, onde além deste capítulo, um detalhamento dos algoritmos de aprendizado de máquina é apresentada no capítulo 2, além das técnicas e métricas de avaliação dos resultados do treinamento destes modelos.

O capítulo 3 descreve a metodologia proposta, desde o tratamento dos dados, passando por etapas de ajuste de informações, treinamento e teste dos modelos e análise de desempenho dos classificadores.

Os resultados e discussões são apresentadas no capítulo 4, com os detalhes da correlação dos dados, análise de sensibilidade dos modelos de Floresta Aleatória e Perceptron Multicamadas, combinação dos modelos, influência do desbalanceamento de informações no banco de dados e finalizando com a análise dos resultados.

No Capítulo 5, são expostas as conclusões deste trabalho e também as possibilidades de desenvolvimentos futuros.

2 APRENDIZADO DE MÁQUINA

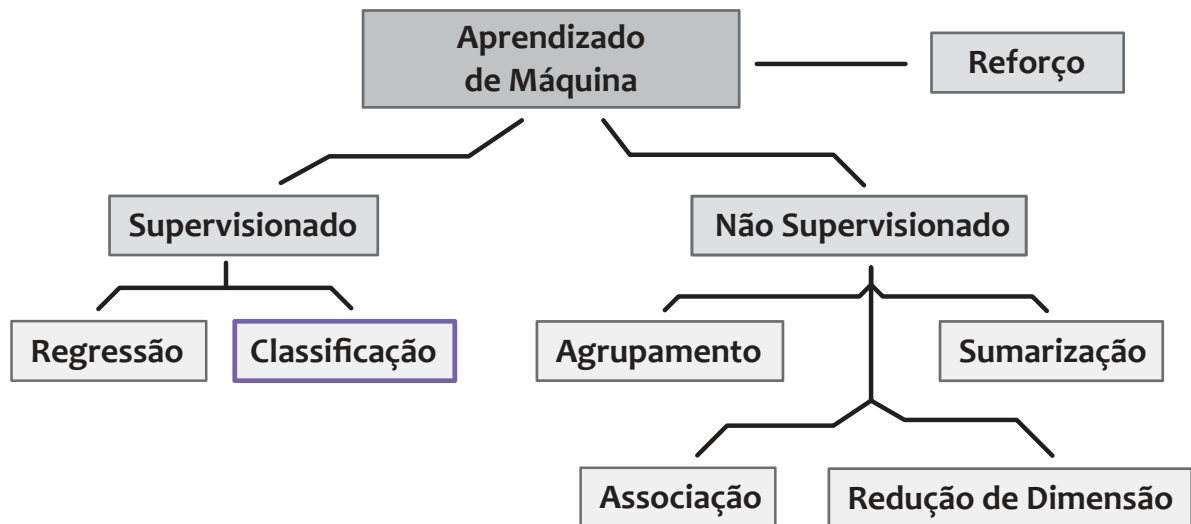
Neste capítulo, serão apresentados o conceito e os tipos de aprendizado de máquina, incluindo dois modelos (algoritmos) de aprendizado e tipos de agrupamento de modelos, que são utilizados neste trabalho. Também se apresenta formas de avaliação do desempenho destes modelos.

O Aprendizado de Máquina (*Machine Learning* - ML), pode ser descrito como um segmento da inteligência artificial, que aplica métodos computacionais para desenvolver estruturas capazes de promover relações implícitas entre dados relacionados a um problema e o resultado/solução deste problema (MOHRI; ROSTAMIZADEH; TALWALKAR, 2018; BENESTY *et al.*, 2009).

2.1 TIPOS DE APRENDIZADO

O ML pode ser dividido em três tipos de aprendizado: supervisionado, aprendizado não-supervisionado e aprendizado por reforço (MARSLAND, 2015). A Figura 2 apresenta os tipos de aprendizados e algumas de suas técnicas, com destaque para a técnica utilizada neste trabalho: classificação.

Figura 2 – Tipos de aprendizado de máquina, com destaque ao método utilizado neste trabalho.



Fonte: Elaborado pelo autor (2021).

No aprendizado supervisionado, o algoritmo tem o objetivo de associar atributos dependentes a atributos não dependente de um problema, sendo que os atributos dependentes são conhecidos. Assim, durante o processo de treinamento do algoritmo de aprendizado, é possível avaliar comparativamente a resposta do algoritmo com o resultado real. As

técnicas de classificação e regressão são exemplos deste tipo de aprendizado (SATHYA; ABRAHAM, 2013).

O aprendizado não-supervisionado trata de algoritmos que buscam encontrar similaridades em atributos não dependentes para poder inferir em um atributo dependente não pre-estabelecido. Normalmente, este tipo de aprendizagem precisa de uma interpretação dos padrões encontrados pelo algoritmo. Incluem-se neste trabalho técnicas de aprendizado não-supervisionado a agrupamento (*Clustering*), associação, redução de dimensão e sumarização (CELEBI; AYDIN, 2016).

O aprendizado por reforço apresenta característica dos dos aprendizados anteriores, resultando em algoritmos capazes de se adaptar a mudanças no meio em que está envolvido. Assim, este algoritmo é capaz de fazer inferências baseando-se apenas em seus atributos independentes, contudo, pode receber recompensas ou penalizações dependendo da sua resposta (SUTTON; BARTO, 2018).

2.2 MODELOS DE APRENDIZADO SUPERVISIONADO

Nesta seção, serão apresentados os modelos de aprendizado supervisionado utilizados neste trabalho. São modelos classificadores: Floresta Aleatória (*Random Forest* - RF), e Perceptron Multicamadas (*Perceptron Multilayer* - MLP).

2.2.1 Floresta Aleatória

O RF é um método de conjunto, que utiliza um grande número de árvores de decisão na busca pelo objetivo, classificação ou regressão. Como o RF é baseado em árvores de decisão, ele acaba absorvendo suas características. Assim, serão abordadas as características das árvores de decisão e em sequência as características de RF.

Árvores de decisão consiste em uma metodologia baseada em hierarquia e na importância de seus ramos. Estes modelos conseguem utilizar como atributos tanto elementos numéricos como categóricos. Além disso, têm a capacidade de seleção de atributos, o que lhe confere robustez quanto a variáveis sem importância e/ou com ruídos (AWAD; KHANNA, 2015).

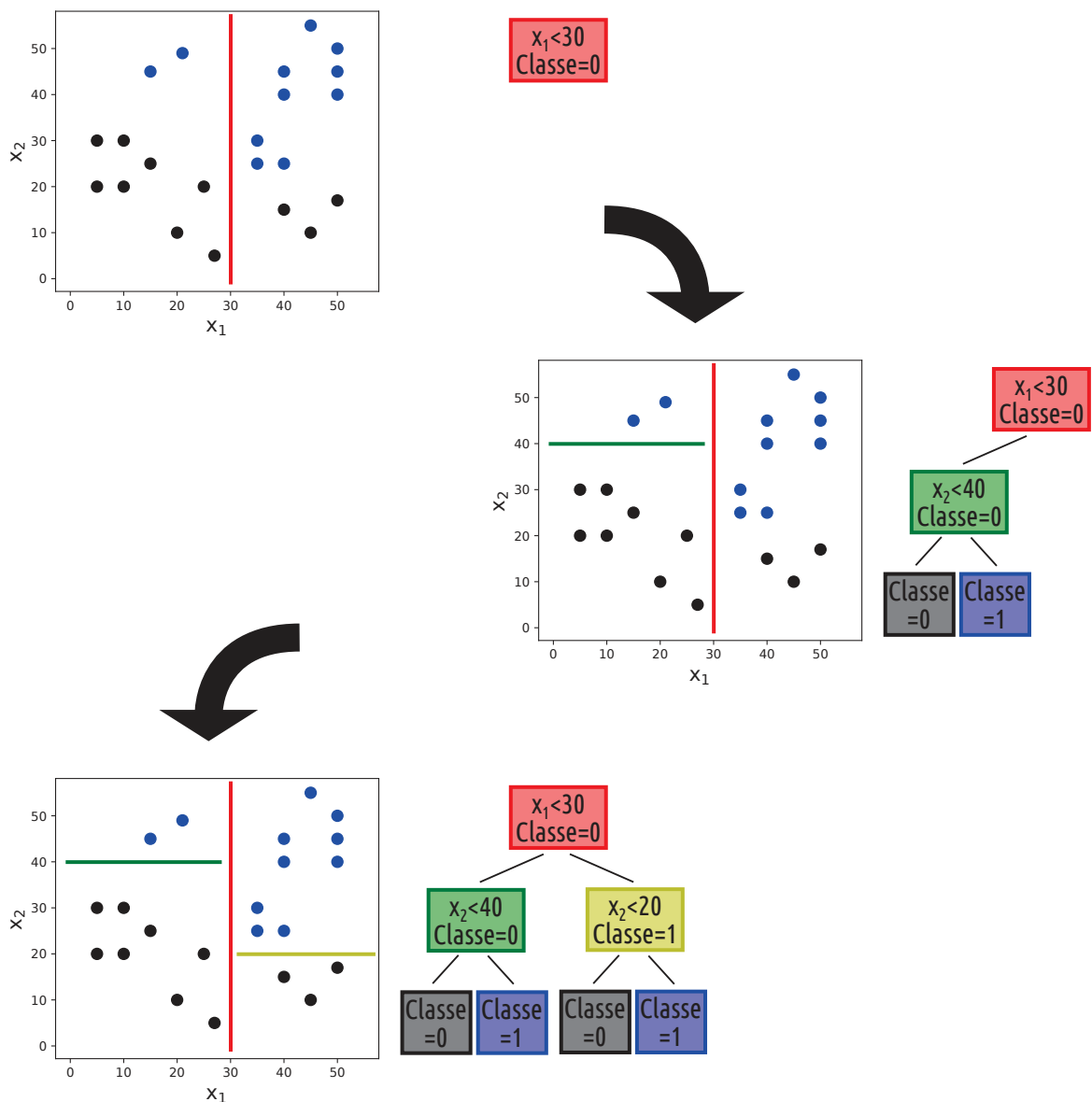
As árvores apresentam alguns elementos estruturais muito importantes, sendo eles: os nós, que são elementos responsáveis por armazenar informação; o nó raiz, que é o primeiro nó da árvore; os nós folha, também chamados de nó terminal, sendo eles os últimos nós da árvore contendo uma informação de classe ou valor (BREIMAN, 2001).

Em cada nó de uma árvore, existe uma pergunta que deve ser avaliada como verdadeira ou falsa, e cada resposta leva a um próximo nó, que pode conter outra pergunta, ou ser um nó folha e representar a decisão final da árvore. Cada nó, com exceção do nó folha, representa um corte ortogonal no espaços, assim, no decorrer de uma árvore, vários

nós são traçados, criando regiões menores, onde as informações ali contidas, tendem a ser as mesmas, também conhecidas como pureza da região.

Para tornar o entendimento mais fácil, a Figura 3 apresenta um gráfico em duas dimensões, em que os pontos de cor preta representam a classe 0 e as marcações em azul representam a classe 1. Os retângulos associados aos nós da árvore estão na mesma cor dos seus respectivos cortes no gráfico. Assim, pode-se observar que as divisões (*split*) ocorrem até que o nó folha seja um nó puro, ou seja, contenha apenas amostras de uma única classe.

Figura 3 – Exemplo simples da estrutura de uma árvore decisão binária, que realiza a classificação da classe 0 e da classe 1, representadas respectivamente pelas cores preta e azul.



Fonte: Elaborado pelo autor (2021).

A pureza dos nós são avaliadas a cada divisão, e quanto mais puro o nó, melhor será a predição. É comum, em problemas de classificação realizar as avaliações das divisões com o Índice de Gini, entropia cruzada ou erro de classificação incorreta (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

A estrutura de treinamento de uma árvore de decisão é permite que ocorram divisões dos nós até que sejam encontrados outros nós onde todos os pontos pertencendo a uma única classe, ou seja, com a máxima pureza no nó, ou até que cada nó folha apresente apenas um único ponto, o que pode levar ao *overfitting* do modelo. Em alguns casos, porém, árvores com pouca profundidade, ou poucas divisões, podem deixar de absorver as regras intrínsecas ao problema. Assim, existem diferentes formas de limitar o crescimento de uma árvore, que podem ocorrer durante o crescimento da árvore ou após o crescimento desta.

As limitações impostas durante o crescimento da árvore podem ser quanto à sua profundidade, grau de pureza da divisão, mínimo de pontos em cada nó folha e também o mínimo de pontos por ramo, para que seja considerada uma divisão. A profundidade limita o número de divisões que um caminho pode conter; por exemplo, quando a árvore é limitada com profundidade 6, isso diz que para quaisquer caminhos percorridos na árvore, o máximo de divisões que poderá ser encontrado antes do nó folha é 6. A limitação baseada no grau de pureza realiza uma simples comparação do grau de pureza do nó com o grau de pureza limite, definido antes do treinamento, e caso esse nó extrapole este limite de pureza, ele é transformado em um nó folha. Algo similar ocorre para a limitação por pontos no nó folha, caso o nó apresente menos amostras/pontos que o limite mínimo e então, este nó é transformado em nó folha. Por fim, a limitação relacionada a amostras por ramo simplesmente impede que uma divisão ocorra caso algum dos nós provenientes desta divisão não apresente o mínimo de amostras previamente definido e também transforma este nó em um nó terminal. É muito usual que em uma mesma configuração de árvore de decisão, sejam levados em consideração mais de um limitador de crescimento (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

As limitações impostas após o crescimento da árvore são conhecidas como poda. Assim, dada a árvore treinada, ocorrem as podas dos galhos mais fracos, que remove de forma contínua os nós, resultando em uma árvore com menos caminhos. A poda é controlada por um parâmetro que é um fator de calibração, sendo assim, caso o parâmetro proporcione uma poda muito intensa, a árvore pode ficar com poucos caminhos e muito rasa, e caso contrário, pode ter pouca ou nenhuma influência na melhoria do desempenho da árvore (JUNIOR, 2018).

A natureza hierárquica das árvores de decisão as tornam muito vulneráveis às divisões de seu nó, visto que um erro cometido em uma divisão próxima ao nó raiz irá ser propagado ao longo de toda a sua extensão. Assim, a árvore pode apresentar grande

variedade de resultados dado a pequenas variações nos dados. Objetivando reduzir este problema de variabilidade de solução, é realizado o emprego de RF (AWAD; KHANNA, 2015).

No RF por meio das técnicas de *Bootstrap* e *Bagging*, realiza-se o treinamento de inúmeras árvores de decisão com banco de treinamentos independentes. Essa estratégia parte do princípio de que um indivíduo (árvore) sozinha é fraca, porém, um grande número de indivíduos fracos, juntos, se tornam fortes (AWAD; KHANNA, 2015). Esse conceito objetiva a mitigação do problema de variabilidade das respostas com árvores de decisão, melhora no desempenho com banco de dados de grande porte, e também diminui impactos de *outliers* nos treinamentos (BELGIU; DRĂGUT, 2016).

O *Bootstrap* é uma técnica de particionamento de banco de dados, de forma a ser possível avaliar a incerteza dos resultados dos treinamentos, com o objetivo de melhorar as predições do modelo. Esta técnica utiliza de escolhas aleatórias de dados para a formação de um novo conjunto, e isso traz a possibilidade de que um mesmo conjunto de atributos, independentes ou dependentes, sejam selecionados mais de uma vez, além de outras informações que sequer façam parte de algum subconjunto utilizado no treinamento. Outra característica é a formação de subconjuntos com a mesma quantidade de informação do conjunto base (inicial) ou de uma fração deste conjunto. A Figura 4 apresenta um exemplo de *Bootstrap* para um conjunto base com seis informações, compostas por três atributos independentes e a sua respectiva classe. Nesta ilustração, é possível identificar que os subconjuntos são formados por um número de informações igual à metade do número de informações do conjunto inicial. Também existe a repetição de informação nos subconjuntos 2 e 4, assim como a ausência da informação 6 de todos os subconjuntos criados (LIAW; WIENER *et al.*, 2002; EFRON; TIBSHIRANI, 1994).

O *Bagging* (ou *Bootstrap Aggregating*) é uma técnica responsável por criar sucessivos preditores treinados de forma independente, por meio de bancos de treinamentos oriundos do *Bootstrap*. Assim, de posse de diversos modelos treinados, a resposta final do modelo preditor é a resposta obtida com a maioria das árvores separadamente, para o caso de um modelo preditor de classificação. Em casos de regressão, a resposta única do conjunto de preditores é encontrada como a média das respostas de todas as árvores individualmente (BREIMAN, 1996a).

A utilização da técnica de *Bootstrap* e *Bagging* para a criação das árvores dá origem à possibilidade de utilização de uma forma de validação dos preditores, ainda durante sua criação e treinamento. Como os dados de formação dos subconjuntos são escolhidos aleatoriamente, isso ocorre com uma probabilidade de $\frac{1}{3}$ dos dados totais não estarem compondo os subconjuntos de treinamentos dos preditores, sendo, então, considerados dados fora do saco (*Out-Of-Bag* - OOB). Assim, o OOB pode ser utilizado como critério de parada para o modelo. O OOB somente é aplicado nos preditores que não possuem

Figura 4 – Imagem da técnica *Bootstrap* aplicada para a criação de seis subconjuntos compostos por um número de dados igual à metade do total de dados que formam o conjunto inicial.

Info.	X	Y	Z	Classe
1	0,7	0,9	0,8	1
2	0,5	0,6	0,1	0
3	0,8	0,2	0,8	0
4	0,9	0,4	0,7	0
5	0,7	0,5	0,1	1
6	0,2	0,2	0,9	1

Banco de Dados

Info.	X	Y	Z	Classe
1	0,7	0,9	0,8	1
2	0,5	0,6	0,1	0
3	0,8	0,2	0,8	0

Subconjunto 1

Info.	X	Y	Z	Classe
5	0,7	0,5	0,1	1
2	0,5	0,6	0,1	0
5	0,7	0,5	0,1	1

Subconjunto 2

Info.	X	Y	Z	Classe
1	0,7	0,9	0,8	1
3	0,8	0,2	0,8	0
5	0,7	0,5	0,1	1

Subconjunto 3

Info.	X	Y	Z	Classe
1	0,7	0,9	0,8	1
1	0,7	0,9	0,8	1
4	0,9	0,4	0,7	0

Subconjunto 4

Info.	X	Y	Z	Classe
2	0,5	0,6	0,1	0
4	0,9	0,4	0,7	0
5	0,7	0,5	0,1	1

Subconjunto 5

Info.	X	Y	Z	Classe
4	0,9	0,4	0,7	0
3	0,8	0,2	0,8	0
1	0,7	0,9	0,8	1

Subconjunto 6

Fonte: Elaborado pelo autor (2021).

