

Manutenção Preditiva de Bombas de Injeção Utilizando *Deep Learning*

Joaquín Figueroa, Rubén Benites, Marcelo Martins

LabRisco – Laboratório de Análise, Avaliação e Gerenciamento de Risco

Escola Politécnica da Universidade de São Paulo – Departamento de Engenharia Naval e Oceânica

Enrique López Droguett

Universidad de Chile – Departamento de Ingeniería Mecánica

Renato Mendes

Petrobras

ABSTRACT

A injeção de água em um poço petrolífero aumenta a pressão no reservatório, evitando seu declínio acelerado na recuperação do petróleo. Uma falha na bomba de injeção pode comprometer a produção de petróleo durante o tempo que esta fique parada. Por isso, prever falhas com antecedência e adotar políticas de manutenções preditivas fará com que sua disponibilidade alcance valores superiores aos atuais. Utilizando técnicas de *Machine Learning* (ML) e *Deep Learning* (DL), é possível identificar a existência de padrões complexos nos parâmetros monitorados que fornecem informações relevantes a respeito do estado do equipamento. Este trabalho apresenta um modelo para a previsão do estado de operação de uma bomba de injeção de água utilizando redes neurais CNN+LSTM. Para a previsão de ocorrência de falhas com dois dias de antecedência, todas as métricas de desempenho atingem valores acima de 98,8%.

1. INTRODUÇÃO

Na indústria há um constante esforço em otimizar processos. A manutenção é um destes processos. Como alternativa aos paradigmas mais antigos de planejamento, a manutenção baseada na condição (CBM) [1] propõe fazer o monitoramento constante das máquinas e utilizar os dados obtidos para fazer uma estimativa do tempo até a próxima falha e planejar manutenções fundamentado nessa informação. Nos últimos anos, algoritmos de ML e DL têm sido utilizados com grande sucesso para analisar os dados e prever falhas [2], [3]. As aplicações mais comuns têm relação com detecção e previsão de falhas em elementos rotativos. Neste contexto, as redes *Convolutional Neural Networks* (CNN) e *Long Short-Term Memory Networks* (LSTM) (cujas bases teóricas são explicadas em [4]) são algoritmos de DL que têm se destacado bastante pelos seus resultados, estabilidade e múltiplas possibilidades de aplicação. Elas também têm sido utilizadas na área de confiabilidade [5], [6].

O objetivo deste trabalho é analisar o desempenho de um algoritmo CNN+LSTM para previsão de falhas no longo prazo em uma bomba de injeção de água em poço de petróleo. Este artigo contém cinco seções. A seção 2 descreve o estudo de caso e apresenta o modelo de previsão de falhas. A seção 3 mostra os resultados obtidos pelo algoritmo. Na seção 4 os resultados são analisados e discutidos. Finalmente, a seção 5 mostra as principais conclusões obtidas do trabalho realizado, juntamente com as possíveis melhorias e passos do trabalho.

2. DESCRIÇÃO

Nesta seção se descreve a base de dados utilizada e o modelo de previsão de estado de saúde.

2.1 Base de Dados

Os dados utilizados para a previsão do estado de operação provêm de uma bomba centrífuga e seu motor elétrico. O sistema é utilizado para injeção de água em poços de petróleo. Sensores são colocados no sistema para monitorá-lo, dos quais 15 são utilizados neste trabalho, para desenvolvimento do modelo de previsão. As variáveis utilizadas incluem vazão, pressão, vibração e temperatura. Elas têm taxas de

amostragem diferentes, portanto, os dados foram agrupados em janelas de 10 minutos, onde o valor associado é a média dentro da janela. Assim, se consegue padronizar os dados para eles serem usados no algoritmo, obtendo uma base de 90.965 dados. Cada dado tem uma etiqueta associada a um dos quatro estados de operação presentes na base. Os dados são organizados em janelas de tempo para treinar a rede. Cada intervalo de 4 horas corresponde a 24 tempos consecutivos contendo as quinze variáveis. A etiqueta da matriz é o estado da bomba num tempo futuro. A base de dados é dividida em três partes: 64% dela é utilizada para treinamento; 16% para validação; e o 20% restantes para testes.

2.2 Modelo de Previsão de Estado de Saúde

Para a previsão do estado de saúde, se utilizou uma rede neural com diversas camadas, sendo: duas camadas CNN: uma camada LSTM: e uma camada FFN. A rede recebe os dados de entrada e os classifica de acordo com o estado de saúde 48 horas a frente. A classificação feita pelo algoritmo é comparada com a etiqueta, que mostra a classificação verdadeira. O erro gerado pela discordância em algumas classificações é a função objetivo do algoritmo, a qual é minimizada no processo de treinamento pela atualização dos pesos. A configuração da rede é a seguinte: a primeira convolução pega a imagem de entrada com 64 filtros de 1x15. A segunda convolução se compõe de 32 filtros de 1x7. O produto dessa segunda convolução é transferido a uma camada LSTM de 50 unidades. Depois, a informação passa por uma camada FFN de 256 neurônios, para finalmente estabelecer a camada de saída com 4 neurônios, cada um deles representando um estado de funcionamento da bomba. Após treinamento, o modelo é testado. O processo é repetido cinco vezes, para obter resultados médios, pois o processo de treinamento tem fontes de aleatoriedade, como por exemplo, a inicialização dos pesos. Dado o fato que os dados estão desbalanceados, as diversas classificações que o modelo faz são mostradas numa matriz de confusão. Para o trabalho apresentado, se utilizam duas métricas derivadas da matriz de confusão: o *precision* e *recall*, descritas em [7].

