

Aprendizado de Máquina Aplicado à Previsão do Estado de Alarmes em Plataforma de Submarino

Rafael Soares Ferreira¹, Maurício Carvalho Mathias de Paulo¹ e Antonio Eduardo Carrilho da Cunha¹

¹Instituto Militar de Engenharia (IME), Rio de Janeiro/RJ - Brasil

Resumo—A previsão do estado de alarmes é um pilar crítico da manutenção preditiva em sistemas reais complexos. Este trabalho analisa o desempenho de modelos de aprendizado de máquina na previsão de estado de alarmes a partir de variáveis operacionais do sistema de controle e monitoramento de um submarino, sobre 161 dias de operação com 346 atributos. A metodologia avalia técnicas de extração e seleção de atributos aplicadas à previsão do estado de quatro alarmes em um período imediatamente após o índice, por *Random Forest* e redes neurais *Long Short-Term Memory* (LSTM), sob resoluções temporais de 15 segundos a 1 hora. Adotando o *F1-Score* como critério de mérito, os melhores modelos alcançaram valores entre 0,6365 e 0,9759. O *baseline* de persistência, contudo, supera os modelos treinados em dez das doze configurações, indicando que os índices elevados refletem a inércia temporal do rótulo. Os ganhos efetivos concentram-se no alarme de menor inércia.

Palavras-Chave—aprendizado de máquina, previsão de séries temporais, plataforma de submarino.

I. INTRODUÇÃO

A modelagem preditiva do estado de alarmes, aplicada à manutenção de meios navais, visa mitigar a indisponibilidade forçada e resguardar a integridade estrutural e operativa de sistemas críticos, como propulsão e navegação. Falhas não planejadas nesses vetores geram severas restrições táticas, além de vulnerabilidades de segurança física para a tripulação. Atualmente, os sistemas de supervisão e controle de bordo consolidam grandes volumes de dados de processo. Porém, operam de forma reativa, demandando a integração de algoritmos preditivos para otimização de sua eficácia [1]. O estabelecimento dessa abordagem pode reduzir a frequência da manutenção corretiva, um gargalo crítico que onera a cadeia logística e limita a projeção de poder das unidades navais em operações prolongadas.

O Sistema Integrado de Gerenciamento de Plataforma (do inglês, *Integrated Platform Management System*, IPMS) permite monitorar e controlar os subsistemas de um navio, como propulsão, sistemas de combate, ventilação, iluminação e outros, a partir de uma interface centralizada [2]. O IPMS utilizado pelas marinhas de guerra se diferencia dos sistemas de automação integrados (do inglês, *Integrated Automation Systems*, IAS) comerciais por incluir recursos específicos para controle de danos em combate, além de ser projetado para atender às necessidades específicas de cada navio e missão. No caso de um submarino, o sistema coleta dados de sensores nos diversos equipamentos e emite mensagens de alarme quando identifica situações anormais.

O *Machine Learning* (ML), um subconjunto da Inteligência Artificial, emergiu como o motor central dos sistemas de

previsão de falhas na Manutenção Preditiva (do inglês, *Predictive Maintenance*, PdM), superando a necessidade de modelos matemáticos complexos específicos para cada ativo [1], [3] e [4]. Os algoritmos de ML são capazes de aprender padrões e relações complexas entre as condições da máquina e o estado de falha, sendo que modelos baseados em dados podem gerar previsões sem a necessidade de modelagem física de cada subsistema.

Neste trabalho, investiga-se o desempenho de modelos de aprendizado de máquina associados a técnicas de engenharia de atributos para a detecção precoce do estado de alarmes em subsistemas de plataforma. A abordagem proposta utiliza algoritmos *Random Forest* (RF) e redes neurais *Long Short-Term Memory* (LSTM) sob variadas resoluções temporais, operando a partir de uma massa de dados restrita extraída de valores históricos utilizados na geração de relatórios operacionais de bordo. O escopo do estudo concentra-se em estabelecer o aspecto preditivo, prescindindo da instalação de instrumentação física dedicada ao monitoramento da integridade do ativo, substituindo-a por uma abordagem de inferência indireta fundamentada na fusão contextual de variáveis operacionais secundárias. Diante disso, a solução apresenta-se alinhada às restrições de engenharia militar e naval, pois demanda reduzido aporte de recursos e mitiga riscos operacionais ao garantir uma implementação não intrusiva, ou seja, sem a necessidade de intervenção ou re-homologação do sistema de controle consolidado.

O artigo está organizado da seguinte forma: a Seção II revisa os trabalhos relacionados; a Seção III descreve os materiais e métodos, o conjunto de dados e os modelos utilizados; a Seção IV apresenta os experimentos; a Seção V analisa os resultados; e a Seção VI apresenta a conclusão do trabalho.