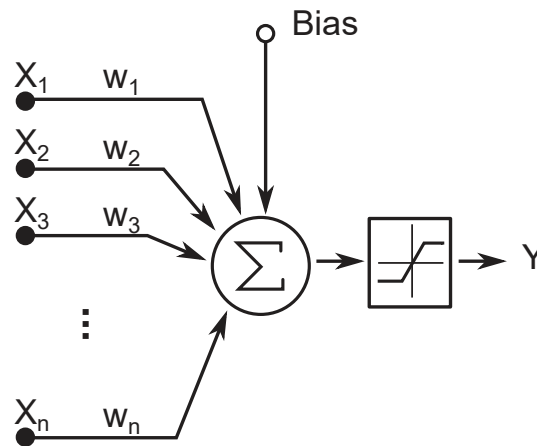
dados de treinamento que pertencem ao OOB (BREIMAN, 1996b).

2.2.2 Perceptron Multicamadas

O MLP é uma categoria dentro de ANN. O perceptron de camada única, ou apenas perceptron, é a estrutura mais básica de uma rede neural, capaz de resolver apenas problemas linearmente separáveis. A estrutura de um perceptron, apresentado na Figura 5, é composta pelas entradas (X_1 a X_n), os pesos sinápticos (w_1 a w_n), o bias, uma função somatório, uma função de ativação e a saída (Y) (HAYKIN, 1994).

O funcionamento do perceptron é relativamente simples de compreender. Os valores de entrada são multiplicados pelos seus respectivos pesos sinápticos. Em sequência, ocorre o somatório dos valores de entrada multiplicados pelos pesos e do bias, gerando o que pode ser chamado de potencial de ativação, que então é submetido a uma função, que vai gerar a resposta ou saída do perceptron. A Figura 6 apresenta a representação de um perceptron em camadas de entrada e saída. Na camada de saída, estão contidos o somatório e a função de ativação, enquanto na camada de entrada não é realizado nenhum tipo de processamento. Assim, devido ao fato de uma camada só ser contabilizada se

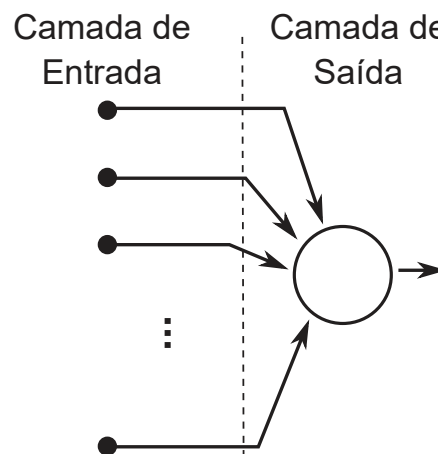
Figura 5 – Estrutura de um perceptron.



Fonte: Elaborado pelo autor (2021).

ela possuir algum tipo de processamento, o perceptron também pode ser chamado de perceptron de camada única ou neurônio artificial (SILVA, 2019a).

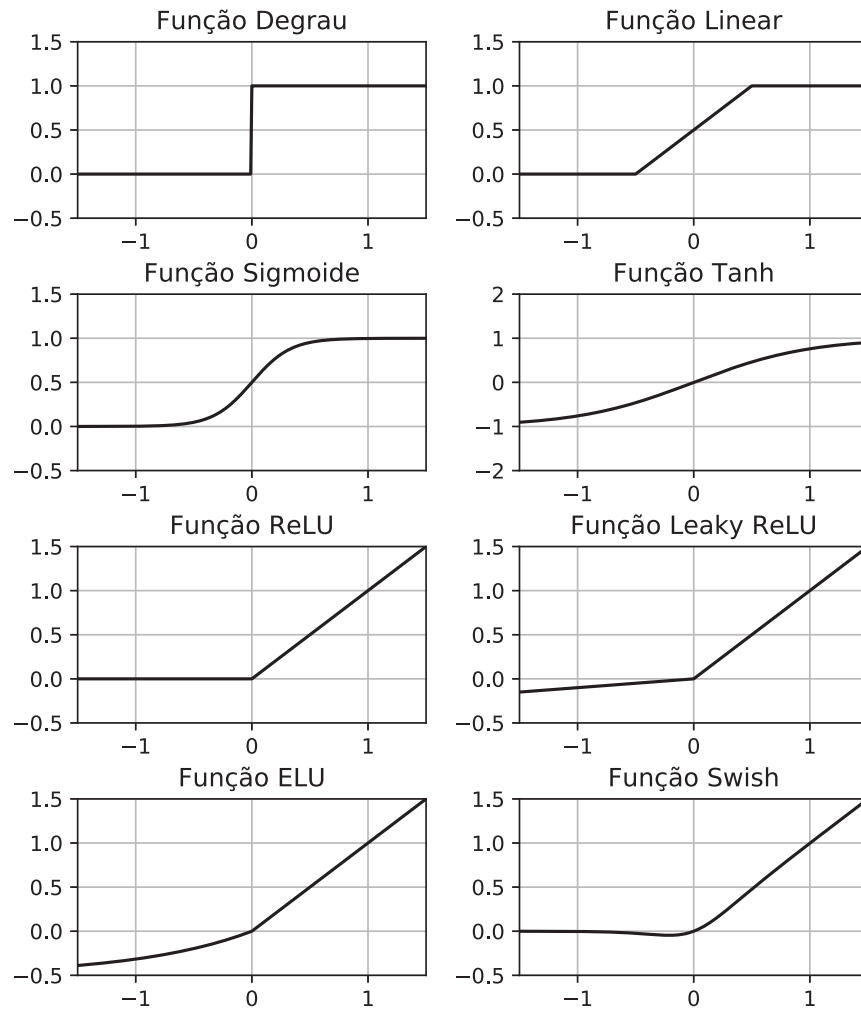
Figura 6 – Estrutura de camadas de um perceptron.



Fonte: Elaborado pelo autor (2021).

As funções de ativação presentes aos neurônios artificiais desempenham um papel muito importante de processar o sinal que chega até eles (SHARMA, 2017). Essas funções comumente apresentam saturação fora de uma faixa de valores, ou seja, apresentam saída limitada, sendo alguns exemplos destas funções: função degrau, linear, sigmoide e tangente hiperbólica (SILVA, 2019a). Contudo, a necessidade de melhores desempenhos de treinamento levou ao desenvolvimento de funções de ativação que podem não apresentar saturação, como a função Ativação Linear Retificada (*Rectified Linear Unit function* - ReLU), a função *Leaky* ReLU, a função Unidade Linear Exponencial (*Exponential Linear Unit* - ELU) e a função *Swish* (AGOSTINELLI *et al.*, 2015; CLEVERT; UNTERTHINER; HOCHREITER, 2016; MAAS; HANNUN; NG, 2013; RAMACHANDRAN; ZOPH; LE, 2017). A

Figura 7 – Ilustrações de funções de ativação.



Fonte: Elaborado pelo autor (2021).

O MLP generaliza o conceito de perceptron, aumentando o número de camadas e também o número de neurônios artificiais em sua composição. Isto possibilita a resolução de problemas com vários níveis de complexidade. Os variados tipos de função de ativação, presentes em cada neurônio artificial da rede possibilitam a utilização desta rede para diversos tipos de problemas, como classificação, regressão e outros (HAYKIN, 1994). O MLP apresenta como característica estrutural a densa ligação entre suas camadas, ou seja, cada neurônio de uma camada está conectado a todos os outros neurônios das camadas adjacentes.

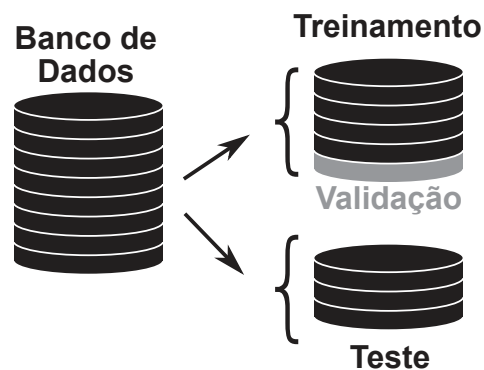
Com a possibilidade de expansão do modelo MLP, há a necessidade metodologias capazes de realizar o treinamento dos parâmetros dessa rede com eficiência e rapidez. O algoritmo *backpropagation*, desenvolvido pra realizar a calibragem dos parâmetros da rede, se mostrou muito eficiente e foi fundamental para o crescimento destes modelos. Este algoritmo consiste de duas etapas: passo para frente (*forward pass*) e o passo para trás

(*backward pass*). A etapa *forward pass* acontece com os dados de entrada passando por toda a rede, bem como por operações e funções de ativação, até que um valor seja entregue na saída. Esse valor encontrado pela rede é comparado com o valor esperado, o que acarreta em um erro. Então a etapa *backward pass* trata de retornar com o erro por toda a rede, fazendo ajustes em seus parâmetros de modo a melhorar o resultado em cada um conjunto de passo para frente e passo para trás. Este algoritmo também é conhecido como algoritmo de retro-propagação.

Os otimizadores são métodos capazes de modificar os parâmetros internos de um modelo de aprendizado de máquina. São algoritmos que buscam melhorar os resultados dos modelos minimizando funções de cálculo de erro. Existem diversos otimizadores disponíveis para otimizar classificadores MLP, são alguns deles: SGD, Adam, Adagrad, Adamax, Adadelta e RMSProp. Existem estudos direcionados a encontrar o melhor otimizador (BERA; SHRIVASTAVA, 2020; SHAF *et al.*, 2018).

Assim como a técnica OOB do RF, o MLP também pode passar por uma validação durante o processo de treinamento do modelo. O banco de dados original passa pela divisão em grupos de dados de treinamento e teste. Após isso, o grupo de teste passa por mais uma divisão, onde uma parcela de dados que seriam utilizado para treinamento serão utilizado para a validação. É importante frisar que o modelo não tem contato com o banco de teste em nenhum momento do treinamento.

Figura 8 – Ilustração de uma divisão do banco de dados em grupos de treinamento, validação e teste.



Fonte: Elaborado pelo autor (2021).

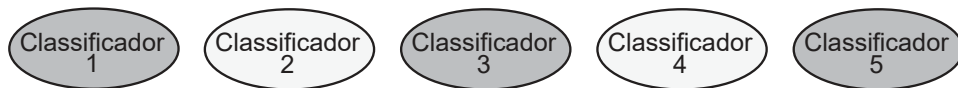
2.2.3 Combinação de Classificadores

A combinação de classificadores parte da ideia de combinação de conhecimentos e opiniões especializadas. As combinações comumente têm duas decisões baseadas em lógica e estatística (TULYAKOV *et al.*, 2008). O modelo de aprendizagem RF, apresentado na Seção 2.2.1, faz uso de técnicas de combinação de classificadores, o *Bootstrap* e *bagging*.

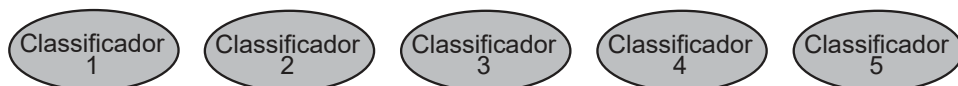
Duas abordagens de combinação de classificação, baseadas em lógica, serão abordadas neste trabalho, sendo elas: Maioria Simples e Unanimidade. A Figura 9 apresenta um exemplo de cada uma destas combinações. A combinação com decisão por maioria simples, em uma classificação binária, define como resultado a classe que é dada como saída em pelo menos 50% + 1 dos classificadores. E a combinação com decisão por unanimidade, como já se infere de forma intuitiva, apresenta uma classe como resultado se e somente se todos os classificadores entregarem a mesma classe como resposta a uma determinada entrada.

Figura 9 – Representação da combinação de classificadores por maioria simples e por unanimidade.

MAIORIA SIMPLES



UNANIMIDADE



Fonte: Elaborado pelo autor (2021).

2.3 AVALIAÇÃO

A criação e treinamento do modelo precisa ser seguida de uma boa metodologia de avaliação. Uma boa métrica de avaliação é fundamental para a construção e definição de um modelo de aprendizado de máquina. Existem diversas métricas de avaliação de modelos preditivos (FÜRNKRANZ; FLACH, 2003). Este trabalho concentra as avaliações dos modelos de classificação em métricas que são bem explicadas por meio da definição da matriz de confusão, sendo elas: acurácia, precisão, revocação ¹ e pontuação F1.

A Figura 10 apresenta a composição da matriz de confusão. O Verdadeiro Negativo (*True Negative* - TN), indica o total de amostras que deveriam ser preditas como negativo e realmente foram preditas assim. O Verdadeiro Positivo, (*True Positive* - TP), representa as amostras preditas como positivas e que realmente pertencem a esta classe. O Falso Positivo (*False Positive* - FP), indica o número de vezes que o modelo deveria prever a classe como negativa, porém, o modelo prediz como sendo da classe positiva. O Falso Negativo (*False Negative* - FN), mostra o número de vezes que a classe negativa foi predita enquanto o correto seria prever a classe positiva.

¹ Revocação é a fração de informações verdadeiras recuperadas, também conhecida como sensibilidade

Figura 10 – Matriz de confusão.

		Valores Preditos	
		Negativo	Positivo
Valores Reais	Negativo	TN (Verdadeiro Negativo)	FP (Falso Positivo)
	Positivo	FN (Falso Negativo)	TP (Verdadeiro Positivo)

Fonte: Elaborado pelo autor (2021).

A acurácia é uma relação entre as previsões corretas e o total de previsões realizadas. Em alguns casos onde os dados são balanceados, a acurácia pode ser suficiente para qualificar um modelo, ou seja, quando as observações positivas apresentam o mesmo número de observações negativas ou números próximos. A acurácia é apresentada na Equação (2.1).

$$Acur = \frac{TP + TN}{TP + TN + FN + FP} \quad (2.1)$$

A precisão está relacionada à razão entre todas as observações positivas previstas corretamente e todas as previsões positivas, ou seja, a taxa de precisão elevada indica a reduzida taxa de falso positivos. A Equação (2.2) apresenta o seu equacionamento.

$$Prec = \frac{TP}{TP + FP} \quad (2.2)$$

A taxa de revocação ou taxa de sensibilidade é definida pela relação entre as observações positivas previstas corretamente e a soma de todas as observações reais positivas, FN e TP. Assim, é possível saber a taxa de previsões corretas de vertimento com relação a todas as vezes que o modelo deveria prever a necessidade do vertimento. A Equação (2.3) modela o cálculo da taxa de revocação.

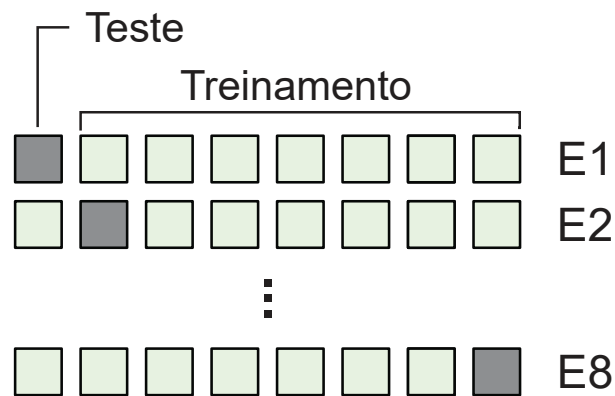
$$Revoc = \frac{TP}{TP + FN} \quad (2.3)$$

A Pontuação F1 é mais utilizada em avaliações de treinamentos com bancos de dados desbalanceados, pois é calculada pela média ponderada da precisão e revocação, dessa forma, levando em conta tanto os falso positivos quanto os falso negativos encontrados nas previsões. A Pontuação F1 é apresentada na Equação (2.4).

$$PontF1 = \frac{2 \times Revoc \times Prec}{Revoc + Prec} \quad (2.4)$$

Além da análise dos modelos realizada com as métricas de avaliação, é muito usual realizar a validação cruzada (*Cross-Validation - CV*), em treinamentos de modelos de aprendizado de máquina (BERRAR, 2019). A CV objetiva avaliar a capacidade de generalização dos modelos e consiste no particionamento do banco de dados em porções menores chamadas de *fold*, como apresentado na Figura 11. Desta forma, um *fold* é utilizado para o teste e o restante é utilizado no treinamento do modelo. Esse procedimento acontece até que todas as possibilidades de conjuntos de treinamento e teste tenham sido utilizadas. Ao final de todas as etapas de teste, é realizado uma média dos valores obtidos para as métricas de avaliação.

Figura 11 – Exemplo de validação cruzada com a divisão do banco de dados em oito subconjuntos.

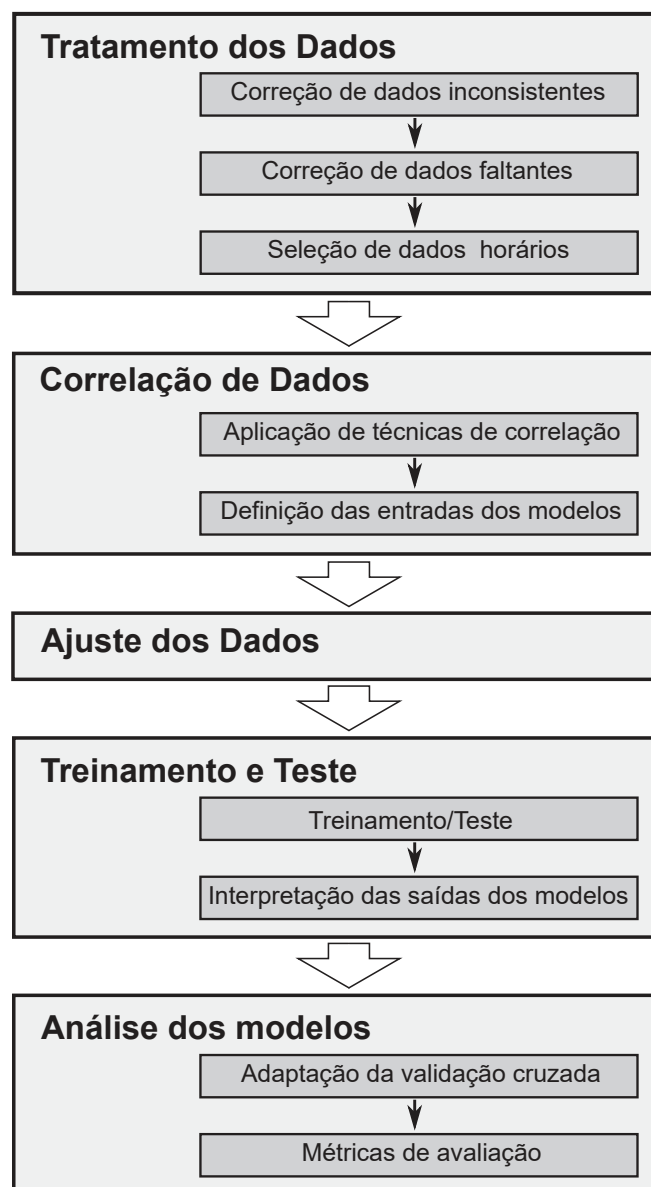


Fonte: Elaborado pelo autor (2021).

3 METODOLOGIA

Neste capítulo é descrita a metodologia proposta para obtenção de um modelo inteligente, capaz de realizar a previsão da condição operativa de vertimento em uma UHE. Essa metodologia pode ser dividida em cinco etapas: (i) tratamento dos dados, (ii) correlação de dados, (iii) ajuste dos dados, (iv) treinamento e teste, e (v) análise dos modelos, como apresentado na Figura 1. As seções deste capítulo trazem, de forma detalhada, cada uma destas etapas.

Diagrama 1 – Metodologia proposta para desenvolvimento e avaliação de modelos de previsão de vertimento utilizando aprendizagem de máquina.



Fonte: Elaborado pelo autor (2021).

3.1 TRATAMENTO DOS DADOS

Os dados utilizados neste trabalho são referentes à medições automatizadas realizada em estações de telemetria sob concessão da usina e também medições realizadas na própria usina. A aquisição destas informações está sujeita a falhas, gerando problemas com os dados, como por exemplo, os valores podem apresentar taxa de amostragem diferente da padrão, valores incoerentes e até mesmo valores faltantes. Essas falhas podem ser provenientes de diferentes fontes, como uma falta de calibração do equipamento, ruídos, falha de comunicação durante a transmissão de informação, parada para manutenção ou até mesmo defeito do equipamento.

Com o intuito de resolver os problemas apresentados, com relação aos dados, é proposto um tratamento das informações baseado nos problemas já conhecidos e que aparecem no banco de dados. Desta maneira, a primeira correção, quanto a dados inconsistentes, busca mitigar informações que apresentem valores negativos, sendo realizada a exclusão destes dados.

A segunda etapa de tratamento das informações, refere-se aos dados faltantes. Os dados podem apresentar discretização horária ou a cada quinze minutos, e isso pode trazer alguns problemas de falha: as informações podem não ser coletadas na hora exata, mas somente nas frações de hora; também podem ocorrer intervalos de horas ou até mesmo dias em que não há nenhum valor mensurado. Como os dados são sequenciais e amostrados em tempos específicos, não é possível simplesmente eliminá-lo. Assim, o algoritmo trata todas as informações como se estas apresentassem discretização a cada quinze minutos, pois é a maior taxa de amostragem, e nas datas que não possuem informação, é aplicado a interpolação linear, para a composição destes dados ausentes, seguindo a Equação 3.1.

$$y = y_0 + (y_1 - y_0) \frac{x - x_0}{x_1 - x_0} \quad (3.1)$$

Onde (x_0, y_0) e (x_1, y_1) dois diferentes pontos de informação.

A interpolação linear aplicada aos dados, sendo considerada uma amostragem de 15 minutos, evita que informações reais sejam perdidas. Contudo, apenas serão utilizadas informações horárias. Logo, a terceira etapa do tratamento consiste na eliminação dos dados em fração de hora, de forma que seja criado um banco de dados com informações de hora exata.

3.2 CORRELAÇÃO DE DADOS

Decorrida a etapa de tratamento dos dados históricos, estes são submetidos a uma análise de correlação linear e não-linear, com o objetivo de selecionar características de maior relevância para o modelo de aprendizado.

O coeficiente de correlação de Pearson (ρ) é o primeiro índice definido para avaliar a correlação entre os dados do problema (BENESTY *et al.*, 2009). Este índice determina a correlação linear entre duas variáveis de escala por meio da Equação 3.2:

$$\rho = \frac{\sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (3.2)$$

Onde, x_i e y_i são os valores medidos das variáveis, \bar{x} e \bar{y} são as médias aritméticas das variáveis e N é o número de dados analisados. Este coeficiente (ρ) assume valores entre -1 e 1, no qual representam:

- $\rho = 1$: correlação perfeita positiva entre duas variáveis;
- $\rho = -1$: correlação perfeita negativa entre duas variáveis, ou seja, caso uma aumente a outra diminui;
- $\rho = 0$: variáveis que não dependem linearmente uma da outra.

Visando uma análise de correlação robusta e devido à técnica de correlação de Pearson investigar apenas correlações lineares entre os dados, a técnica de regressão polinomial é aplicada para verificar se há alguma correlação não-linear nos dados. Os resultados da regressão polinomial são analisados com o coeficiente de determinação (R^2), utilizando a Equação 3.3.

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3.3)$$

Onde,

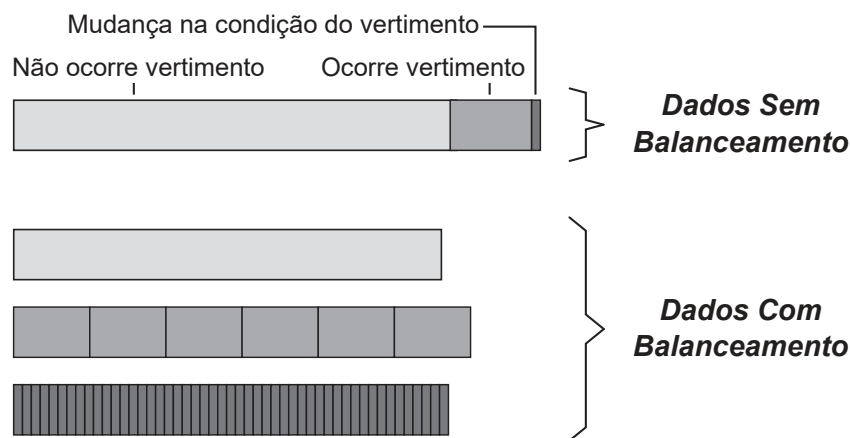
$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

3.3 AJUSTE DOS DADOS

O banco de dados utilizado para o treinamento do modelo de previsão apresenta um desbalanceamento muito grande com relação a vertimento e não vertimento na usina, visto que o vertimento normalmente acontece em épocas mais chuvosa do ano, ou seja, muito dependente das estações do ano. Ao longo de um ano, o número de horas onde não ocorrerá vertimento é significativamente maior que os momento com vertimento. Assim, uma aplicação de um conjunto de dados sem tratamento relacionado ao equilíbrio das classes, pode impactar na etapa de treinamento do modelo.

Desta maneira, foi aplicada uma estratégia de balanceamento do conjunto de dados, de forma que no treinamento, as indicações de vertimento, não vertimento e mudança na condição de vertimento passassem a apresentar quantidade de informações com pouca diferença. Essa estratégia se baseia na simples repetição de dados com menor frequência no banco de dados original, até que este conjunto de dados apresente quantidade de dados igual ou minimamente superior. A Figura 12 ilustra o desequilíbrio observado em um banco de dados histórico de vertimento e o ajuste implementado com o intuito de balancear os dados e melhorar o treinamento do modelo proposto.

Figura 12 – Ilustração do balanceamento de dados.



Fonte: Elaborado pelo autor (2021).

Este tratamento ocorre apenas no banco de dados exclusivamente de treinamento do modelo, ou seja, esse balanceamento ocorre após a divisão do grupo de treinamento e teste. A validade da implementação deste tratamento baseado em balanceamento de dados é analisada de forma comparativa neste trabalho. Somente alguns modelos foram treinados com dados modificados com o procedimento de balanceamento.

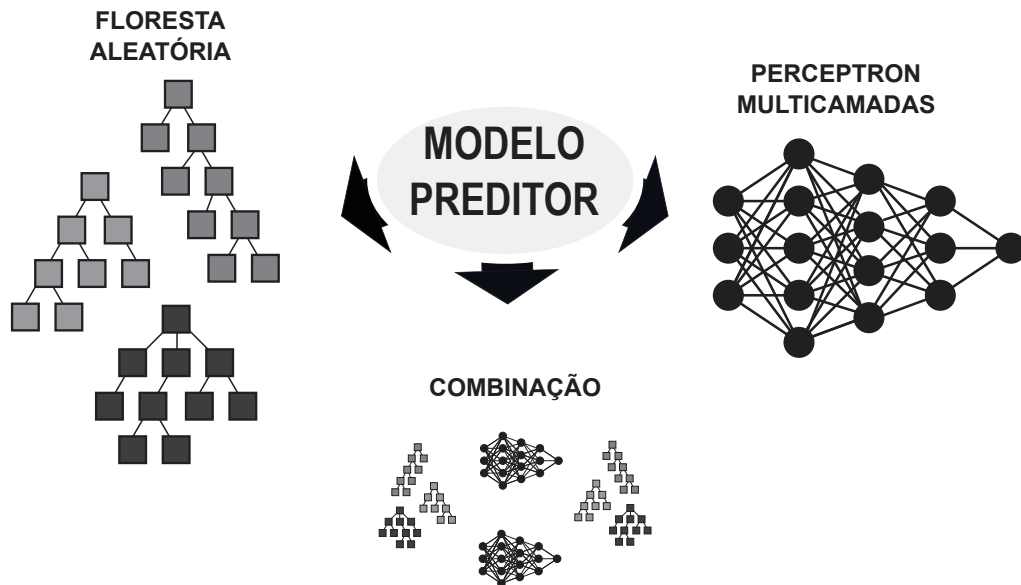
3.4 TREINAMENTO E TESTE DOS MODELOS

Os dados utilizados para treinamento e teste dos modelos de aprendizado de máquina, passam primeiramente por um pré-processamento, onde é aplicado uma normalização dos dados, entre zero e 120% do valor máximo presente no histórico, para cada tipo de conjunto de informações. Ou seja, a normalização é aplicada de forma separada para valores de nível do rio, vertimento e os demais. Assim, é garantido que os dados de entrada tenha uma contribuição numérica igual (SINGH; SINGH, 2020) e o modelo de treinamento se torna mais robusto a possíveis medições com valores superiores aos até então estabelecidos no banco de dados original.

Os modelos de previsão de vertimento são baseados em RF, MLP e também na combinação desses dois anteriores, como ilustrado na Figura 13. As principais motivações

para o uso das metodologias de aprendizado de máquina se devem ao fato de ambas serem utilizadas na literatura especializada como ferramentas de classificação (AWAD; KHANNA, 2015) e apresentarem bons desempenho com dados em forma tabular (YEUNG *et al.*, 2010; CERRADA *et al.*, 2016).

Figura 13 – Ilustração dos modelos preditores.



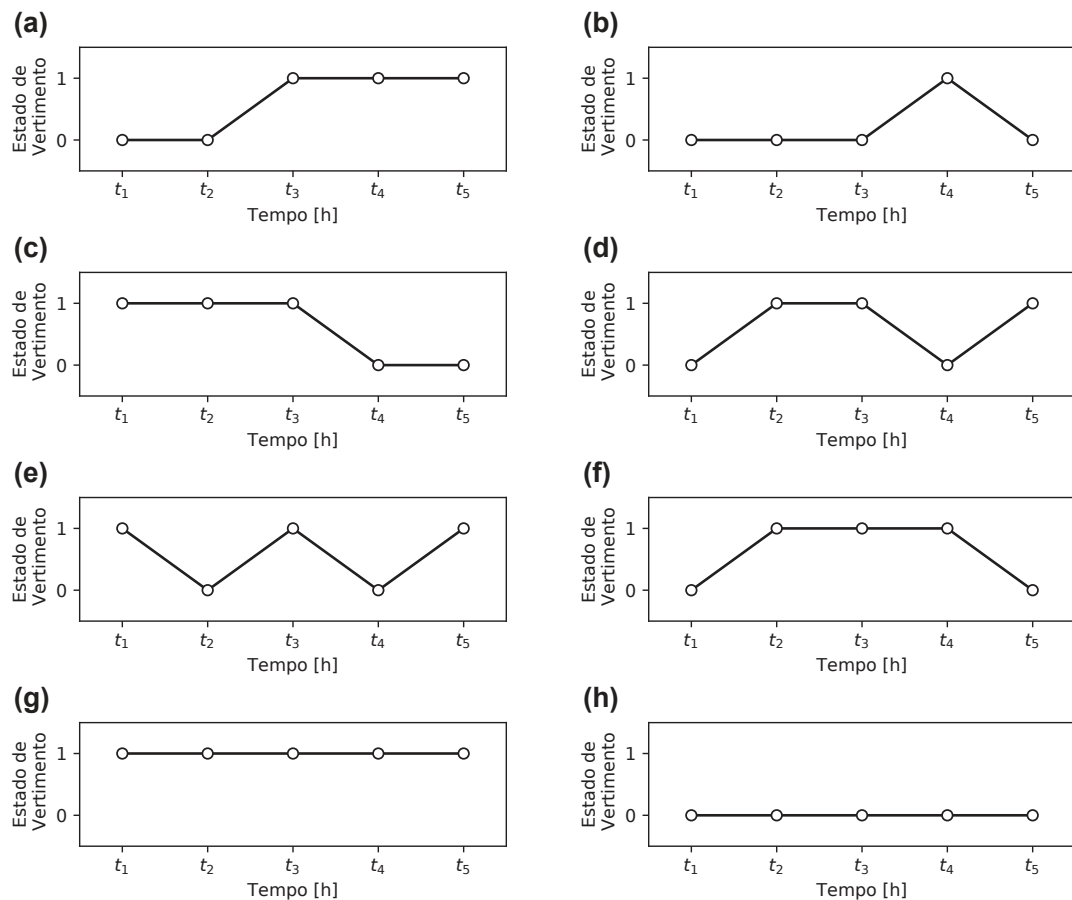
Fonte: Elaborado pelo autor (2021).

A saída do modelo de previsão consiste em cinco valores binários em sequência, que representam a condição da operação de vertimento por cinco horas à frente, onde o valor 0 (falso) indica a ausência de vertimento e um valor 1 (verdadeiro) indica a necessidade de vertimento. Alguns perfis de resposta do modelo são expostos na Figura 14, onde as seis primeiras figuras são momentos de mudança na condição do vertimento e as Figuras 14(g) e 14(h) apresentam configurações comuns de vertimento e ausência de vertimento, respectivamente.

As Figuras 14(b), 14(d), 14(e) e 14(f) representam momentos raramente encontrados no histórico de vertimento, pois essas alterações no estado do vertimento em um período muito curto não são comuns, podem representar até mesmo decisões equivocadas quanto a tomada ou retomada do vertimento. Além disso, essas configurações são difíceis de serem alcançadas pelo modelo de previsão e podem complicar o processo de aprendizagem do modelo de previsão.

Assim, para melhorar o desempenho do treinamento e manter a qualidade da resposta do preditor, as cinco horas a serem previstas são tratadas como um único valor binário, em que a lógica se mantém idêntica, 0 corresponde à ausência de vertimento em todas as horas preditas enquanto o 1 corresponde à necessidade de vertimento em pelo menos uma das horas preditas.

Figura 14 – Exemplo de configuração de dados de vertimento preditos para 5 horas futuras.



Fonte: Elaborado pelo autor (2021).

Ao longo da operação da usina, a tomada de decisão muitas vezes está associada à necessidade de alteração da condição de vertimento. Essas alterações são consideradas como situações de grande importância para a operação, ou seja, a decisão de quando é necessário iniciar ou interromper o vertimento de água do reservatório da usina, é crucial para os operadores. Logo, esses momentos de mudança no vertimento também são analisados neste trabalho. As Figuras 14(a) a 14(f) são exemplos de mudança na condição de vertimento, e elas sempre serão consideradas vertimento, pois para isso, apenas uma das horas preditas deve apresentar o valor 1.

Como a tomada de decisão quanto ao vertimento ou não normalmente é realizada pelo operador da UHE, algumas decisões distintas podem ser tomadas para cenários de condições hidrológicas, climáticas e de operação semelhantes. Assim, a utilização de modelos de previsão de vertimento podem auxiliar na uniformidade dessas decisões.

3.5 ANÁLISE DE DESEMPENHO

A análise de treinamento dos modelos de aprendizado de máquina precisa considerar a generalização do modelo, ou seja, a capacidade de fazer previsões de boa qualidade independentemente dos dados usados durante o treinamento e dos dados utilizados para teste. Assim, é utilizada a técnica de validação cruzada para a divisão do banco de dados em treinamento e teste. Contudo, os dados que acusam vertimento são minoria no banco de dados, que é predominantemente formado com dados que apontam a condição de não vertimento.

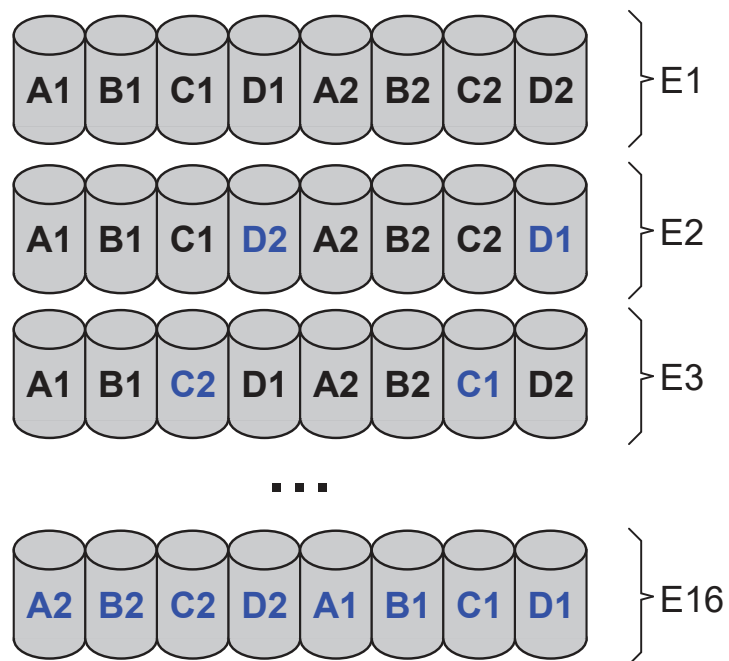
A condição climática está fortemente relacionada com as ocorrências de vertimento. Logo, as informações históricas das quatro estações do ano devem aparecer tanto nos dados de treinamento quanto nos dados de teste dos modelos de previsão.