II. TRABALHOS RELACIONADOS

Na literatura especializada, o prognóstico de estados de alarme e a análise de integridade de ativos são abordados sob o escopo da Manutenção Preditiva (*Predictive Maintenance*) e do Prognóstico de Falhas (*Failure/Fault Prediction*) [5], cujo objetivo central é antecipar falhas funcionais em sistemas críticos a partir de fluxos contínuos de dados sensoriais. Em [1] propõe-se uma arquitetura de fusão multinível na qual o *Random Forest* atua na seleção de atributos e em previsões preliminares; redes LSTM modelam as dependências temporais de longo prazo inerentes às variáveis físicas dos sensores, e o *XGBoost* consolida a classificação e a emissão definitiva do alerta de falha.

Sob uma perspectiva comparativa, em [6] avaliou-se de forma concorrente os modelos *Random Forest*, LSTM e *Support Vector Machines* (SVM) em sistemas de processamento

de alta pressão — cuja dinâmica operacional é análoga à de subsistemas hidráulicos e de atuação pneumática empregados em submarinos —, com ênfase no pré-processamento em larga escala e na engenharia de atributos, validando a eficácia dos classificadores na geração antecipada de alarmes a partir de séries amostradas em intervalos de 3 a 5 segundos.

Em síntese, a literatura converge para a associação entre a seleção de atributos baseada em árvores de decisão e o tratamento de dependências temporais por arquiteturas recorrentes. No presente estudo, estendeu-se a aplicação do algoritmo baseado em árvores de decisão para integrar, concomitantemente com a seleção de atributos, um mecanismo de janelamento temporal de dados pretéritos, estabelecendo um horizonte preditivo complementar em relação ao instante de referência, em paralelo com a predição via LSTM.

III. MATERIAIS E MÉTODOS

Nesta seção, apresenta-se a descrição do delineamento experimental e a estrutura do fluxo de tratamento de dados adotado. O objetivo central deste estudo consiste na análise comparativa de modelos preditivos operando com um vetor de 346 atributos, os quais representam variáveis de supervisão e controle da plataforma de um submarino convencional (e.g., leituras térmicas, correntes elétricas e estados de atuação de válvulas). Essas variáveis são aplicadas ao prognóstico de quatro alarmes distintos, selecionados a partir de um universo de mais de 750 instâncias homologadas no IPMS.

A avaliação dos modelos emprega um limiar de decisão fixo em 0,5, sem qualquer procedimento de otimização de limiar, de modo que o desempenho reportado não incorpora o ajuste realizado sobre a parcela reservada para os testes. O fluxo de trabalho do método é apresentado na Fig. 1.

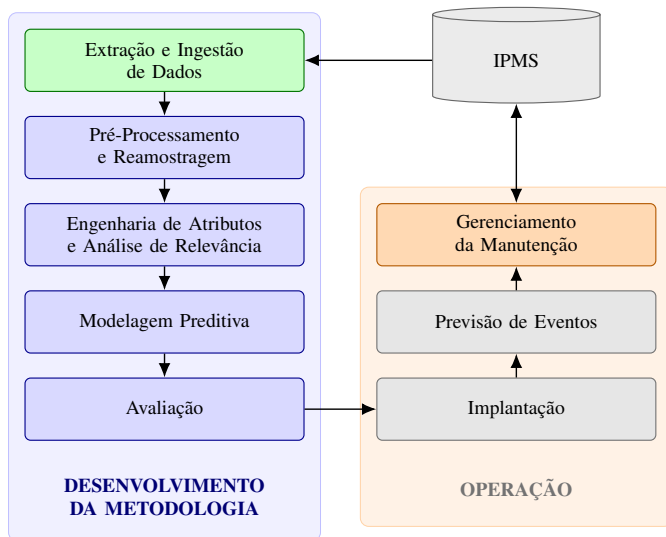


Fig. 1. Fluxograma do método.

A. Extração e Ingestão de Dados

A etapa inicial compreendeu a extração de séries temporais históricas armazenadas em um servidor de historização (*historization*).

Após a ingestão desses arquivos em um ambiente integrado de desenvolvimento, os dados foram padronizados em matrizes estruturadas (*dataframes*), nas quais cada vetor de

linha representa o registro cronológico (*timestamp*) de coleta. As ocorrências de alarmes foram representadas por atributos binários, utilizados para o treinamento supervisionado do modelo e para a produção de métricas de avaliação.

B. Pré-Processamento e Reamostragem

A fim de mitigar a assincronia original das leituras, realizou-se uma nova amostragem dos dados, garantindo a homogeneidade temporal do *dataset*. Para o vetor de atributos preditores, foram calculados os valores correspondentes a cada intervalo com base nas características intrínsecas dos atributos e alarmes.

Instantes em que não se tinha o registro de valor válido naquele momento exato foram tratados por meio do método de imputação de valores, como apresentado em [4], mantendo a consistência lógica da reamostragem. Períodos operacionais em que não havia um valor pretérito válido para ser imputado foram descartados da análise.

A seleção de alvos preditivos baseou-se no isolamento de parte dos alarmes mais frequentes no período amostrado, visando garantir um volume de dados estatisticamente robusto para o treinamento das arquiteturas de aprendizado de máquina. Desse grupo, filtraram-se os alarmes que manifestaram atividade contínua pelo maior intervalo de tempo.