Com base nestas afirmações, e considerando um banco de dados históricos que apresente as quatro estações de um ano completa, ou seja, um banco de dados que contenha dados anuais completos ou quase completos, de modo que seja possível dividir o banco de dados em $4 \times N$ partes, onde N é o número de anos existente no banco de dados. Desta maneira, ao considerar 4 partes sequenciais está observando informações de quatro estações, mas não necessariamente de um mesmo ano. Assim pode-se trocar partes de estações equivalentes, modificando a composição do banco de treinamento e teste.

A Figura 15 apresenta a divisão de um banco de dados que apresente informações históricas de dois anos. Como são dois anos de dados, tem-se $N = 2$, e a divisão é realizada em oito partes. A consideração sequencial é importante para a montagem dos diferentes grupos de treinamento e teste. Como na imagem, a parte $A1$ apenas pode ser trocada de posição com a parte $A2$, dessa forma, é garantido que estas partes representam a mesma estação do ano ou mesmo período de um ano. Essas trocas são realizadas até que todas as formações diferentes de grupos de treinamento e teste encontradas, sendo que são 16 formações possíveis para um banco de dados com informações de dois anos.

As métricas de avaliação apresentadas na Seção 2.3 são aplicadas para todos os modelos de aprendizagem de máquina e para cada um dos treinamentos realizados com a validação cruzada adaptada. Os resultados são apresentados em formato de tabelas e também gráficos *boxplot*, para que seja possível verificar a influência dos diferentes grupos de treinamento e teste formados seguindo a divisão do banco de dados da Figura 15.

Figura 15 – Ilustração da validação cruzada adaptada para dados de dois anos.



Fonte: Elaborado pelo autor (2021).

4 RESULTADOS E DISCUSSÃO

Neste capítulo, serão apresentadas as análises de correlação utilizadas para a definição das variáveis de entrada dos modelos de previsão, a estrutura do modelo, análises de sensibilidade dos modelos baseados em *Random Forest* e também dos modelos baseados em *Multilayer Perceptron*. Além disso, é realizado o estudo do impacto nos resultados devido a diferentes formas do tratamento do banco de dados de treinamento dos modelos e da utilização de agrupamentos de modelos de previsão.

A metodologia foi toda implementada em linguagem Python. Os treinamentos e teste dos modelos de previsão foram realizados no ambiente Google Colaboratory (Google Colab), que é um serviço de nuvem oferecido gratuitamente e hospedado pelo Google. Os treinamentos do modelo de *Random Forest* se deram com a utilização da biblioteca *Scikit-learn* versão 0.22.2, e com a classe *RandomForestClassifier*. O modelo *Multilayer Perceptron* foi treinado utilizando-se da biblioteca *Tensorflow* versão 2.4.1, com a classe *Sequential*.

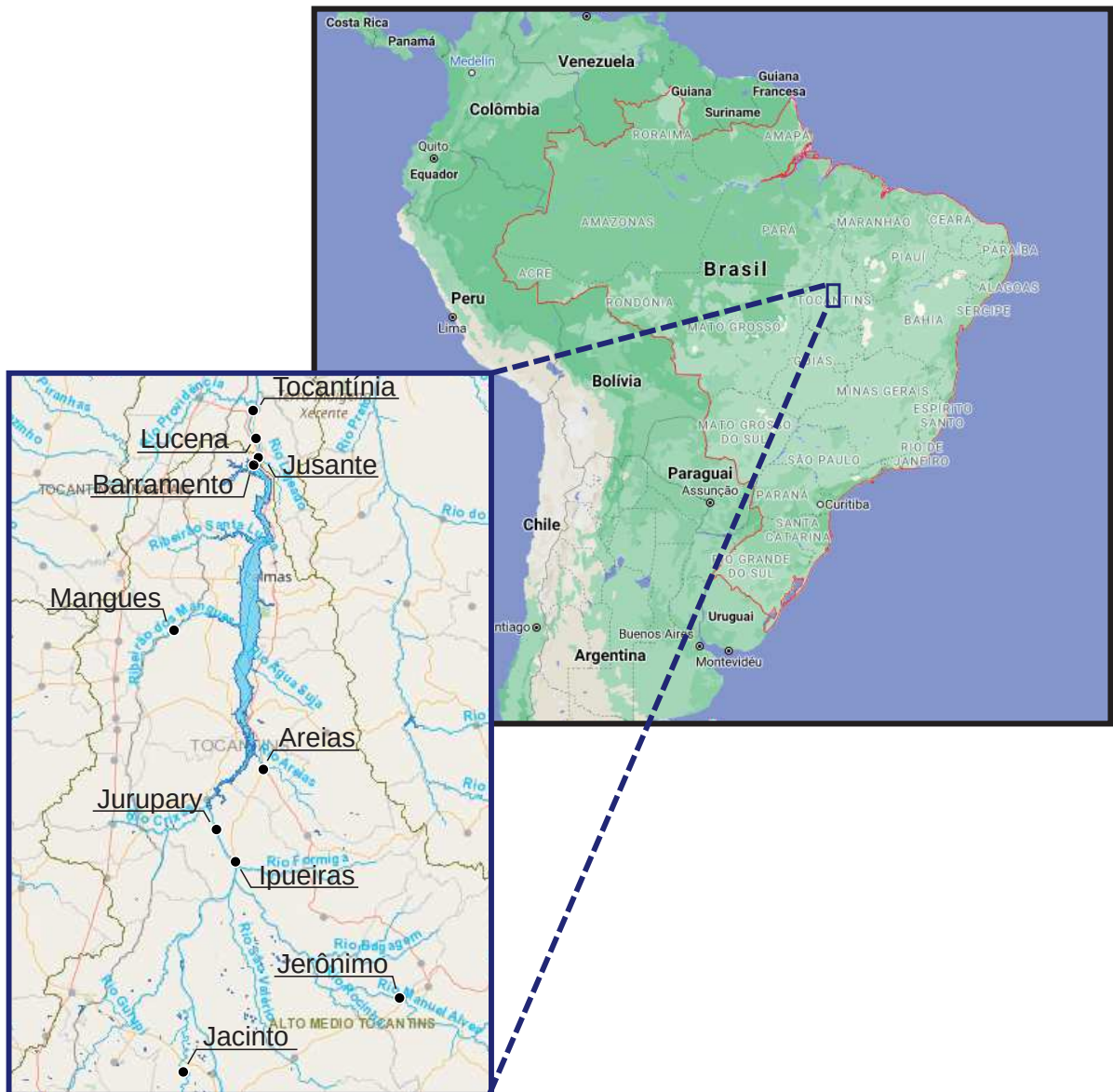
4.1 ANÁLISE DE CORRELAÇÃO

Os banco de dados utilizado neste trabalho é composto por informações compreendidas no período de 13 de agosto de 2018 a 21 de julho de 2020, e foram fornecidas pela concessionária que administra a UHE Lageado, contudo os dados das estações de telemetria também podem ser encontrados no site da ANA (SNIRH, 2020). Os dados de controle hidráulico da usina apresentam discretização horária e as informações de telemetria apresentam discretização de quinze minutos.

Visando à determinação das variáveis a serem utilizadas como entrada do modelo de previsão de vertimento, a primeira análise realizada tratou das posições das estações de telemetria em relação à localização da UHE, como ilustrado na Figura 16. Desta forma, com uma análise simples, as estações de telemetria, Jusante, Lucena e Tocantínia foram excluídas dos estudos por se tratarem de estações localizadas a jusante da barragem da UHE. A estação Barramento não apresenta dados que se diferem dos dados referentes às informações de controle hidráulico da própria UHE, logo, por ser redundante a utilização destes dados, as informações da estação Barramento também foram excluídas do estudo.

Com a simples análise quanto a localização e redundância das estações de telemetria, quatro das dez estações foram retiradas dos estudos, permanecendo as estações Jacinto, Jerônimo, Ipueiras, Jurupary, Areias e Mangues. Os dados de nível do rio, capturados por estas seis estações de telemetria, foram submetidos a uma análise de correlação com os parâmetros de controle hidráulico da UHE, nível montante e o vertimento. O método de Pearson foi utilizado para os cálculos da correlação linear, enquanto que para a análise não-linear foi utilizada a regressão polinomial de segundo, terceiro, quarto e quinto grau.

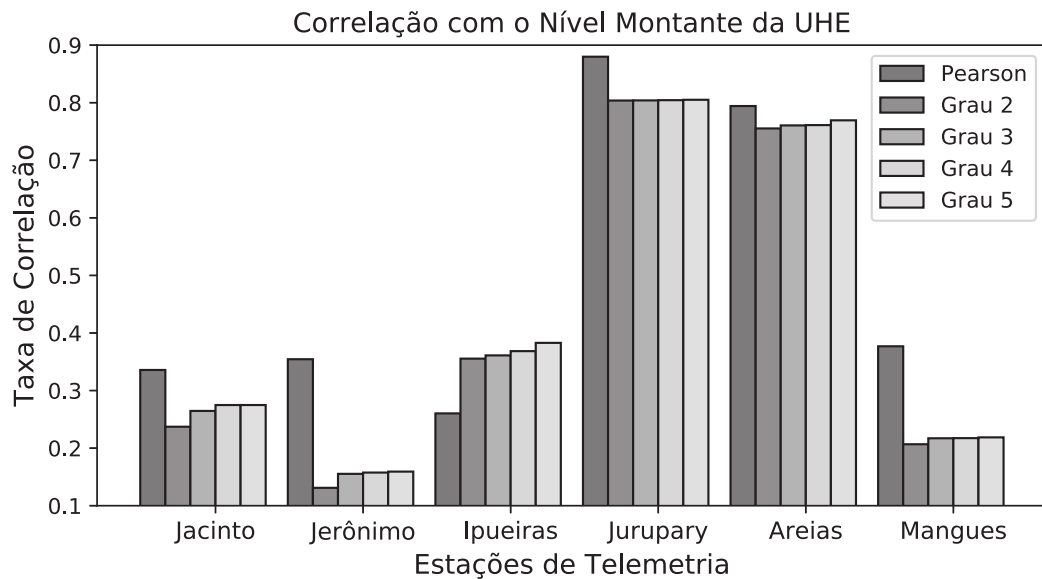
Figura 16 – Localização das estações de telemetria.



Fonte: Elaborado pelo autor (2021).

Os resultados dos cálculos de correlação entre o nível das estações e o nível a montante da represa estão apresentados no Gráfico 1, no qual se observa em destaque os valores de correlação encontrados para as estações de Jurupary e Areias. Jurupary apresentou maior pontuação referente a correlação de Pearson, com o valor igual a 0,879, e o menor valor foi encontrado para a regressão com polinômio de grau 2, cujo valor é 0,804. A estação de Areias obteve correlações um pouco menores que as encontradas com Jurupary, sendo sua melhor pontuação igual a 0,794, também referente ao método de Pearson e sua menor correlação igual a 0,755, também para a regressão polinomial de grau 2. Assim, essas duas estações de telemetria se destacaram frente às demais, que em nenhum momento obtiveram pontuações de correlação superiores a 0,4.

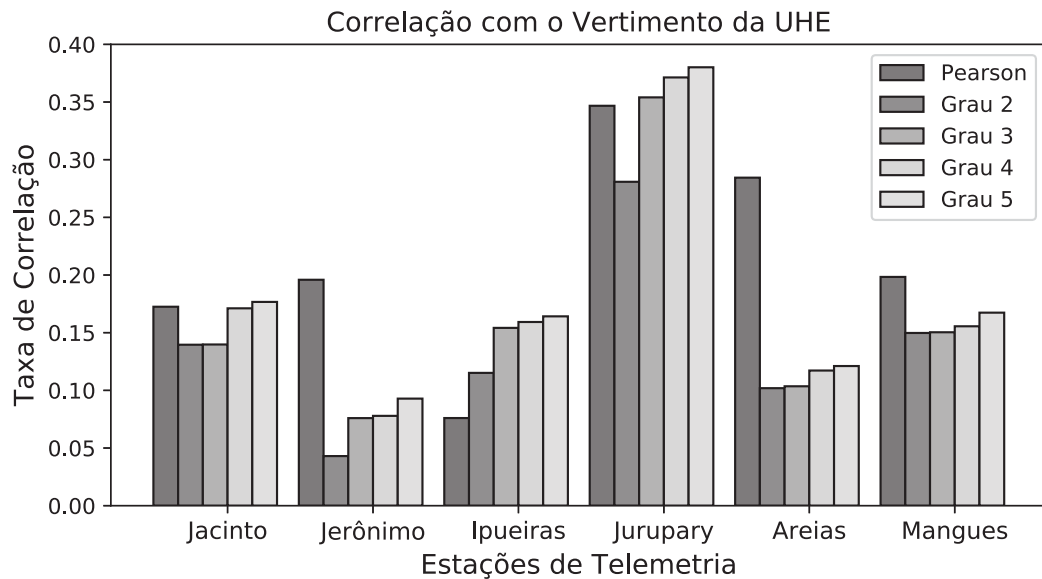
Gráfico 1 – Valores de correlação do nível medidos nas estações de telemetria com o nível a montante da UHE.



Fonte: Elaborado pelo autor (2021).

O Gráfico 2 apresenta os valores calculados de correlação dos níveis medidos nas estações de telemetria com os valores de volume vertido mensurados na UHE, no qual é possível verificar que, de modo geral, as correlações são baixas, sempre apresentando valores inferiores a 0,4. No entanto, assim como na correlação com o nível a montante da UHE, a estação Jurupary se destacou frente às outras, só apresentando valor menor que 0,35 para a regressão de segundo grau. A estação Areias apresenta o segundo maior valor para a correlação de Pearson, 0,284, o que é 43,4% maior que a terceira maior, encontrada em Mangues com valor igual a 0,198. As demais estações de telemetria não apresentaram expressividade em seus valores de correlação.

Gráfico 2 – Valores de correlação do nível medidos nas estações de telemetria com o vertimento da UHE.



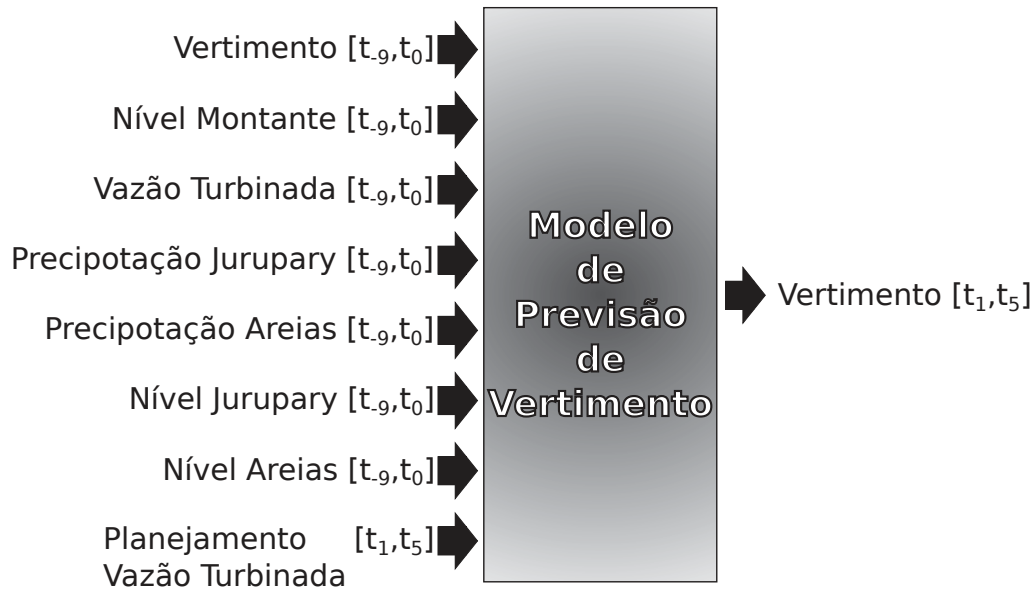
Fonte: Elaborado pelo autor (2021).

4.2 ESTRUTURA DO MODELO DE PREVISÃO DE VERTIMENTO

Com os resultados de correlação entre dados das estações de telemetria e da usina, foi possível identificar que Jurupary e Areias apresentaram melhores correlações com as informações da usina. Assim, a estrutura de entrada do modelo de previsão de vertimento apresenta em sua composição informações de nível medidos nas estações de Jurupary e Areias, e como complementação de informação, dados de precipitação referentes a estas estações também são utilizados como dados de entrada do modelo. As informações de nível a montante, vertimento e vazão turbinada completam as sete características de entrada do modelo, como apresentado na Figura 17.

Todas as sete características de entrada são compostas por informações referentes às dez horas anteriores ao momento inicial da previsão, e tal quantidade de dados foi definida por estudos previamente realizados. Além das informações de horas passadas, o modelo também recebe como entrada cinco horas futuras de programação de vazão turbinada na usina. Esses valores esperados são fornecidos por um modelo de otimização que concede o cronograma de operação com base nas metas de geração atribuídas à UHE. Na saída, o modelo fornece valores binários que se referem a condição operativa com relação ao vertimento nas cinco horas de previsão.

Figura 17 – Estrutura de entradas e saídas do modelo de previsão do vertimento.



Fonte: Elaborado pelo autor (2021).

4.3 SENSIBILIDADE DO MODELO *RANDOM FOREST*

Objetivando entender o efeito da variação de parâmetros internos dos modelos de aprendizagem de máquina utilizados, é realizada uma análise de sensibilidade quanto a alguns desses parâmetros. No modelo baseado em *Random Forest*, são realizadas análises com relação a dois diferentes parâmetros: quanto ao número de árvores (estimadores) que compõem a floresta e quanto ao número máximo de dados utilizados para o treinamento de cada uma das árvores que compõem a floresta. As demais características são as mesmas para todos os treinamentos realizados para este modelo:

- A qualidade da divisão do nó, ou *split*, é medida pela Função Gini (Índice Gini);
- As árvores de decisão não possuem limite de profundidade;
- É definido como 2 o valor mínimo de amostras para que o nó possa ser dividido;
- Uma divisão só será considerada se deixar pelo menos uma amostra de treinamento por ramo;
- O número de atributos utilizados na busca pela melhor divisão é igual à raiz quadrada do total de atributos;
- É utilizada a técnica de *bootstrap* na divisão dos subconjuntos de treinamentos (conjunto de treinamento de cada árvore);
- A funcionalidade *Out-Of-Bag* é utilizada.