Por fim, cumpre destacar que, por razões de segurança orgânica e salvaguarda de informações institucionais, as descrições originais e as *tags* técnicas das variáveis foram substituídas por pseudônimos, sem prejuízo da validade metodológica e da reprodutibilidade científica do experimento.

C. Engenharia de Atributos e Análise de Relevância

Na etapa subsequente, as séries temporais são redimensionadas em magnitudes distintas para a extração de descritores estatísticos no domínio do tempo.

Os dados resultantes foram submetidos, inicialmente, a uma análise de correlação linear por meio do coeficiente de Pearson, visando quantificar a dependência estatística direta entre os preditores gerados e a ativação dos alarmes, de forma similar ao que foi apresentado em [6]. Complementarmente, de forma a mapear interações complexas e interdependências não lineares negligenciadas pelo método linear, implementa-se uma avaliação do grau de relevância das variáveis por meio do cálculo do escore de importância dos atributos (*feature importance*), obtido via classificador *Random Forest*.

A partir dessas análises, define-se uma estratégia de filtragem para cada variável de alarme, estruturando três subconjuntos distintos de preditores para o treinamento dos modelos. O primeiro subconjunto isola os atributos que superam um limiar do coeficiente de correlação de Pearson definido para o experimento. O segundo subconjunto é composto pelas variáveis selecionadas por critério análogo aplicado aos escores de importância gerados pelo *Random Forest*. Por fim, o terceiro conjunto é constituído pela união dos dois grupos anteriores (*Union*), expurgando-se elementos duplicados de modo a mitigar redundâncias e os efeitos de multicolinearidade no pipeline preditivo.

D. Modelagem Preditiva

A partir dos subconjuntos de atributos estruturados nas etapas precedentes, visando introduzir a capacidade

de prognóstico temporal ao classificador *Random Forest*, implementou-se uma estratégia de janelamento que separa os últimos valores da série temporal para a previsão. Esta configuração permite que o algoritmo capture a assinatura dinâmica e a tendência evolutiva das variáveis operacionais ao longo de um intervalo de tempo contínuo em relação ao alarme avaliado.

De forma concomitante, as bases de dados reservadas para treino foram utilizadas no treinamento de uma arquitetura de redes neurais recorrentes do tipo LSTM. Para fins de pareamento e análise comparativa entre as abordagens de aprendizado estatístico e aprendizado profundo, a rede LSTM foi parametrizada com o mesmo horizonte de atraso temporal estabelecido no modelo anterior. Os modelos investigados são apresentados a seguir:

- **Random Forest:** É uma técnica supervisionada de *ensemble* que constrói múltiplas árvores de decisão independentes, cada uma treinada sobre subconjuntos aleatórios de amostras e de variáveis, consolidando as projeções por votação majoritária [6], [7]. Destaca-se pela capacidade de processar dados de alta dimensionalidade e pela menor propensão ao sobreajuste (*overfitting*) do que uma árvore isolada [1], [6]. Na manutenção preditiva, além da classificação da falha, o modelo é frequentemente empregado na seleção de atributos (*feature selection*), isolando os sinais operacionais mais críticos e fornecendo as primeiras previsões do estado do alarme [1].
- **Redes Neurais (LSTM):** É uma arquitetura recorrente de aprendizado profundo (*Deep Learning*) projetada para o processamento de entradas sequenciais e séries temporais [6], [8]. Sua célula retém informação por longos horizontes, permitindo identificar correlações distantes no tempo em conjuntos de dados complexos [1], [5]. Na manutenção preditiva, a rede é alimentada com o histórico contínuo dos sensores e encarregada de capturar os padrões temporais e as tendências de degradação, fornecendo prognósticos antecipados que subsidiam o agendamento das intervenções [6], [8].

E. Avaliação

A validação experimental das arquiteturas foi conduzida mediante o levantamento de métricas clássicas de desempenho: acurácia, precisão, sensibilidade (*recall*) e *F1-Score*. Considerando o contexto operacional de plataformas de defesa, buscou-se uma avaliação equilibrada, visando à confiabilidade de um eventual sistema de previsão. Portanto, adotou-se o *F1-Score* como métrica central de avaliação.

F1-Score é uma métrica que estabelece a média harmônica simples entre precisão e sensibilidade. Seu emprego se justifica quando há a necessidade de buscar o equilíbrio ótimo entre a capacidade do modelo de identificar corretamente as falhas (sensibilidade) e a confiabilidade de seus alertas, minimizando a ocorrência de alarmes falsos (precisão) [8]. Dessa forma, a métrica atribui pesos iguais a ambas as medidas de desempenho, sendo ideal para cenários em que falsos positivos e falsos negativos possuem relevância analítica equivalente.

Como referência de comparação, adotou-se uma *baseline* de persistência, que atribui ao índice subsequente o último estado observado do alarme ($\hat{y}(t+1) = y(t)$). Por não

envolver treinamento, esse procedimento quantifica a parcela do desempenho atribuível à inércia temporal do rótulo, de modo que um modelo apenas evidencie uma capacidade preditiva efetiva ao superá-lo.

IV. EXPERIMENTOS

O conjunto de dados foi particionado por corte temporal da série ordenada, sem embaralhamento de amostras, em segmentos contíguos e mutuamente exclusivos: um segmento inicial destinado ao ajuste dos modelos, um segmento intermediário de validação e um segmento final reservado ao teste, de modo que cada partição corresponda a um período de operação distinto e posterior ao anterior. Todas as etapas que antecedem o cálculo final das métricas utilizam exclusivamente os dados anteriores ao segmento de teste: os coeficientes de correlação de Pearson e os escores de importância do *Random Forest*, que definem os subconjuntos de preditores, são calculados apenas sobre o segmento inicial, e o critério de parada antecipada (*early stopping*) da rede LSTM considera apenas o segmento de validação. A parcela final da série é empregada unicamente no cálculo das previsões e métricas reportadas.

Devido a restrições de arquitetura de TI da plataforma, o subconjunto de dados extraídos não abrange a totalidade das variáveis do IPMS, correspondendo a um escopo selecionado de registros contínuos originalmente estruturados para alimentar o módulo gerador de relatórios operacionais de bordo — uma ferramenta de apoio à decisão da equipe de operação. Após a consolidação, o conjunto extraído totalizou 209.299.057 registros, distribuídos entre 754 tipos de alarmes e 346 atributos operacionais. Um exemplo dos dados tratados no experimento é apresentado na Fig. 2.

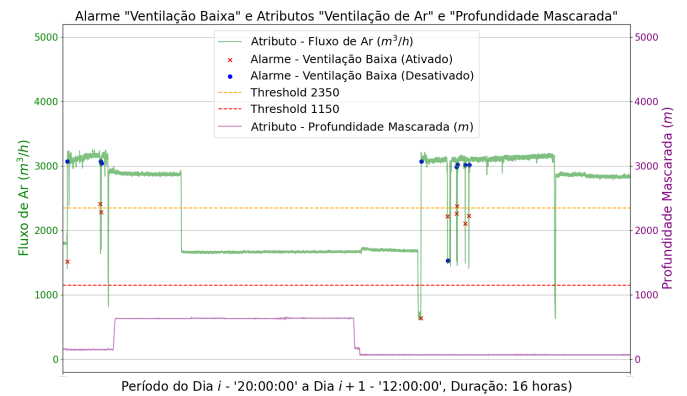


Fig. 2. Alarme de ventilação baixa e atributos de ventilação de ar e profundidade mascarada.

Os sensores são registrados em intervalos irregulares, portanto foi aplicada uma reamostragem para a frequência uniforme de 1 Hz, na qual se calculou a média aritmética das leituras no intervalo de 1 segundo. Em contrapartida, para as variáveis de estado (alarmes), aplicou-se o operador de valor máximo no intervalo, salvaguardando e privilegiando o estado ativo (sinalização de falha latente). A imputação de valores ausentes foi realizada com base no último valor válido (*forward fill*). Entradas de dados em que não se pôde aferir um valor válido foram removidas.

Foram extraídos os 10% dos alarmes mais frequentes durante o período de leitura de aproximadamente 161 dias.

Dentre estes, 20 foram selecionados por apresentarem atividade em um maior intervalo temporal. Finalmente, dentre os mais frequentes e ativos, foram amostradas quatro instâncias utilizadas neste experimento, apresentadas na Tabela I.

TABELA I
AMOSTRA DE ALARMES FREQUENTES E ATIVOS NO PERÍODO

Descrição	Frequência	Período de Ativ. em Dias
Motor Y - Falha de pressão	323	158
Motor Z - Falha de pressão	348	158
Motor X - Falha geral	7063	158
Ventilação baixa	1450	161

A partir das leituras desses quatro alarmes, geraram-se diferentes granularidades temporais: 15s, 1min (60s), 15min (900s), 30min (1800s) e 1h (3600s). A consolidação dos valores nesses intervalos seguiu critérios análogos aos da reamostragem inicial (média para atributos e valor máximo para alarmes).

Em seguida, foram extraídas as seguintes características das variáveis operacionais: valores máximo e mínimo, média aritmética, valor eficaz (*Root Mean Square* — RMS), desvio padrão, variância e a amplitude de variação (Δ).

Foram calculados os coeficientes de correlação de Pearson e os graus de importância em *Random Forest*, relacionados a cada um dos quatro alarmes selecionados nas diversas granularidades temporais. Nesta etapa, foram utilizados os valores das médias dos intervalos de atributos e valores máximos de alarmes.

A partir daí, utilizou-se um limiar estatístico definido pela média acrescida do desvio padrão dos valores calculados para a seleção dos atributos.

Um exemplo de atributos com maior relevância em ambos os métodos para os alarmes selecionados com reamostragem de 1s e granularidade amostral de 15s pode ser visto nas Tabelas II e III.

Para o modelo em *Random Forest*, foi aplicada uma defasagem unitária ($t - 1$), que correlaciona o estado imediatamente anterior ao evento de falha e uma abordagem de janelamento expandido que abrange 20 períodos pretéritos ($t - 20$ a $t - 1$), conforme mostrado na Fig. 3. De forma concomitante, submeteu-se a base de dados ao treinamento de uma arquitetura de redes neurais recorrentes do tipo LSTM com *lookback* de 20 leituras anteriores, conforme mostrado na Fig. 4.