A primeira análise de sensibilidade diz respeito ao número de árvores/estimadores que formam a floresta. Foram definidos seis diferentes valores a serem atribuídos ao número de estimadores, são eles: 10, 50, 100, 500, 1000 e 1500 estimadores. Para esta etapa de análise, o número máximo de dados utilizados para o treinamento de cada uma das árvores foi mantido como sendo 10% do número total de informações no banco de treinamento. A Tabela 2 apresenta a média de todas as métricas, calculadas para os cinco treinamentos e testes, realizados para cada configuração de estimadores no modelo. Assim, é possível observar que a variação dos resultados entre diferentes números de árvores é muito pequena. Os melhores resultados referentes a cada métrica foram destacados em vermelho, com exceção do tempo. O modelo treinado com a configuração de 500 estimadores apresentou o melhor valor para três das seis métricas de avaliação; também apresentou o segundo melhor valor quanto a acurácia I, com uma porcentagem de acerto apenas 0,06% menor que o modelo que utilizou apenas 10 estimadores. Contudo, ainda comparando estes dois modelos, o valor médio de acurácia II encontrado pelo modelo com 500 estimadores se difere em aproximadamente 8% do modelo com 10 estimadores. O modelo treinado com 1000 estimadores, nas métricas que se destaca, não é nem 0,5% melhor que o modelo com 500 estimadores, e ainda precisa do dobro do tempo de treinamento, em média. Portanto, o valor de 500 estimadores será utilizado na sequência da análise de sensibilidade.

Tabela 2 – Valores médios com cinco diferentes formações - primeira análise de sensibilidade RF.

Nº de Árvores	Ac. I	Ac. II	Ac. Média	Pont. F1	Precisão	Revocação	Tempo [s]
10	0,9454	0,5092	0,7273	0,8599	0,7809	0,9603	1
50	0,9423	0,5929	0,7676	0,8602	0,7756	0,9735	2
100	0,9421	0,5845	0,7633	0,8601	0,7759	0,9720	3
500	0,9448	0,5891	0,7670	0,8643	0,7814	0,9736	15
1000	0,9442	0,5932	0,7687	0,8637	0,7812	0,9733	30
1500	0,9440	0,5909	0,7674	0,8628	0,7800	0,9729	46

Fonte: Elaborado pelo autor (2021).

Dada a definição do número de árvores que serão utilizadas nos modelos de *Random Forest*, é realizada a segunda análise de sensibilidade, relacionada ao número máximo de informações que formam os subconjuntos utilizados no treinamento de cada árvore. Nesta análise, os subconjuntos podem ser formados com 10%, 30%, 40%, 50%, 70% ou 90% do número máximo de informações presentes no banco de treinamento. É importante ressaltar que na formação dos subconjuntos, utilizando-se da técnica *bootstrap*, pode existir informações repetidas em um mesmo subconjunto.

Na Tabela 3, estão dispostos os resultados médios para as métricas de avaliação, obtidos para a segunda análise de sensibilidade para os modelos de *Random Forest*. A coluna “Amostras *bootstrap*” refere-se à variação do volume dos subconjuntos de treinamentos.

Nesta tabela, os valores com destacados em vermelhos são os melhores por métrica. Os resultados de acurácia I, apresentou uma tendência em melhorar a medida que o subconjunto aumenta de volume de informações. Todavia, com os resultados obtidos para a acurácia II, pode-se observar um comportamento contrário, ou seja, os valores para essa métrica diminuem a medida que o subconjunto de informações cresce.

Tabela 3 – Valores médios com cinco diferentes formações - segunda análise de sensibilidade RF.

Amostras <i>bootstrap</i>	Ac. I	Ac. II	Ac. Média	Pont. F1	Precisão	Revocação	Tempo [s]
10%	0,9405	0,5893	0,7649	0,8573	0,7726	0,9733	15
30%	0,9567	0,5339	0,7453	0,8875	0,8226	0,9661	20
40%	0,9655	0,4985	0,7320	0,9039	0,8537	0,9624	21
50%	0,9665	0,5066	0,7365	0,9064	0,8581	0,9623	23
70%	0,9672	0,4830	0,7251	0,9080	0,8635	0,9590	27
90%	0,9687	0,4729	0,7208	0,9122	0,8725	0,9574	32

Fonte: Elaborado pelo autor (2021).

Apesar de a variação das médias dos resultados das métricas continuar baixa, pode-se afirmar que o modelo *Random Forest* se apresentou mais sensível às variações relativas aos subconjuntos de treinamento das árvores do que à variação do número de árvores presente em cada floresta. O comportamento do tempo de treinamento se comporta como o esperado, visto que o aumento nos dados utilizados do treinamento das árvores causa o aumento do tempo necessário para que se conclua o treinamento. Ainda na Tabela 3, observa-se que os melhores resultados por métrica ficaram divididos entre os modelos com subconjuntos de treino igual a 10% e 90% dos dados de treinamento.

Assim, sabendo da importância dos acertos em momentos de mudança na condição operativa do vertimento, representada pela acurácia II e de forma indireta na acurácia média, optou-se por detalhar os resultados de cada um dos cinco treinamentos, com relação às acurácias I, II e Média, apresentados na Tabela 4. Logo, os resultados detalhados nesta tabela ressaltam o melhor desempenho do modelo com menor volume de dados para treinar cada árvore da floresta.

4.4 SENSIBILIDADE DO MODELO *MULTILAYER PERCEPTRON*

A análise de sensibilidade para o modelo MLP buscou avaliar o impacto nos resultados devidos ao número de camadas ocultas e de neurônios artificiais presentes nestas camadas. As demais características deste modelo foram mantidas as mesmas durante todas as execuções de treinamento e teste apresentadas neste documento, e são elas:

- As camadas do modelo são densamente conectadas;

Tabela 4 – Valores de acurácia I, II e Média para os cinco treinamentos/testes realizados na segunda análise de sensibilidade do modelo RF.

Amostras <i>bootstrap</i>	Treinamento/ Teste	Acurácia I	Acurácia II	Acurácia Média
10%	1	0,908	0,647	0,778
	2	0,976	0,472	0,724
	3	0,972	0,520	0,746
	4	0,876	0,635	0,756
	5	0,970	0,673	0,821
90%	1	0,951	0,518	0,735
	2	0,986	0,317	0,652
	3	0,972	0,453	0,714
	4	0,976	0,449	0,701
	5	0,978	0,627	0,803

Fonte: Elaborado pelo autor (2021).

- Função de ativação no neurônio do tipo ReLU (função de Unidade Linear Retificada);
- O treinamento dos parâmetros internos da rede foi realizado através do *Adam Optimizer*;
- Os resultados são validados, durante o treinamento, com 20% dos dados separados para a validação;
- O treinamento visa minimizar o erro quadrático médio (MSE) entre os dados observados e previstos.
- O número de épocas de treinamento é limitado em 50.

As características fixas do modelo foram escolhidas com base em estudos prévios e também com base em publicações que demonstram o bons resultados de modelos MLP que utilizam os atributos definidos, como em (JIANG *et al.*, 2018; WANG *et al.*, 2018), nos quais é demonstrado a eficiência da função de ativação ReLU. Em (KINGMA; BA, 2014), se apresenta e demonstra a efetividade do *Adam Optimizer*, e em (WANG *et al.*, 2019), este otimizador se destaca quando utilizado para banco de dados estruturados.

A Tabela 5 apresenta a análise de sensibilidade do modelo com a variação do número de camadas ocultas e também do número de neurônios em cada camada, incluindo o número de neurônios artificiais presentes nas camadas ocultas de um mesmo modelo. Foram utilizados modelos que apresentam de uma a quatro camadas ocultas. Além disso, foram definidos quatro patamares de parâmetros treináveis (parâmetros ajustáveis internamente no modelo) de forma que o número de neurônios artificiais fosse determinados de modo que o número de parâmetros treináveis ficasse próximo destes patamares. Assim, é possível identificar a sensibilidade do modelo quanto ao número de camadas ocultas de

uma forma mais justa. As configurações de número de camadas ocultas e de neurônios em cada uma pode ser visto na primeira coluna, com o valor de parâmetros treináveis em cada configuração exposto na segunda coluna desta tabela.

Tabela 5 – Média dos resultados de cinco treinamentos diferentes para cada configuração do modelo MLP.

C. Ocultas/ Neurônios	Parâmetros Treináveis	Ac. I	Ac. II	Ac. Média	Pont. F1	Precisão	Revocação	Tempo [s]
1/126	10211	0,745	0,562	0,653	0,553	0,388	0,980	38
2/68	10205	0,834	0,583	0,709	0,648	0,497	0,971	39
3/54	10319	0,886	0,530	0,708	0,732	0,604	0,963	41
4/46	10217	0,890	0,520	0,705	0,736	0,620	0,932	43
1/332	26894	0,796	0,669	0,733	0,624	0,474	0,973	40
2/128	26855	0,925	0,643	0,784	0,799	0,687	0,965	42
3/97	26874	0,885	0,698	0,791	0,737	0,607	0,953	46
4/82	27065	0,903	0,441	0,672	0,756	0,629	0,961	47
1/526	42611	0,786	0,536	0,661	0,592	0,428	0,976	41
2/169	42424	0,921	0,602	0,762	0,792	0,676	0,962	46
3/127	42804	0,926	0,554	0,740	0,800	0,696	0,958	50
4/106	42617	0,937	0,519	0,728	0,826	0,746	0,942	50
1/989	80114	0,810	0,646	0,728	0,633	0,486	0,973	48
2/245	80120	0,918	0,509	0,714	0,793	0,717	0,923	51
3/180	79745	0,879	0,584	0,731	0,719	0,593	0,951	56
4/150	80105	0,934	0,398	0,666	0,813	0,731	0,932	60

Fonte: Elaborado pelo autor (2021).

As mesmas métricas de avaliação e os cinco grupos de treinamentos utilizados para os modelos de RF também foram utilizados para o MLP. Na Tabela, estão expostos os

valores da média para resultado dos cinco treinamentos e o melhor resultado para cada métrica (por coluna) está destacado em vermelho, com exceção do tempo de treinamento do modelo. Afirmações podem ser realizadas a partir da análise desta tabela, sendo enumeradas a seguir:

1. Os resultados de acurácia I são significativamente menores nos treinamentos de modelos com apenas uma camada oculta;
2. Os resultados de acurácia II são inferiores nos treinamentos de modelos com quatro camadas ocultas;
3. A precisão dos modelos é normalmente baixa, o que indica a ocorrência de muito falsos positivos, levando a possíveis decisões do operador que podem levar a desperdício de recursos;
4. A taxa de revocação dos modelos é alta, logo, o modelos têm uma alta taxa de acertos nos momentos em que o vertimento deve ser previsto;
5. Os tempos não apresentaram muitas diferenças, mas se pode perceber que o aumento no número de camadas ocultas aumenta o tempo de treinamento;
6. Dois modelos se destacaram com os melhores resultados, são eles: o modelo com 3 camadas ocultas e 97 neurônios em cada, e o modelo com 4 camadas ocultas e 106 neurônios em cada.

Os resultados individuais de acurácia (I, II e Média) dos cinco treinamentos dos dois melhores modelos MLP estão expostos na Tabela 6. Como mencionado anteriormente, a acurácia II representa a capacidade de previsões corretas em momentos em que se deve ocorrer mudança na condição de vertimento. Assim, pode-se verificar para o modelo com mais camadas ocultas que para uma acurácia I acima de 0,9, os valores de acurácia II não foram superiores a 0,6, porém, para o modelo com três camadas ocultas, houve melhores valores de acurácia II e mais equilíbrio entre as acurácias I e II, observados com a métrica acurácia Média.

4.5 SENSIBILIDADE À MODIFICAÇÃO NO BANCO DE TREINAMENTO

Esta análise de sensibilidade diz respeito ao tratamento do banco de treinamentos. Assim, os modelos são treinados com o banco de treinamento sem passar por nenhuma modificação, mas são treinados com o banco de treinamentos após passar pelo balanceamento de informações. Porém este último implica em um aumento no volume de dados contido no banco de treinamento, tal volume é quase três vezes maior que o volume original. Logo, buscando uma análise mais robusta, o banco de treinamentos também passará uma

Tabela 6 – Valores de acurácia I, II e Média para os cinco treinamentos/testes realizados na segunda análise de sensibilidade do modelo MLP.

C. Ocultas Neurônios	Treinamento/ Teste	Acurácia I	Acurácia II	Acurácia Média
3/97	1	0,882	0,640	0,761
	2	0,888	0,699	0,794
	3	0,943	0,747	0,845
	4	0,790	0,494	0,642
	5	0,921	0,909	0,915
4/106	1	0,951	0,446	0,699
	2	0,949	0,398	0,674
	3	0,878	0,667	0,772
	4	0,948	0,583	0,766
	5	0,961	0,500	0,730

Fonte: Elaborado pelo autor (2021).

repetição completas das suas informações em três vezes. Assim, o banco de treinamento dos modelos poderá passar pelo balanceamento de dados, será utilizado sem nenhuma modificação e também passará por uma repetição de informações que o deixará com três vezes o volume de dados inicial.

Nesta análise, serão utilizados todos os 16 diferentes grupos de treinamento e teste produzidos com a técnica de validação cruzada. Serão utilizados um modelo de RF e um de MLP, sendo o modelo RF composto de 500 árvores e subgrupos de treinamento das árvores gerados com 10% do número de dados de treinamento. O modelo MLP é composto de 97 neurônios artificiais em cada uma de suas 3 camadas ocultas.

O Quadro 1 apresenta a nomenclatura dada aos modelos a serem treinados e suas características quanto ao método de aprendizagem e também quanto a modificação do banco de treinamento.

Quadro 1 – Descrição dos modelos e do tratamento realizado no banco de dados de treinamento.

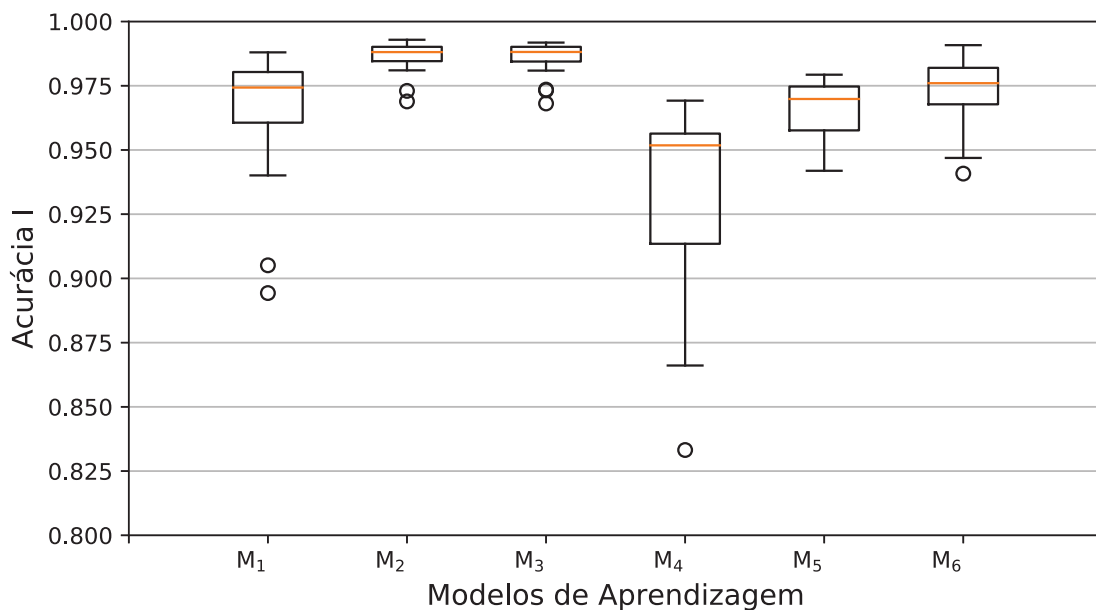
Nome do Modelo	Aprendizagem de Máquina	Modificação no Banco de Treinamento
M ₁	RF	Balanceamento dos dados
M ₂	RF	Sem modificações
M ₃	RF	Triplicar os dados
M ₄	MLP	Balanceamento dos dados
M ₅	MLP	Sem modificações
M ₆	MLP	Triplicar os dados

Fonte: Elaborado pelo autor (2021).

Os resultados de acurácia I encontrados com os 16 treinamentos realizados com

cada um dos seis modelos estão apresentados no Gráfico 3. É possível observar que os três primeiros modelos, baseados em RF, se comportam melhor que os três últimos, baseados em MLP, visto que os primeiros apresentam valores de mediana mais altos e também menor dispersão de resultados. Além disso, analisando os treinamentos que utilizaram dados balanceados, percebe-se o aumento da dispersão dos resultados independente do modelo de aprendizagem abordado.

Gráfico 3 – O gráfico boxplot com os valores de acurácia I para a validação cruzada feita com os modelos do Quadro 1.

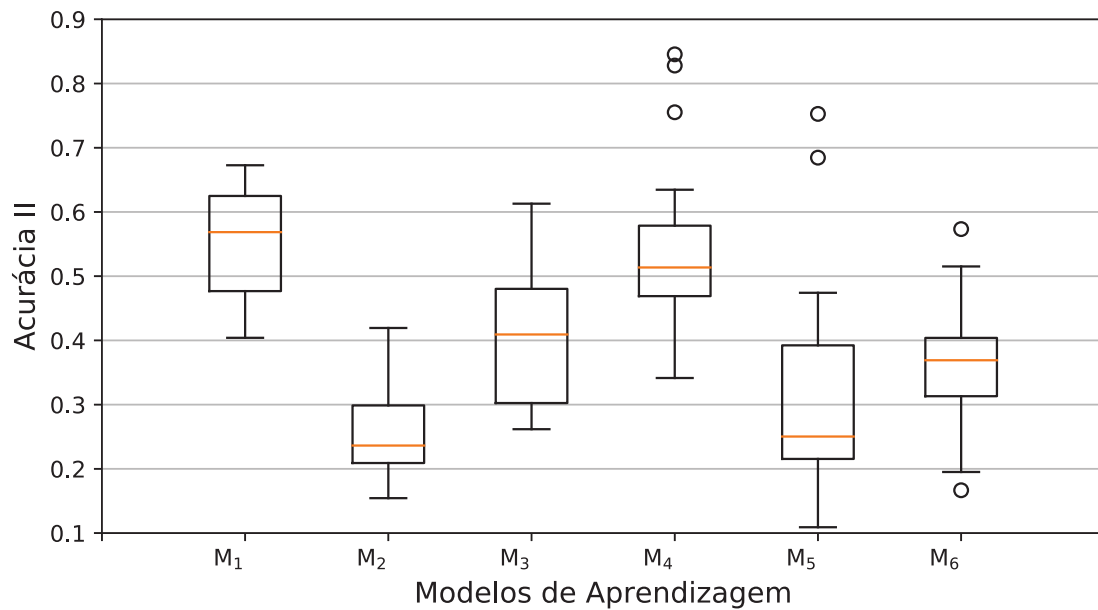


Fonte: Elaborado pelo autor (2021).