As configurações dos modelos foram mantidas fixas em toda a grade de experimentos. O *Random Forest* empregou 100 árvores (1.000 na análise de relevância de atributos), semente aleatória fixa e ponderação de classes inversamente proporcional à frequência observada (*balanced*), adotada para compensar o desbalanceamento entre instantes com e sem alarme ativo. A rede LSTM combinou duas camadas recorrentes de 128 unidades e uma camada densa de saída única, otimizada por Adam com taxa de aprendizado de 5×10^{-4} , lotes de 1.024 amostras e limite de 50 épocas, com parada antecipada pela perda de validação (paciência de dez épocas e restauração dos melhores pesos). Nesse caso, o desbalanceamento foi tratado na função de perda pela entropia cruzada binária focal com balanceamento de classes ($\gamma = 2$ e α igual à proporção da classe majoritária no segmento de treinamento), que reduz a contribuição das amostras facilmente classificadas.

TABELA II
10 ATR. COEF. CORRELAÇÃO DE PEARSON > (MÉDIA + DESVIO PADRÃO)

Atributos (Motor Y - Falha de pressão)	Correlação
Válvula de Bloqueio de Saída Y Fechada	0,950761
Válvula de Bloqueio de Entrada Y Fechada	0,950761
Válvula de Bloqueio de Saída Z Fechada	0,949330
Válvula de Bloqueio de Entrada Z Fechada	0,949328
Válvula de Bloqueio de Entrada W Fechada	0,948647
Válvula de Bloqueio de Saída W Fechada	0,948646
Válvula de Bloqueio de Saída X Fechada	0,945638
Válvula de Bloqueio de Entrada X Fechada	0,945636
Válvula de Exaustão de Saída 2	0,861081
Válvula de Exaustão de Saída 4	0,861081
Atributos (Motor Z - Falha de pressão)	Correlação
Válvula de Bloqueio de Saída Z Fechada	0,918758
Válvula de Bloqueio de Entrada Z Fechada	0,918756
Válvula de Bloqueio de Entrada W Fechada	0,918301
Válvula de Bloqueio de Saída W Fechada	0,918300
Válvula de Bloqueio de Saída Y Fechada	0,916228
Válvula de Bloqueio de Entrada Y Fechada	0,916228
Válvula de Bloqueio de Saída X Fechada	0,912066
Válvula de Bloqueio de Entrada X Fechada	0,912064
Pressão da praça dos diesel	0,832628
Válvula de Exaustão de Saída 4	0,831314
Atributos (Motor X - Falha geral)	Correlação
Velocidade do motor X	0,571712
Pressão do óleo de lubrificação X	0,571700
Válvula de Bloqueio de Entrada X Aberta	0,565177
Válvula de Bloqueio de Saída X Aberta	0,565173
Medição da tensão de saída do retificador 1	0,564335
Pressão da água doce de saída X	0,563224
Temperatura do cilindro A6 X	0,561558
Maior temperatura de cilindro X	0,561010
Temperatura do cilindro A5 X	0,560154
Temperatura do cilindro A2 X	0,559963
Atributos (Ventilação Baixa)	Correlação
Fluxo de ventilação B	0,151972
Nível baixo do tanque de água do mar	0,087398
Compressor em funcionamento	0,084329
Temperatura do circuito de resfriamento	0,074574
Pressão do acumulador hidráulico de alta pressão	0,074070
Temperatura de entrada da água de resfriamento	0,073376
Temperatura da Seção 2 Conversor 1	0,071231
Nível do tanque de compensação A	0,068775
Nível do tanque de compensação B	0,067851
Temperatura da Seção 3 Conversor 1	0,067555

V. RESULTADOS E DISCUSSÃO

A análise de desempenho das arquiteturas foi conduzida sob diferentes horizontes temporais de predição. Para cada variável de alarme catalogada na Tabela I, os perfis de resposta foram mapeados graficamente em função das estratégias de janelamento dinâmico e da natureza dos subconjuntos de preditores selecionados. Os valores de *F1-Score* (eixo Y) agrupados por critério de seleção e dimensionalidade no eixo X para cada um dos métodos de predição utilizados (RF(1) - Entrada: Índice - 1, RF(20) - Entrada: Índice - 20 e LSTM(20) - rede neural recorrente com *lookback* de 20 entradas) encontram-se discriminados nas Figs. 5 para o alarme “Motor Y - Falha de pressão”, 6 para “Motor Z - Falha de pressão”, 7 para “Motor X - Falha geral” e 8 para “Ventilação baixa”.

Complementarmente, a Tabela IV sintetiza as configurações e os hiperparâmetros dos modelos que alcançaram desempenho máximo em termos de *F1-Score* no contexto do experimento, viabilizando a análise do impacto decorrente das modificações estruturais e da aplicação de modelos.

Os resultados experimentais revelam que os alarmes “Motor Y - Falha de pressão” e “Motor Z - Falha de pressão” —

TABELA III
10 ATR. RF DE GRAU DE IMPORTÂNCIA > (MÉDIA + DESVIO PADRÃO)

Atributos (Motor Y - Falha de pressão)	Grau Import.
Válvula de Bloqueio de Entrada Y Fechada	0,070090
Válvula de Bloqueio de Entrada W Fechada	0,056598
Válvula de Bloqueio de Saída W Fechada	0,053275
Válvula de Bloqueio de Saída Y Fechada	0,049466
Válvula de Bloqueio de Saída Z Fechada	0,048647
Válvula de Bloqueio de Entrada Z Fechada	0,046272
Válvula de Exaustão de Saída 2	0,044794
Válvula de Bloqueio de Saída X Fechada	0,041626
Válvula de Exaustão de Saída 4	0,037276
Pressão da praça dos diesel	0,034192
Atributos (Motor Z - Falha de pressão)	Grau Import.
Válvula de Bloqueio de Saída W Fechada	0,054404
Válvula de Bloqueio de Entrada W Fechada	0,053515
Válvula de Bloqueio de Entrada Y Fechada	0,052210
Válvula de Bloqueio de Entrada Z Fechada	0,050917
Válvula de Bloqueio de Saída Z Fechada	0,050668
Válvula de Exaustão de Saída 2	0,042419
Pressão do óleo de lubrificação Z	0,040719
Válvula de Bloqueio de Saída Y Fechada	0,038215
Válvula de Bloqueio de Saída X Fechada	0,036441
Válvula de Exaustão de Saída 4	0,032882
Atributos (Motor X - Falha geral)	Grau Import.
Válvula de Bloqueio de Entrada X Aberta	0,039510
Válvula de Bloqueio de Saída X Aberta	0,036195
Válvula de Bloqueio de Entrada W Aberta	0,035810
Válvula de Bloqueio de Saída Y Aberta	0,034254
Válvula de Bloqueio de Saída W Aberta	0,030614
Válvula de Bloqueio de Entrada Y Aberta	0,027968
Válvula de Bloqueio de Saída Z Aberta	0,025377
Válvula de Bloqueio de Entrada Z Aberta	0,020843
Válvula de Exaustão de Saída 3	0,019421
Temperatura do cilindro B3 X	0,017963
Atributos (Ventilação baixa)	Grau Import.
Fluxo de ventilação B	0,103118
Fluxo de ventilação A	0,055350
Pressão da água do mar W	0,039426
Pressão da água do mar Y	0,027833
Pressão da água do mar X	0,024218
Tensão média de célula 1A	0,019117
Menor tensão de 1B	0,018774
Menor tensão de 1A	0,017247
Menor nível de eletrólito 1A em mm	0,016823
Tensão média de célula 1B	0,016106

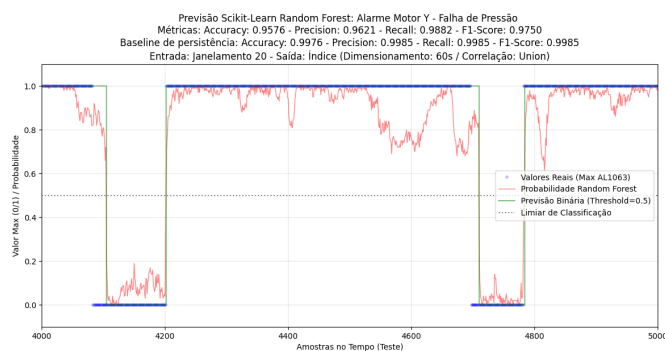


Fig. 3. RF(20) — alarme “Motor Y - Falha de pressão”, dimens.: 60.

associados a anomalias de pressão em uma mesma etapa operativa do motor — exibem elevada previsibilidade ponto a ponto, com *F1-Score* superior a 0,97 em todas as arquiteturas avaliadas. Esse desempenho, contudo, é explicado pela inércia temporal do rótulo: o *baseline* de persistência atinge 0,9985 e 0,9986 nesses alarmes, superando todos os modelos treinados, conforme a Tabela IV. Somada à baixa sensibilidade às variações na arquitetura dos modelos e na dimensionalidade dos dados, essa constatação indica que tais alarmes não discriminam arquiteturas e podem ser preteridos

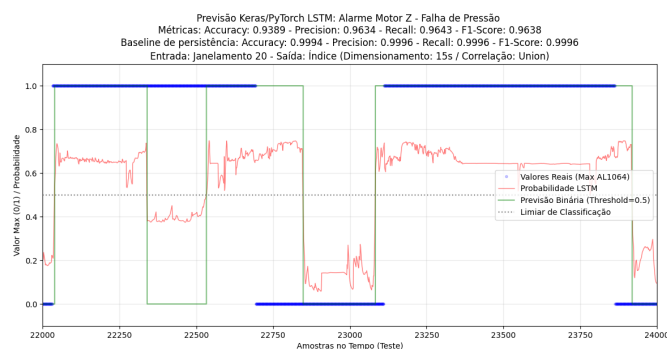


Fig. 4. LSTM(20) — alarme “Motor Z - Falha de pressão”, dimens.: 15.