O Gráfico 4 apresenta os resultados de acurácia II, onde observa-se grande dispersão dos resultados para todos os seis modelos. M1 e M4, que passaram pelo balanceamento dos dados de treinamento, apresentaram seus valores de mediana mais expressivos que os demais. O modelo M4 apresentou três pontos de *outliers* dos quais, dois são superiores a 0,8.

A precisão dos modelos, encontrada com as simulações, pode ser encontrada no Gráfico 5. Novamente, os resultados para M1 e M4 apresentaram maior dispersão e menores valores de mediana. Os modelos baseados em MLP foram os que apresentaram menor precisão quando comparados com os modelos baseados em RF que tiveram os mesmos tipos de tratamento do banco de treinamento. Os modelos M2 e M3, de destacaram, com resultados superiores a 0,95, o que indica que estes dois modelos são mais seguros para evitar falso positivos. Este gráfico também permite observar que a utilização do tratamento (balanceamento) dos dados de treinamento está interferindo negativamente na precisão dos modelos, ou seja, esses modelos ficam mais propensos a indicar necessidade

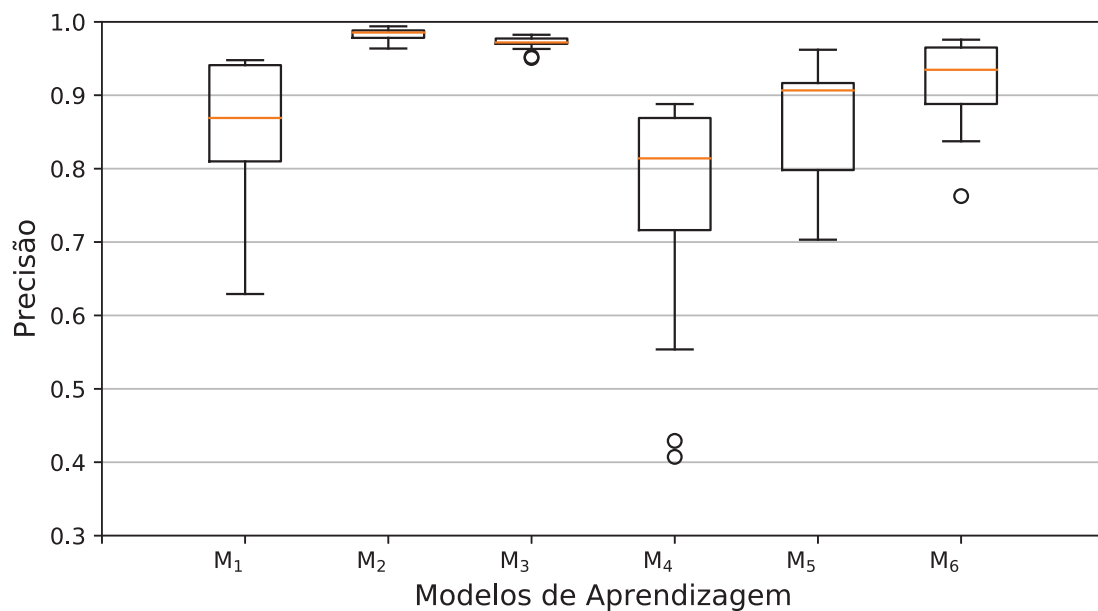
Gráfico 4 – O gráfico boxplot com os valores de acurácia II para a validação cruzada feita com os modelos do Quadro 1.



Fonte: Elaborado pelo autor (2021).

de vertimento quando na verdade, deveriam indicar a não necessidade de vertimento.

Gráfico 5 – O gráfico boxplot com os valores de precisão para a validação cruzada feita com os modelos do Quadro 1.

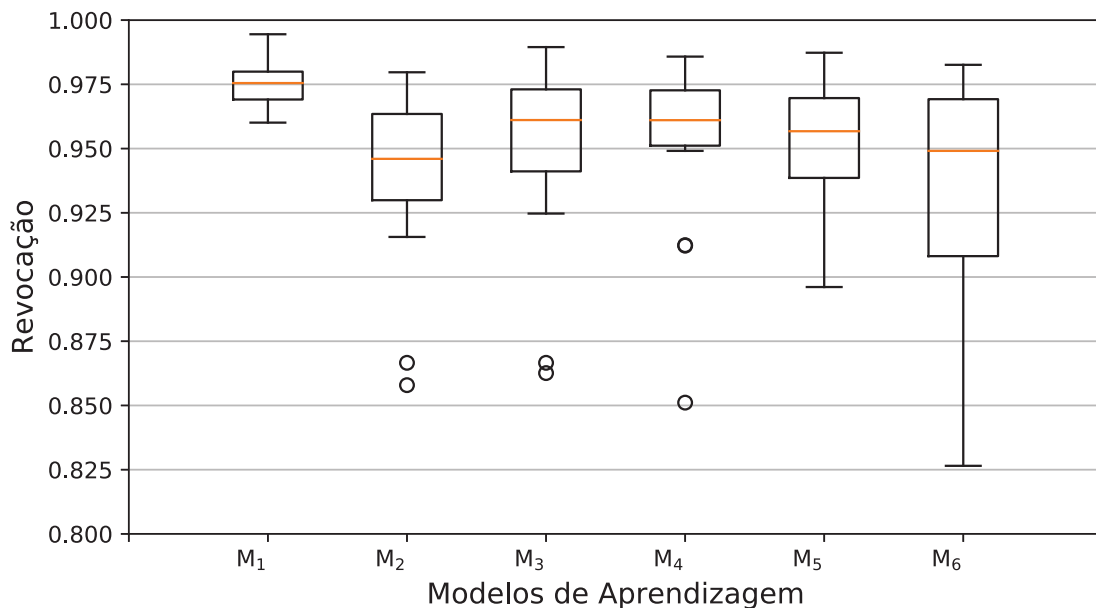


Fonte: Elaborado pelo autor (2021).

O Gráfico 6 expõe os resultados referentes à revocação dos modelos treinados.

O destaque positivo é do modelo M1, o qual apresentou baixa dispersão de resultados e a maior mediana entre os seis modelos analisados, sendo esta superior a 0,975. O comportamento dos modelos M1 e M2, que passaram pelo balanceamento dos dados, apresentou menor dispersão quando comparados a outras formas de modificação ou não modificação. Contudo, para os seis modelos, as medidas de mediana foram superiores 0,94, o que deixa afirmar que são bons modelos em prever o vertimento quando realmente deve acontecer.

Gráfico 6 – O gráfico boxplot com os valores de revocação para a validação cruzada feita com os modelos do Quadro 1.

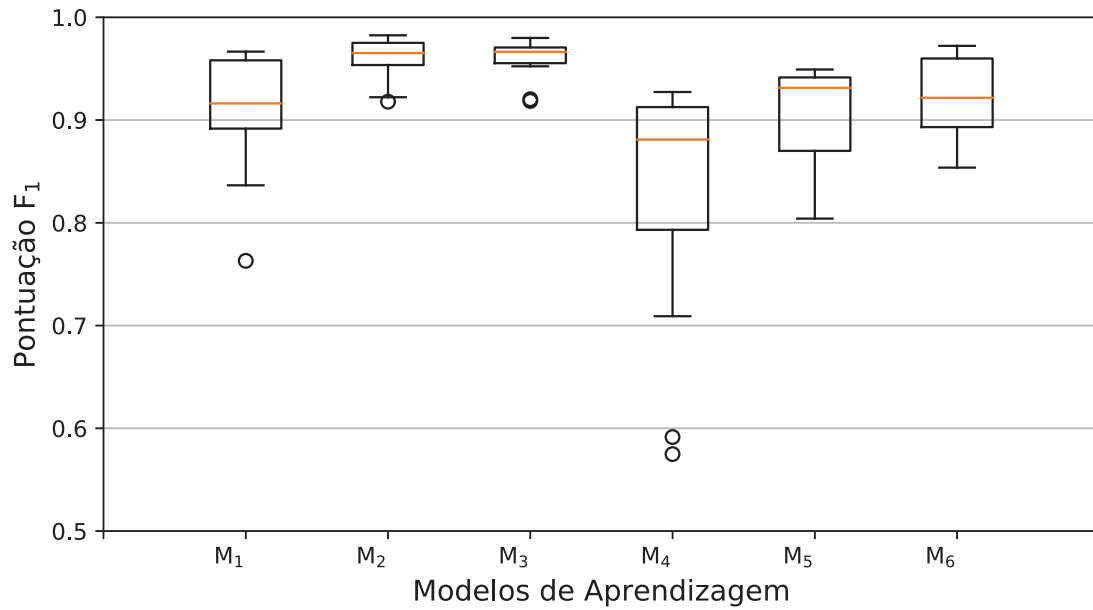


Fonte: Elaborado pelo autor (2021).

A pontuação F1, apresentada no Gráfico 7, representa uma média ponderada entre os resultados de precisão e revocação e dessa forma, pode-se obter um resultado que leva em consideração tanto os resultados falso positivos quanto falso negativos. Contudo, como o desbalanceamento no grupo de teste é muito grande, isso acaba refletindo também no resultado da pontuação F1, visto que os Gráficos 5 e 7 apresentam muita similaridade no formato de suas caixas. Porém, ocorreu um suave deslocamento para cima, com todos os modelos, pois os valores de revocação são mais elevados que os valores de precisão.

No Gráfico 8, onde estão expostos os tempos de treinamento dos modelos, observa-se os modelos RF com custos computacionais muito menores que os modelos MLP. Em ambos os métodos de aprendizagem, é possível identificar que o treinamento mais rápido foi com o banco de treinamento sem modificações, seguido do treinamento com balanceamento dos dados e o mais lento é quando se triplica o banco de treinamento. Esse comportamento era esperado, pois esta é a ordem crescente de dados por tipo de banco de treinamentos.

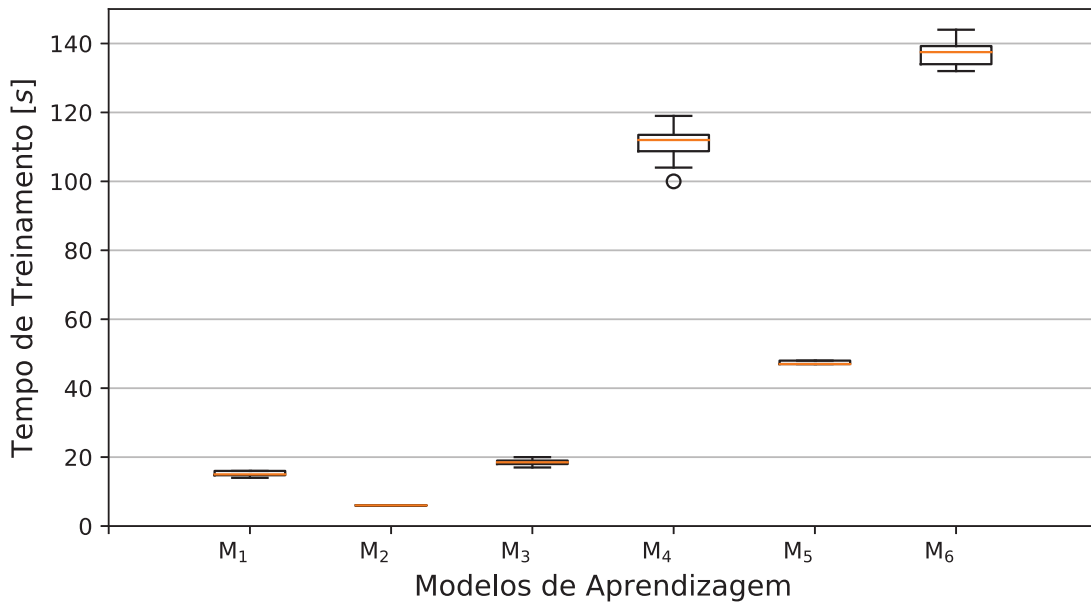
Gráfico 7 – O gráfico boxplot com os valores de pontuação F1 para a validação cruzada feita com os modelos do Quadro 1.



Fonte: Elaborado pelo autor (2021).

Em âmbito geral, os modelos RF apresentaram tempos de treinamentos inferiores a 20 segundos, enquanto os modelos MLP, em alguns casos, passaram de dois minutos de tempo de treinamento.

Gráfico 8 – O gráfico boxplot com os valores de tempo de treinamento para a validação cruzada feita com os modelos do Quadro 1.



Fonte: Elaborado pelo autor (2021).

4.6 ANÁLISE DO AGRUPAMENTO DE MODELOS

Nesta seção são definidos oito diferentes grupos de modelos de aprendizagem, avaliados na seção anterior, de modo a avaliar seus resultados e identificar possíveis melhorias nos resultados. O Quadro 2 apresenta a nomenclatura dada a cada um dos oito grupos definidos, quais modelos os compõem e qual a forma de decidir a resposta deste grupo ("Tipo de Agrupamento").

Quadro 2 – Definição dos agrupamentos de modelos de aprendizagem.

Nome do Modelo	Composição do Grupo	Tipo de Agrupamento
M ₇	M ₁ , M ₂ , M ₃	Maioria
M ₈	M ₄ , M ₅ , M ₆	Maioria
M ₉	M ₁ , M ₂ , M ₃ , M ₄ , M ₆	Maioria
M ₁₀	M ₁ , M ₃ , M ₄ , M ₅ , M ₆	Maioria
M ₁₁	M ₁ , M ₄	Unanimidade
M ₁₂	M ₃ , M ₆	Unanimidade
M ₁₃	M ₁ , M ₃	Unanimidade
M ₁₄	M ₄ , M ₆	Unanimidade

Fonte: Elaborado pelo autor (2021).

Os modelos M₇ e M₁₃ são formados apenas com modelos RF, já os modelos M₈ e M₁₄ são formados apenas por modelos MLP. Os modelos M₉ e M₁₀ são definidos com 5 modelos,

sendo excluído do primeiro o modelo MLP que não passou pro tratamento do banco de dados e do segundo é excluído o modelo RF que também não passou por tratamento do banco de treinamento. Os modelos M_{11} e M_{12} são formados por dois modelos, um RF e outro MLP, que passaram pelo mesmo tratamento do banco de treinamento.

São entregues como resposta dos modelos de previsão apenas duas condições, a primeira onde não há vertimento e a segunda quando há vertimento, como apresentado no Capítulo 3. Porém, há a possibilidade de em certos momentos, quando a previsão é de vertimento, as cinco horas previstas pelo modelo podem não apresentar o mesmo nível lógico, ou seja, em três das cinco horas o modelo prevê vertimento. Assim, dada a importância desses momentos, o Quadro 3 foi estruturado para auxiliar no entendimento das decisões do agrupamento com decisão por maioria simples.

Quadro 3 – Definição dos resultados para combinação de modelos com decisão por maioria simples.

Contador⁰	Contador³	Contador⁵	Agrupamento
*			0
	*		3
		*	5
*	*		3
	*	*	3
*		*	3
*	*	*	3

Fonte: Elaborado pelo autor (2021).

O Contador⁰ representa o número de modelos que deram como saída a condição de não vertimento (zero), o Contador⁵ representa o número de modelos que entregaram como resposta a condição de vertimento (cinco) e o Contador³ representa aqueles modelos que deram a condição de vertimento como resposta, mas não apresentaram esta condição em todas as cinco horas de previsão. No Quadro 3, o símbolo “*” representa o contador com maior valor, e quando há mais de uma marcação por linha, significa que esses contadores apresentam igualmente o maior valor.

No Quadro 4 a resposta do agrupamento dado cada entrada possível dos dois modelos envolvidos. Os momento em que há a previsão de vertimento, porém não em todas as cinco horas de previsão, são tratadas apresentadas com valor 3, simplesmente por ser um valor entre zero e cinco. Assim, podemos verificar que, dado um dos modelos geradores do agrupamento apresenta 5 como resposta e o outro apresenta o 3, então o agrupamento apresenta como resposta o valor 3, que indica o vertimento assim como o valor 5, contudo, indica momentos de mudança na condição de vertimento. Nestes momentos, um dos modelos geradores do agrupamento apresenta 0 como resposta e o outro apresenta 3 ou 5 e então, este modelo não vai ter resposta, visto que a decisão por

vertimento ou não vertimento, não é unânime.

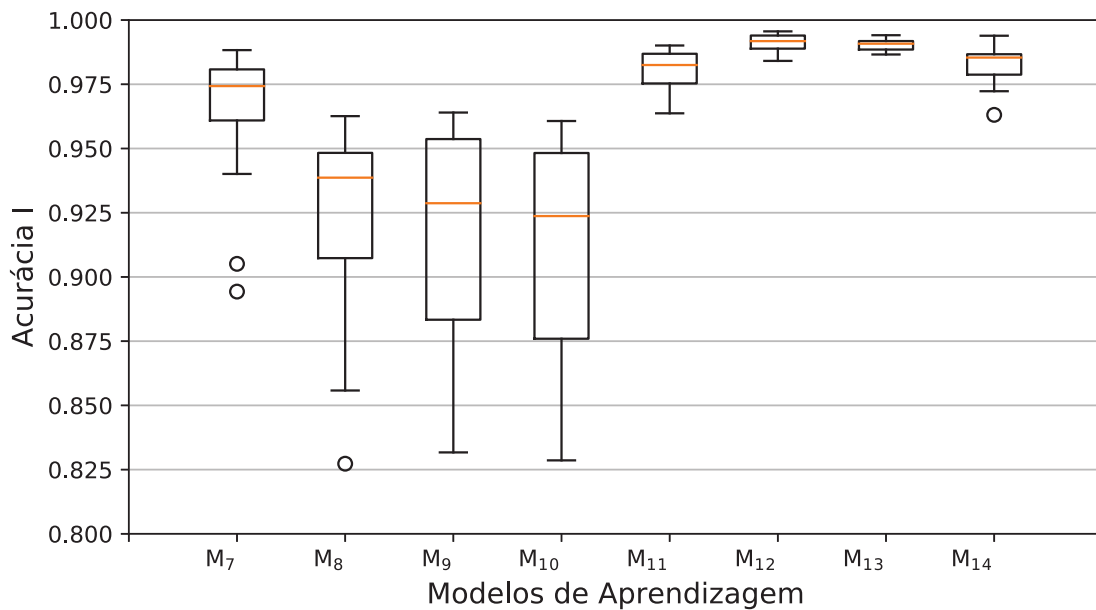
Quadro 4 – Combinação dos resultados para combinação de modelos com decisão por unanimidade.

Modelo A	Modelo B	Agrupamento
0	0	0
0	5	-
5	0	-
5	5	5
3	5	3
5	3	3
3	3	3
0	3	-
3	0	-

Fonte: Elaborado pelo autor (2021).

Os Gráficos de 9 a 13 apresentam os resultados das avaliações dos modelos descritos no Quadro 2. No Gráfico 9 é possível notar que o comportamento do modelo M_7 é muito semelhante ao M_1 (Gráfico 3), que é um de seus formadores. Os modelos M_8 , M_9 e M_{10} , que utilizam a decisão por maioria, apresentou grande dispersão dos resultados. Entretanto, os modelos com decisão por unanimidade apresentaram resultados muito altos e pouco dispersos, com destaque ao modelo M_{13} .

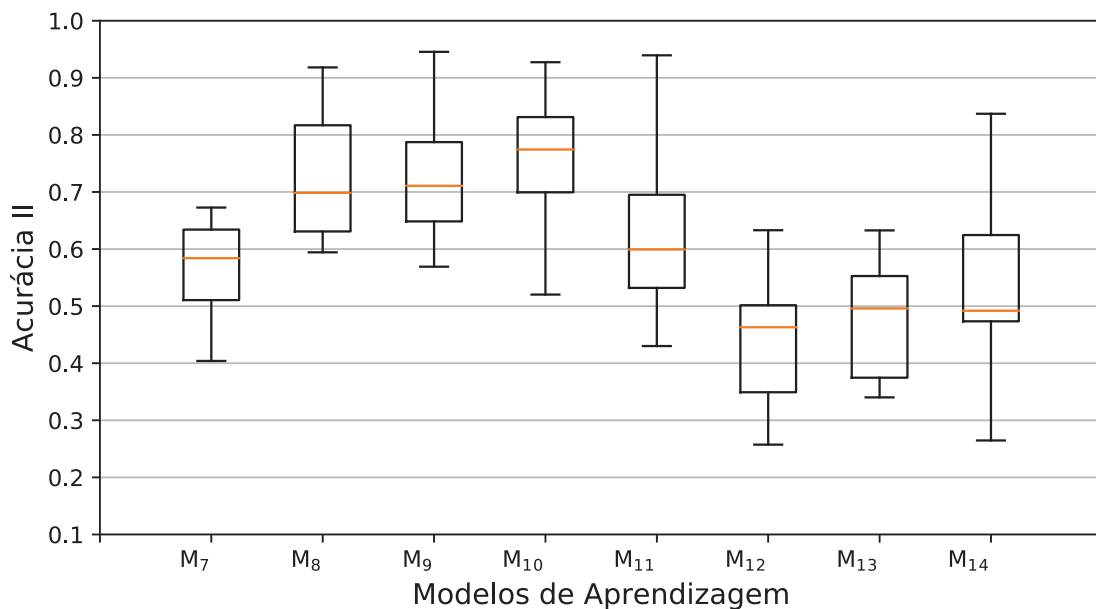
Gráfico 9 – O gráfico boxplot com os valores de acurácia I para a validação cruzada feita com os agrupamentos de modelos.



Fonte: Elaborado pelo autor (2021).

A acurácia II, referente aos momentos de mudança na condição do vertimento, está apresentada no Gráfico 10. Pode-se observar que houve elevação nos resultados em comparação com os modelos de M_1 a M_6 (Gráfico 4). O comportamento de bons resultados para acurácia I e resultados ruins para acurácia II podem ser notados nos modelos M_{12} e M_{14} . No entanto, os modelos M_{11} e M_{14} apresentaram elevados valores de acurácia I e uma grande dispersão de valores em acurácia II, o que permite inferir que em alguns treinamentos houve bons resultados para ambas medidas de acurácia, com destaque ao modelo M_{12} , composto de um RF e um MLP, ambos com tratamento de balanceamento dos dados.

Gráfico 10 – O gráfico boxplot com os valores de acurácia II para a validação cruzada feita com os agrupamentos de modelos.



Fonte: Elaborado pelo autor (2021).

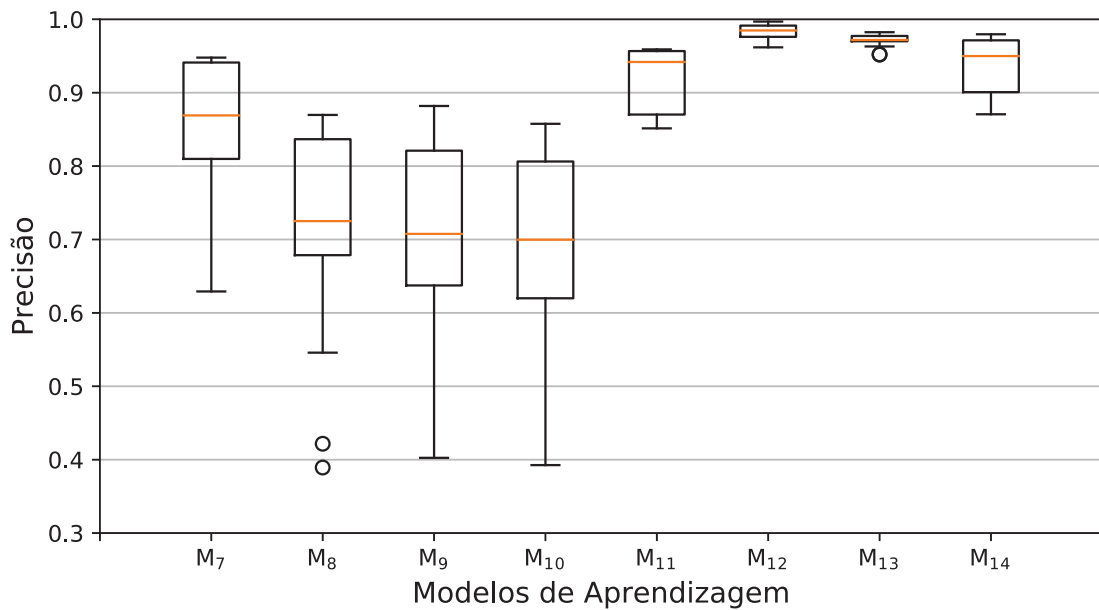
O Gráfico 11 mostra que a precisão para os grupos com decisão por maioria comum foram dispersos e alcançaram valores muito baixos, como nos modelos M_8 e M_9 . Os modelos de M_{11} a M_{14} apresentaram valores superiores a 0,85 para todas as avaliações com esta métrica.

Os valores de revocação, apresentaram melhora nos resultados para todos os agrupamentos, visto que os valores foram comumente superiores a 0,95, como pode ser visto no Gráfico 12.

Os resultados de pontuação F1 se apresentaram com formato similar ao da precisão, porém com um certo deslocamento verticalmente para cima no gráfico, devido aos altos valores de revocação.

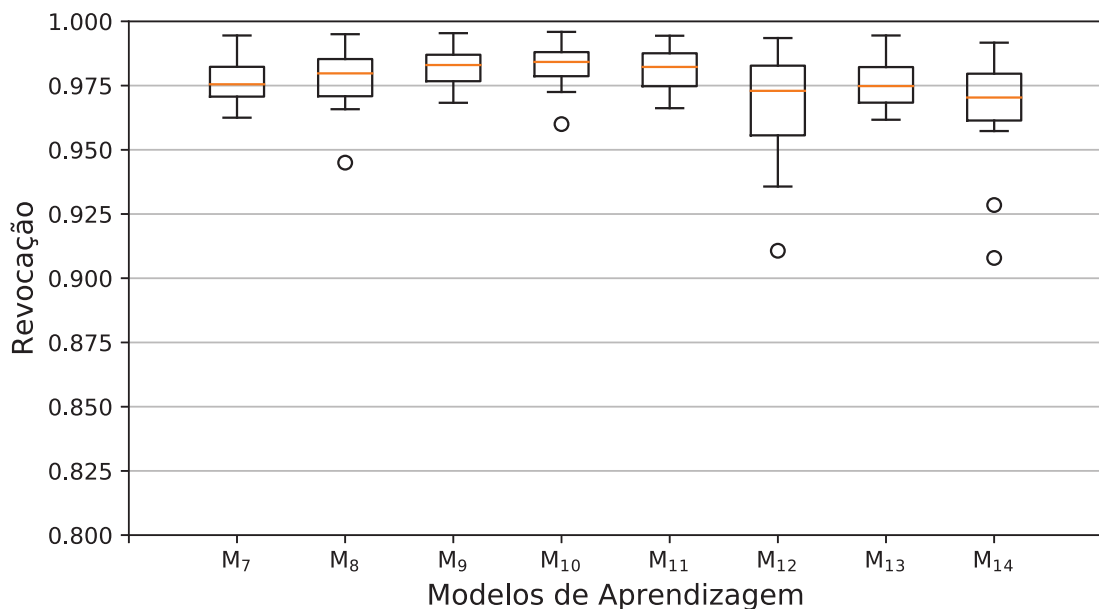
O agrupamento com decisão baseada no resultado unânime dos modelos envolvidos,

Gráfico 11 – O gráfico boxplot com os valores de precisão para a validação cruzada feita com os agrupamentos de modelos.



Fonte: Elaborado pelo autor (2021).

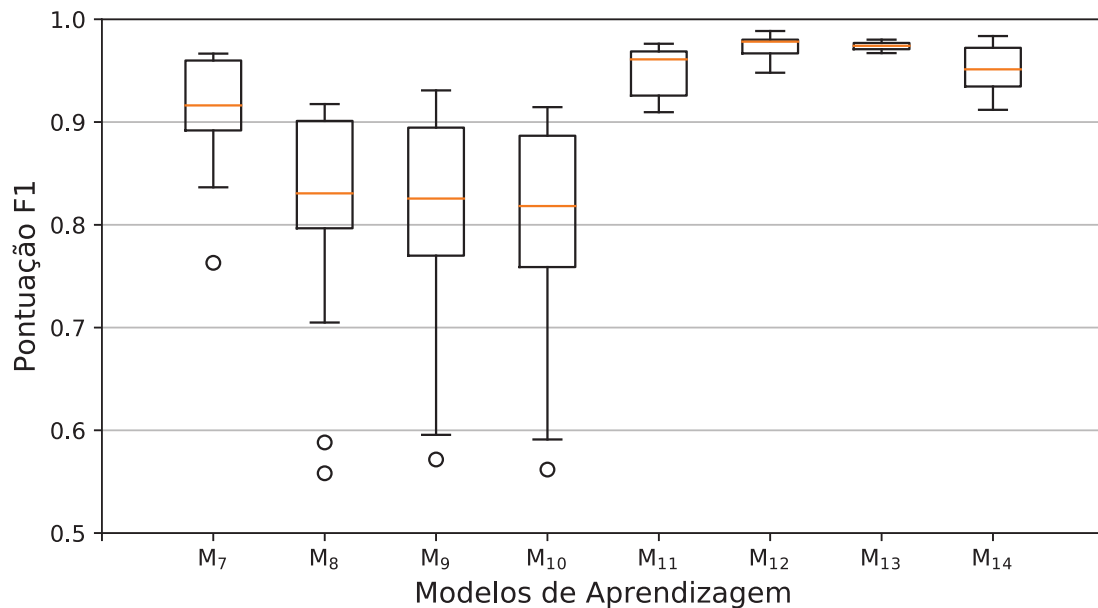
Gráfico 12 – O gráfico boxplot com os valores de revocação para a validação cruzada feita com os agrupamentos de modelos.



Fonte: Elaborado pelo autor (2021).

normalmente apresenta momentos sem resposta, ou seja, não informará a condição operativa de vertimento em alguns momentos ao longo do ano. Assim, o Quadro 5 apresenta os resultados dos treinamentos/testes os 16 grupos de validação cruzada, para cada um dos

Gráfico 13 – O gráfico boxplot com os valores de pontuação F1 para a validação cruzada feita com os agrupamentos de modelos.



Fonte: Elaborado pelo autor (2021).

quatro modelos com decisão com base em unanimidade. Estão expostos o número de amostra de testes que o modelo não apresentou resposta e também o valor em porcentagem, com relação ao número total do banco de teste. Logo, pode-se observar que os modelos M_{11} e M_{14} , normalmente apresentaram valores mais altos de perdas de previsão do que os modelos M_{12} e M_{13} .

O desbalanceamento de informações refletem tanto no banco de dados de treinamento quanto no banco de dados de teste, conseqüentemente os resultados tendem a ser diferentes. As Tabelas 7 e 8 apresentam os resultados das métricas de avaliação para todos os modelos, para dois grupos de validação cruzada diferentes, sendo o primeiro grupo de teste composto por 71,41% saídas indicando não vertimento, 27,14% indicando vertimento e 1,55% das saídas indicavam os momentos de troca. Nesta tabela, observa-se que os valores de acurácia II fica ruim quando o modelo apresenta bons valores de acurácia I, ou o contrário. O modelo que obteve o resultado mais equilibrado foi o M_{11} . Porém este é um modelo de agrupamento baseado em unanimidade, e deixa sem resposta 5,45% dos momentos de previsão (Etapa 1 do Quadro 5).

A Tabela 8 apresenta os resultados para um grupo de validação cruzada de teste com 79,73% saídas indicando não vertimento, 19,18% indicando vertimento e 1,09% das saídas indicavam os momentos de troca. Nesta tabela vários modelos apresentaram bons desempenhos, porém os modelos de agrupamento, M_9 , M_{10} , M_{11} e M_{14} se destacaram. Os modelos M_{11} e M_{14} obtiveram resultados ainda mais altos, contudo há uma perda

Quadro 5 – Quantidade de horas sem resposta de previsão para os modelos com decisão por unanimidade.

Etapa	Modelo M ₁₁		Modelo M ₁₂		Modelo M ₁₃		Modelo M ₁₄	
	Sem Previsão	% do Total	Sem Previsão	% do Total	Sem Previsão	% do Total	Sem Previsão	% do Total
1	463	5,45	220	2,59	467	5,50	362	4,26
2	779	9,17	228	2,68	788	9,27	335	3,94
3	832	9,79	494	5,81	616	7,25	750	8,83
4	1048	12,34	273	3,21	988	11,63	661	7,78
5	222	2,61	81	0,95	78	0,92	249	2,93
6	228	2,68	115	1,35	79	0,93	208	2,45
7	436	5,13	471	5,54	184	2,17	475	5,59
8	297	3,50	168	1,98	152	1,79	251	2,95
9	340	4,00	166	1,95	87	1,02	339	3,99
10	349	4,11	114	1,34	55	0,65	364	4,28
11	467	5,50	243	2,86	251	2,95	417	4,91
12	1280	15,07	188	2,21	199	2,34	1365	16,07
13	249	2,93	122	1,44	50	0,59	231	2,72
14	1070	12,59	81	0,95	29	0,34	1126	13,25
15	322	3,79	188	2,21	217	2,55	351	4,13
16	1048	12,34	159	1,87	170	2,00	1123	13,22

Fonte: Elaborado pelo autor (2021).

significativa nas respostas desses modelos, 15,07% e 16,07%, respectivamente. Por outro lado, os outros modelos apresentam resultados um pouco inferiores, mas não possuem essa perda de dados que ocorre com os agrupamentos com decisão baseada em unanimidade.

Tabela 7 – Resultados dos treinamentos/testes com um grupo de validação cruzada com 71,41% de amostras de não vertimento, 27,14% de amostras de vertimento e 1,55% de amostras com mudança na condição de vertimento.

Modelo	Acurácia	Acurácia	Precisão	Revocação	Pontuação
	I	II			F1
M ₁	0,9475	0,6288	0,8599	0,9753	0,9140
M ₂	0,9821	0,2424	0,9939	0,9432	0,9679
M ₃	0,9735	0,4318	0,9816	0,9247	0,9523
M ₄	0,9523	0,4470	0,8880	0,9535	0,9196
M ₅	0,9590	0,4470	0,9082	0,9531	0,9301
M ₆	0,9752	0,3864	0,9691	0,9432	0,9560
M ₇	0,9475	0,6288	0,8599	0,9753	0,9140
M ₈	0,9383	0,6364	0,8386	0,9712	0,9000
M ₉	0,9321	0,7045	0,8178	0,9811	0,8920
M ₁₀	0,9215	0,7045	0,7933	0,9811	0,8773
M ₁₁	0,9758	0,6214	0,9408	0,9792	0,9596
M ₁₂	0,9870	0,3966	0,9955	0,9575	0,9761
M ₁₃	0,9873	0,5377	0,9816	0,9740	0,9778
M ₁₄	0,9844	0,4646	0,9796	0,9658	0,9727

Fonte: Elaborado pelo autor (2021).

Tabela 8 – Resultados dos treinamentos/testes com um grupo de validação cruzada com 79,73% de amostras de não vertimento, 19,18% de amostras de vertimento e 1,09% de amostras com mudança na condição de vertimento.

Modelo	Acurácia	Acurácia	Precisão	Revocação	Pontuação
	I	II			F1
M ₁	0,9861	0,6237	0,9441	0,9901	0,9666
M ₂	0,9929	0,4194	0,9854	0,9797	0,9825
M ₃	0,9918	0,6129	0,9704	0,9895	0,9799
M ₄	0,9653	0,8280	0,8653	0,9814	0,9197
M ₅	0,9785	0,7527	0,9144	0,9861	0,9489
M ₆	0,9816	0,4409	0,9307	0,9826	0,9559
M ₇	0,9861	0,6237	0,9441	0,9901	0,9666
M ₈	0,9589	0,8495	0,8360	0,9919	0,9073
M ₉	0,9617	0,9140	0,8439	0,9954	0,9134
M ₁₀	0,9569	0,9247	0,8269	0,9959	0,9036
M ₁₁	0,9901	0,8987	0,9590	0,9941	0,9762
M ₁₂	0,9938	0,6203	0,9762	0,9935	0,9848
M ₁₃	0,9918	0,6237	0,9704	0,9901	0,9802
M ₁₄	0,9867	0,8353	0,9458	0,9917	0,9682

Fonte: Elaborado pelo autor (2021).

5 CONCLUSÕES E TRABALHOS FUTUROS

Neste capítulo, serão expostas as conclusões relacionadas ao trabalho apresentado e também serão apresentadas propostas de trabalho futuro que visam à melhora da metodologia e dos resultados obtidos.