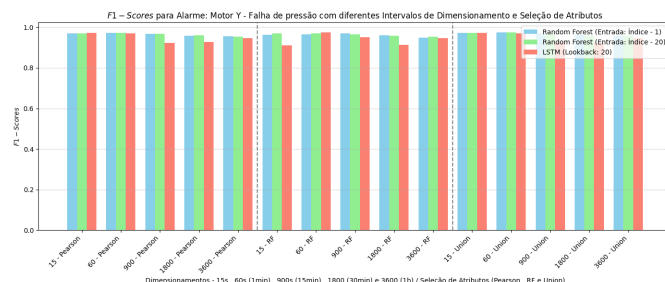


Fig. 5. *F1-Scores* para o alarme “Motor Y - Falha de pressão” com diferentes intervalos de dimensionamento e seleção de atributos.

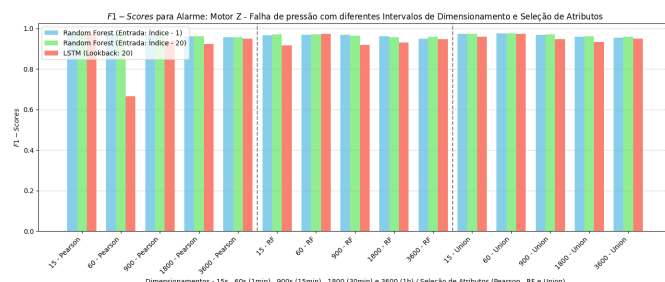


Fig. 6. *F1-Scores* para o alarme “Motor Z - Falha de pressão” com diferentes intervalos de dimensionamento e seleção de atributos.

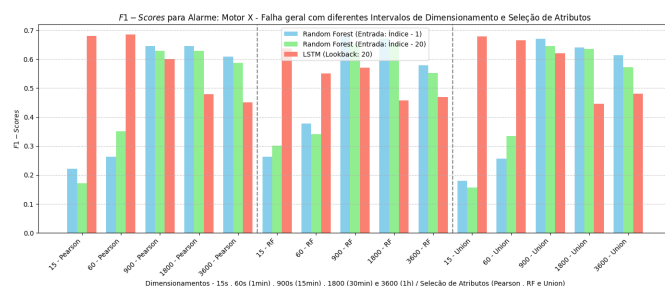


Fig. 7. *F1-Scores* para o alarme “Motor X - Falha geral” com diferentes intervalos de dimensionamento e seleção de atributos.

em análises subsequentes voltadas à otimização de hiper-parâmetros, priorizando-se cenários que contemplem alarmes de naturezas distintas e maior complexidade de modelagem.

Por outro lado, os alarmes “Motor X - Falha geral” e “Ventilação baixa” apresentaram maior dispersão nos resultados, evidenciando uma complexidade estocástica superior. No cenário de “Falha geral”, que agrega múltiplos eventos em uma única variável de saída, o melhor desempenho coube à LSTM(20), com *F1-Score* de 0,6853 na configuração (60 -

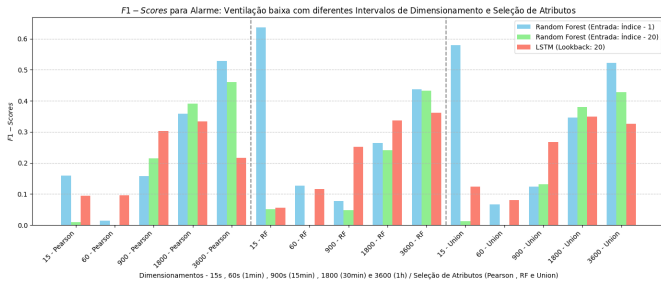


Fig. 8. *F1-Scores* para o alarme “Ventilação baixa” com diferentes intervalos de dimensionamento e seleção de atributos.

TABELA IV
MODELOS E ATRIBUTOS COM MAIOR *F1-SCORE* E SUA PERSISTÊNCIA

Descr. Alarme	<i>F1-Score</i> Máximo (<i>Baseline</i> Persistência) (Dimens. - Critério de Sel.)		
	RF(1) (60 - Union)	RF(20) (60 - Union)	LSTM(20) (60 - RF)
Motor Y Falha de pressão	0,9743 (0,9985)	0,9750 (0,9985)	0,9741 (0,9985)
Motor Z Falha de pressão	0,9759 (0,9986)	0,9758 (0,9986)	0,9742 (0,9996)
Motor X Falha geral	0,6778 (0,7829)	0,6514 (0,7348)	0,6853 (0,9001)
Ventilação baixa	0,6365 (0,6806)	0,4604 (0,2989)	0,3621 (0,2989)

Pears.); ainda assim, foi inferior ao *baseline* de persistência na mesma configuração (0,9001). As únicas configurações em que os modelos superam a persistência ocorrem no alarme “Ventilação baixa” sob dimensionamentos mais longos, em que RF(20) e LSTM(20) alcançam 0,4604 e 0,3621 contra 0,2989 do *baseline*, indicando que há sinal efetivamente aprendido justamente onde o rótulo é menos inercial, ainda que em patamar absoluto reduzido. Nesses cenários, o impacto da escolha do modelo preditivo e da estruturação dimensional dos dados mostrou-se relevante, o que consolida essas variáveis como instâncias de referência (*benchmarks*) para o ajuste fino e a validação de modelos aplicados a falhas de maior complexidade.

Os resultados experimentais evidenciam que a previsão do estado do alarme no índice subsequente é factível sob diferentes granularidades temporais, com *F1-Score* dos melhores modelos variando entre 0,6365 e 0,9759, conforme a Tabela IV. A comparação com o *baseline* de persistência delimita, entretanto, o alcance dessa constatação: em dez das doze configurações avaliadas, a simples repetição do último estado observado supera o modelo treinado, de modo que os valores absolutos de *F1-Score* refletem majoritariamente a inércia temporal dos alarmes e não a antecipação de transições de estado. Esse resultado reforça a pertinência da avaliação por episódio — restrita aos instantes em que o alarme não se encontra previamente ativo —, registrada como trabalho futuro.

VI. CONCLUSÃO

Este trabalho avaliou a previsão do estado de alarmes em um Sistema de Gerenciamento de Plataforma de submarino, a partir de variáveis operacionais e sem instrumentação dedicada. Foram conduzidos experimentos sobre uma grade de três dimensões — modelos de aprendizado de máquina,

resoluções de agrupamento temporal e critérios de seleção de atributos — aplicados a uma amostra de quatro alarmes. O conjunto de dados compreende 161 dias de operação, totalizando 209.299.057 registros distribuídos entre 754 tipos de alarmes e 346 atributos operacionais.

Nos alarmes de falha de pressão, todos os classificadores alcançaram *F1-Score* superior a 0,97 em todas as arquiteturas avaliadas. O *baseline* de persistência, entretanto, atinge valores ainda superiores nesses casos, o que indica que o desempenho decorre da inércia temporal do rótulo e não da antecipação do evento.

Nos alarmes de falha geral e ventilação baixa, o modelo, o agrupamento temporal e o critério de seleção produziram variação expressiva nas métricas. A comparação com a persistência, contudo, qualifica esse resultado: o *baseline* supera o modelo treinado em dez das doze configurações avaliadas, e os modelos só o superam em ventilação baixa sob os dimensionamentos mais longos — o alarme de rótulo menos inercial.

Esta pesquisa constitui a etapa inicial de um projeto mais amplo. Os desdobramentos previstos incluem a adoção de rotulagem por episódio, restrita às transições de alarme e com exclusão dos períodos em que o alarme já se encontra ativo; a extensão do horizonte de predição para além do período imediatamente subsequente; a inclusão de novos algoritmos de aprendizado de máquina; e a otimização de hiperparâmetros. Tais etapas visam avaliar em que medida a abordagem pode subsidiar estratégias de manutenção baseada em condição, bem como sua eventual expansão para os mais de 750 alarmes que integram o IPMS.

REFERÊNCIAS

- [1] S. Sharma, V. Vivek, A. K. Malviya, “Ai-enhanced predictive maintenance in intelligent systems for industries,” in *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET 2024)*. Indore: Institute of Electrical and Electronics Engineers Inc., 2024.
- [2] W. Kaikkonen, “Integrated platform management system on naval ships: Applicability study,” Dissertação (Mestrado em Engenharia Marinha), Satakunta University of Applied Sciences, Rauma, Finlândia, 2020. [Online]. Available: <https://www.theseus.fi/handle/10024/346853>
- [3] A. M. Abood, A. R. Nasser, H. Al-Khazraji, “Predictive maintenance of electromechanical systems based on enhanced generative adversarial neural network with convolutional neural network,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 12, no. 4, pp. 1704–1712, 2023.
- [4] W. Silva M. Capretz, “Assets predictive maintenance using convolutional neural networks,” in *20th IEEE/ACIS International Conference On Software Engineering, Artificial Intelligence, Networking And Parallel/Distributed Computing (SNPD)*. Toyama: IEEE, 2019, pp. 59–66.
- [5] S. Rigas, P. Tzouveli, S. Kollias, “An end-to-end deep learning framework for fault detection in marine machinery,” *Sensors*, vol. 24, no. 16, p. 5310, 2024.
- [6] S. Srisuwan S. Innet, “An investigation of machine learning algorithms for predictive maintenance in high pressure processing systems,” in *2024 21st International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2024, pp. 94–98.
- [7] S. Kaparathi D. Bumblauskas, “Designing predictive maintenance systems using decision tree-based machine learning techniques,” *International Journal of Quality & Reliability Management*, vol. 37, no. 4, pp. 659–686, 2020.
- [8] U. Lützen S. Beji, “Predictive maintenance for offshore wind turbines through deep learning and online clustering of unsupervised subsystems: a real-world implementation,” *Journal of Ocean Engineering and Marine Energy*, vol. 10, pp. 627–640, 2024.