5.1 CONCLUSÃO

As decisões operacionais em uma UHE estão associadas a diversos fatores, como capacidade de geração, disponibilidade de recursos hídricos e segurança da barragem. Este trabalho implementou uma ferramenta capaz de prever a necessidade ou não de vertimento para cinco horas futuras, com o objetivo de auxiliar o operador da UHE na decisão em tempo real. A validação desta metodologia foi feita para a UHE Lajeado, no rio Tocantins, Brasil, porém pode ser expandida para outras usinas com características similares, onde se tenha acesso a informações da operação em tempo real e informações de estações de telemetria a montante da barragem.

Os treinamentos e testes dos modelos foram realizados por meio dos dados históricos de operação da usina e de dados das estações de telemetria. Foram estudados alguns fatores que poderiam influenciar nos resultados dos modelos, como os fatores internos dos modelos de aprendizagem e modificações no banco de treinamento dos modelos, visando corrigir o desbalanceamento de dados e também a estratégia de agrupamento de modelos.

O estudo de correlação possibilitou a redução do número de entradas dos modelos de aprendizagem e conseqüentemente, a diminuição do custo computacional para a realização dos treinamentos. As falhas no dados históricos de telemetria foram superados com os tratamentos utilizados para correção destes, uma vez que os índices demonstram uma alta eficiência na capacidade de previsão.

As análises demonstraram que a técnica baseada em RF obteve melhor desempenho com relação aos modelos MLP, independente do tipo de tratamento utilizado no banco de treinamento. Além disso, para ambos os modelos é possível verificar que valores altos de acurácia geral tendem a refletir em baixos valores de acurácia nos momentos de mudança na condição operativa de vertimento. E os modelos com melhor capacidade preditiva em momentos de troca do estado do vertimento foram encontrados nos treinamentos com o balanceamento dos dados.

Os agrupamentos de modelos com decisão em maioria simples apresentaram uma significativa melhora nas previsões de momentos de mudança na condição de vertimento, contudo para a acurácia geral, os resultados foram mais dispersos e até inferiores aos seus modelos geradores. Todavia, os agrupamentos com decisão por unanimidade apresentaram resultados com melhoras em todas as métricas de avaliação, principalmente o modelo M_{11} , composto pelos modelos RF e MLP com o balanceamento dos dados. O problema destes

últimos modelos são os momentos em que estes não entregam nenhuma resposta, logo, o operador não terá uma previsão base para se basear na tomada de decisão.

Desta forma, os resultados demonstram que a metodologia proposta é capaz de fornecer boa assertividade nas previsões, sendo assim possível auxiliar os operadores das usinas na tomada de decisão quanto às condições de operação de vertimento.

A modelo de previsão de vertimento proposto neste trabalho ainda está em fase de implementação dentro do sistema de controle da hidrelétrica em estudo, assim, não foi possível apresentar resultados da sua aplicação em tempo real na usina.

5.2 TRABALHOS FUTUROS

Como possíveis formas de melhoria da metodologia e conseqüentemente do resultados, são sugeridos alguns temas a serem abordados como trabalhos futuros:

- Validar a metodologia em outras usinas hidrelétricas com características distintas.
- Avaliar o desempenho da metodologia em um volume maior de informações e utilização de processamento paralelo.
- Estudar a aplicação de outros modelos de aprendizagem de máquina com boa capacidade de trabalhar com banco de dados desbalanceado;
- Utilização de técnicas inteligentes para sintonizar os hiperparâmetros dos modelos de aprendizagem de forma mais eficiente;
- Avaliar outras técnicas de balanceamento de dados a serem aplicadas nos dados de treinamento;

REFERÊNCIAS

- ABDAR, M.; YEN, N. Y.; HUNG, J. C.-S. Improving the diagnosis of liver disease using multilayer perceptron neural network and boosted decision trees. **Journal of Medical and Biological Engineering**, Springer, v. 38, n. 6, p. 953–965, 2018.
- ABRITTA, R. *et al.* Hydroelectric operation optimization and unexpected spillage indications. **Energies**, Multidisciplinary Digital Publishing Institute, v. 13, n. 20, p. 5368, 2020.
- AGOSTINELLI, F. *et al.* Learning activation functions to improve deep neural networks. **International Conference on Learning Representations (ICLR)**, 2015.
- ALI, Z. *et al.* Forecasting drought using multilayer perceptron artificial neural network model. **Advances in Meteorology**, Hindawi, v. 2017, 2017.
- ARAGÃO, A. Passos de; ASANO, P. T. L.; RABÊLO, R. de A. L. A reservoir operation policy using inter-basin water transfer for maximizing hydroelectric benefits in brazil. **Energies**, Multidisciplinary Digital Publishing Institute, v. 13, n. 10, p. 2564, 2020.
- AWAD, M.; KHANNA, R. **Efficient learning machines: theories, concepts, and applications for engineers and system designers**. [S.l.]: Springer Nature, 2015.
- BELGIU, M.; DRĂGUȚ, L. Random forest in remote sensing: A review of applications and future directions. **ISPRS Journal of Photogrammetry and Remote Sensing**, Elsevier, v. 114, p. 24–31, 2016.
- BELVEDERESI, C. *et al.* Predicting river flow using an ai-based sequential adaptive neuro-fuzzy inference system. **Water**, Multidisciplinary Digital Publishing Institute, v. 12, n. 6, p. 1622, 2020.
- BENESTY, J. *et al.* Pearson correlation coefficient. In: **Noise reduction in speech processing**. [S.l.]: Springer, 2009. p. 1–4.
- BERA, S.; SHRIVASTAVA, V. K. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. **International Journal of Remote Sensing**, Taylor & Francis, v. 41, n. 7, p. 2664–2683, 2020.
- BERRAR, D. Cross-validation. **Encyclopedia of bioinformatics and computational biology**, Academic, v. 1, p. 542–545, 2019.
- BOSSANYI, E. Short-term wind prediction using kalman filters. **Wind Engineering**, JSTOR, v. 9, n. 1, p. 1–8, 1985.
- BREIMAN, L. Bagging predictors. **Machine learning**, Springer, v. 24, n. 2, p. 123–140, 1996.
- BREIMAN, L. Out-of-bag estimation. Technical report, Statistics Department, University of California Berkeley, 1996.
- BREIMAN, L. Random forests. **Machine learning**, Springer, v. 45, n. 1, p. 5–32, 2001.

CASTRO, C. M. B. D. **Planejamento Energético da Operação de Médio Prazo Conjugando as Técnicas de PDDE, PAR(P) e Bootstrap**. Tese (Doutorado) — Universidade Federal de Juiz de Fora, 2012.

CELEBI, M. E.; AYDIN, K. **Unsupervised learning algorithms**. [S.l.]: Springer, 2016.

CERRADA, M. *et al.* Fault diagnosis in spur gears based on genetic algorithm and random forest. **Mechanical Systems and Signal Processing**, Elsevier, v. 70, p. 87–103, 2016.

CHEN, J. *et al.* Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. **Energy conversion and management**, Elsevier, v. 165, p. 681–695, 2018.

CLEVERT, D.-A.; UNTERTHINER, T.; HOCHREITER, S. Fast and accurate deep network learning by exponential linear units (elus). **International Conference on Learning Representations (ICLR)**, 2016.

DIOP, L. *et al.* Annual rainfall forecasting using hybrid artificial intelligence model: Integration of multilayer perceptron with whale optimization algorithm. **Water Resources Management**, Springer, v. 34, n. 2, p. 733–746, 2020.

EFRON, B.; TIBSHIRANI, R. J. **An introduction to the bootstrap**. [S.l.]: CRC press, 1994.

FINARDI, E.; TAKIGAWA, F.; BRITO, B. Assessing solution quality and computational performance in the hydro unit commitment problem considering different mathematical programming approaches. **Electric Power Systems Research**, Elsevier, v. 136, p. 212–222, 2016.

FREDO, G. L. M.; FINARDI, E. C.; MATOS, V. L. de. Assessing solution quality and computational performance in the long-term generation scheduling problem considering different hydro production function approaches. **Renewable energy**, Elsevier, v. 131, p. 45–54, 2019.

FÜRNKRANZ, J.; FLACH, P. A. An analysis of rule evaluation metrics. In: **Proceedings of the 20th international conference on machine learning (ICML-03)**. [S.l.: s.n.], 2003. p. 202–209.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: MIT press, 2016.

HARBOLA, S.; COORS, V. One dimensional convolutional neural network architectures for wind prediction. **Energy Conversion and Management**, Elsevier, v. 195, p. 70–75, 2019.

HARTMANN, H. *et al.* Seasonal predictions of precipitation in the aksu-tarim river basin for improved water resources management. **Global and Planetary Change**, Elsevier, v. 147, p. 86–96, 2016.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning: data mining, inference, and prediction**. [S.l.]: Springer Science & Business Media, 2009.

- HAYKIN, S. **Neural networks: a comprehensive foundation**. [S.l.]: Prentice Hall PTR, 1994.
- HEMMATI-SARAPARDEH, A. *et al.* Toward generalized models for estimating molecular weights and acentric factors of pure chemical compounds. **international journal of hydrogen energy**, Elsevier, v. 43, n. 5, p. 2699–2717, 2018.
- HOSSEINZADEH, M. *et al.* A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things. **The Journal of Supercomputing**, Springer, p. 1–22, 2020.
- IANNACE, G.; CIABURRO, G.; TREMATERRA, A. Wind turbine noise prediction using random forest regression. **Machines**, Multidisciplinary Digital Publishing Institute, v. 7, n. 4, p. 69, 2019.
- IEA, I. E. A. "**Data & Statistics**". 2020. [online] Available at: <<https://www.iea.org/>>. [Accessed 21 Oct. 2020].
- JIANG, W. *et al.* Multilayer perceptron neural network for surface water extraction in landsat 8 oli satellite images. **Remote Sensing**, Multidisciplinary Digital Publishing Institute, v. 10, n. 5, p. 755, 2018.
- JUNIOR, W. J. A. **Métodos de Otimização Hiperparamétrica: Um Estudo Comparativo Utilizando Árvores de Decisão e Florestas Aleatórias na Classificação Binária**. 2018. Dissertação (Mestrado) - Universidade Federal de Minas Gerais.
- KIM, S. *et al.* Deeprain: ConvLstm network for precipitation prediction using multichannel radar data. **arXiv preprint arXiv:1711.02316**, 2017.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014.
- LIAKOS, K. G. *et al.* Machine learning in agriculture: A review. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 18, n. 8, p. 2674, 2018.
- LIAW, A.; WIENER, M. *et al.* Classification and regression by randomforest. **R news**, v. 2, n. 3, p. 18–22, 2002.
- LIMA, A. M. P. d. **Planejamento de sistemas hidrotérmicos empregando um modelo com sensibilidade ao risco de vertimento**. 2006. Dissertação (Mestrado) - Universidade de São Paulo.
- MAAS, A. L.; HANNUN, A. Y.; NG, A. Y. Rectifier nonlinearities improve neural network acoustic models. In: CITESEER. **Proc. icml**. [S.l.], 2013. v. 30, n. 1, p. 3.
- MARSLAND, S. **Machine learning: an algorithmic perspective**. [S.l.]: CRC press, 2015.
- MAXWELL, A. E.; WARNER, T. A.; FANG, F. Implementation of machine-learning classification in remote sensing: An applied review. **International Journal of Remote Sensing**, Taylor & Francis, v. 39, n. 9, p. 2784–2817, 2018.

- MOHANDES, M. A. *et al.* Support vector machines for wind speed prediction. **Renewable energy**, Elsevier, v. 29, n. 6, p. 939–947, 2004.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. **Foundations of machine learning**. [S.l.]: MIT press, 2018.
- MOSAVI, A.; OZTURK, P.; CHAU, K.-w. Flood prediction using machine learning models: Literature review. **Water**, Multidisciplinary Digital Publishing Institute, v. 10, n. 11, p. 1536, 2018.
- NAZARI-HERIS, M.; MOHAMMADI-IVATLOO, B.; GHAREHPETIAN, G. Short-term scheduling of hydro-based power plants considering application of heuristic algorithms: A comprehensive review. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 74, p. 116–129, 2017.
- NOSHAD, Z. *et al.* Fault detection in wireless sensor networks through the random forest classifier. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 19, n. 7, p. 1568, 2019.
- ONS, O. N. do S. "O Sistema em Números & Geração de Energia". 2021. [online] Disponível em: <<http://www.ons.org.br/>>. [Acessado 19 Janeiro. 2021].
- ORTIZ-GARCÍA, E.; SALCEDO-SANZ, S.; CASANOVA-MATEO, C. Accurate precipitation prediction with support vector classifiers: A study including novel predictive variables and observational data. **Atmospheric research**, Elsevier, v. 139, p. 128–136, 2014.
- PHITAKWINAI, S.; AUEPHANWIRIYAKUL, S.; THEERA-UMPON, N. Multilayer perceptron with cuckoo search in water level prediction for flood forecasting. In: IEEE. **2016 International Joint Conference on Neural Networks (IJCNN)**. [S.l.], 2016. p. 519–524.
- RAMACHANDRAN, P.; ZOPH, B.; LE, Q. V. Searching for activation functions. **arXiv preprint arXiv:1710.05941**, 2017.
- RAMOS, T. P. **Modelagem Híbrida para o Planejamento da Operação de Sistemas Hidrotérmicos Considerando as Não Linearidades das Usinas Hidráulicas**. Tese (Doutorado) — Universidade Federal de Juiz de Fora, 2015.
- RASOULI, K.; HSIEH, W. W.; CANNON, A. J. Daily streamflow forecasting by machine learning methods with weather and climate inputs. **Journal of Hydrology**, Elsevier, v. 414, p. 284–293, 2012.
- SARICA, A.; CERASA, A.; QUATTRONE, A. Random forest algorithm for the classification of neuroimaging data in alzheimer's disease: A systematic review. **Frontiers in aging neuroscience**, Frontiers, v. 9, p. 329, 2017.
- SATHYA, R.; ABRAHAM, A. Comparison of supervised and unsupervised learning algorithms for pattern classification. **International Journal of Advanced Research in Artificial Intelligence**, Citeseer, v. 2, n. 2, p. 34–38, 2013.
- SHAF, A. *et al.* Two classes classification using different optimizers in convolutional neural network. In: IEEE. **2018 IEEE 21st International Multi-Topic Conference (INMIC)**. [S.l.], 2018. p. 1–6.

- SHANG, Y. *et al.* A method of direct, real-time forecasting of downstream water levels via hydropower station reregulation: A case study from gezhouba hydropower plant, china. **Journal of Hydrology**, Elsevier, v. 573, p. 895–907, 2019.
- SHARMA, S. Activation functions in neural networks. **towards data science**, v. 6, 2017.
- SILVA, A. S. D. **Desenvolvimento e Implementação de Algoritmo Inteligente para Detecção e Localização de Faltas em Sistemas LCC-HVDC**. Tese (Doutorado) — Universidade Federal do ABC, 2019.
- SILVA, P. G. C. **Comportamento da Relação Evaporação-Vertimento-Liberação de Água Pelos Reservatórios Sob Diferentes Regras de Operação**. 2019. Dissertação (Mestrado) - Universidade Federal do Rio Grande do Norte.
- SINGH, D.; SINGH, B. Investigating the impact of data normalization on classification performance. **Applied Soft Computing**, Elsevier, v. 97, p. 105524, 2020.
- SINGH, V. K.; SINGAL, S. Operation of hydro power plants-a review. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 69, p. 610–619, 2017.
- SNIRH. **"Rede Hidrometeorológica Nacional"**. 2020. [online] Available at: <<https://www.ana.gov.br/ANA/>>. [Accessed 22 Jul. 2020].
- SUTTON, R. S.; BARTO, A. G. **Reinforcement learning: An introduction**. [S.l.]: MIT press, 2018.
- TALIB, A.; HASAN, Y. A. The application of artificial neural network for forecasting dam spillage events. 2010.
- TULYAKOV, S. *et al.* Review of classifier combination methods. **Machine learning in document analysis and recognition**, Springer, p. 361–386, 2008.
- VOYANT, C. *et al.* Machine learning methods for solar radiation forecasting: A review. **Renewable Energy**, Elsevier, v. 105, p. 569–582, 2017.
- WANG, S.-H. *et al.* Classification of alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. **Journal of medical systems**, Springer, v. 42, n. 5, p. 85, 2018.
- WANG, Y. *et al.* Assessing optimizer impact on dnn model sensitivity to adversarial examples. **IEEE Access**, IEEE, v. 7, p. 152766–152776, 2019.
- WESTIN, L. G. F. *et al.* Evaluating the impact of streamflow rating curve precision on firm energy of hydropower plants. **Water**, Multidisciplinary Digital Publishing Institute, v. 13, n. 8, p. 1016, 2021.
- YAO, S. *et al.* Deepsense: A unified deep learning framework for time-series mobile sensing data processing. In: **Proceedings of the 26th International Conference on World Wide Web**. [S.l.: s.n.], 2017. p. 351–360.
- YEUNG, D. S. *et al.* **Sensitivity analysis for neural networks**. [S.l.]: Springer, 2010.
- ZHOU, T. *et al.* Research on the long-term and short-term forecasts of navigable river's water-level fluctuation based on the adaptive multilayer perceptron. **Journal of Hydrology**, Elsevier, p. 125285, 2020.



25 a 28 de agosto de 2020

Santo André - SP (Virtual)

VIII Simpósio Brasileiro de Sistemas Elétricos

Certificado

O comitê organizador do SBSE gostaria de parabenizá-lo(a) pelo trabalho

123234 Spillage forecast in hydroelectric power plants via machine learning

de autoria de

Pedro Henrique Macedo Nascimento, Ramon Abritta, Frederico Ferreira Panoeiro, Leonardo de Mello Honório, André Marcato, Ivo Chaves Silva Junior

que teve destaque como o **melhor trabalho** dos apresentados na Sessão: SBSE 3.4 - Inteligência Computacional Aplicada a Sistemas Elétricos, do evento VIII Simpósio Brasileiro de Sistemas Elétricos - Virtual, realizado nos dias 25 a 28 de agosto de 2020.

Alfeu Joãosinho Sguarezzi Filho
(CECS_UFABC) - Coordenador Geral do VIII Simpósio Brasileiro de Sistemas Elétricos

Antônio Padilha Feltrin (CECS_UFABC)
- Coordenador Geral do VIII Simpósio Brasileiro de Sistemas Elétricos



Certification by Galoá



ANEXO A – Certificado