3. RESULTADOS

Os resultados do modelo de previsão para 48 horas são apresentados na Tabela 1 e na Figura 1.

Tabela 1 Métricas de desempenho para modelo de previsão em 48 horas

	Normal	Incipiente	Degradada	Critica
Precision	99,98	99,81	100,00	98,83
Recall	99,90	99,93	100,00	99,96

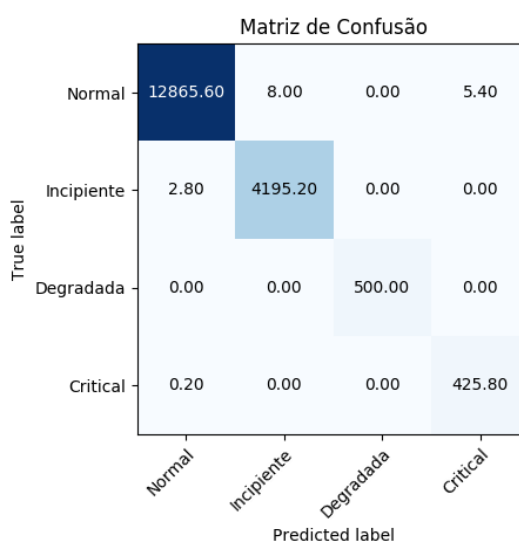


Figura 1 Matriz de confusão do modelo de previsão em 48 horas

Embora o modelo de previsão inicialmente tenha sido ajustado para prever o estado de saúde 48 horas a frente, ele é treinado para fazer previsões em diferentes intervalos, para analisar seu comportamento. Esses resultados são apresentados na Figura 2.



Figura 2 Evolução das medidas de desempenho para previsões em tempos diferentes

4. DISCUSSÃO

Os resultados apresentam as métricas de *precision* e *recall* para todos os estados. Dependendo do estado, uma métrica pode ser mais importante do que a outra. Para o estado de operação normal, é importante que a predição do modelo seja correta, pois um erro implica que o modelo não detectou uma falha. Portanto, a taxa de falsos positivos tem que ser baixa. Em consequência, o *precision* é a métrica mais importante. O contrário é o estado de falha crítica. Nesse caso, o *recall* tem mais importância, pois ele mede quantas instancias de falha crítica são efetivamente detectadas como tal pelo algoritmo. Um erro desse tipo (falso negativo) implica que uma instancia de falha crítica está sendo nomeada como uma instancia de menor importância. Nos casos restantes, as duas métricas têm igual importância.

Para a previsão em 48 horas, pode-se ver na Tabela 1 que, salvo *precision* para a detecção da falha crítica, todas as métricas apresentam valores acima de 0,9980. Isto significa que o algoritmo atinge resultados confiáveis. Isto é consistente com o que os resultados não vêm de só uma execução, mas mostram a média obtida de cinco execuções. Outro ponto de destaque na Tabela 1 tem relação com o estado de falha degradada. Ela apresenta 1,00 de *precision* e *recall*. Isto é importante, pois a detecção precoce de uma falha degradada influi na prevenção de ocorrência de falhas críticas, já que o processo de degradação da bomba faz com que antes de um estado de falha crítica, com alta probabilidade, haja uma falha degradada.

Na Figura 1 pode-se ver a matriz de confusão obtida após as cinco execuções. Nela, pode-se apreciar a grande diferença entre a quantidade de dados das diferentes classes. Além disso, se mostra que os erros mais comuns têm relação com instancias reais de operação normal. Dentro dos possíveis erros, este é o melhor, pois é mais seguro um falso alarme do que uma falha não detectada.

Para a previsão em outros tempos futuros, a Figura 2 mostra que os resultados são contra intuitivos, pois é de esperar que eles sofram uma queda na medida em que a previsão se faz mais para frente. A razão é porque o modelo foi ajustado para um desempenho ótimo em 48 horas, portanto, para cenários diferentes, o

modelo pode não ser ótimo. Por isso, os resultados não seguem a tendência esperada. Apesar disso, os resultados não têm muita diferença e continuam mostrando que o modelo permanece confiável. Para o estado normal, o *precision* atinge valores sempre acima de 0,9996. Isto significa que o modelo tem alta certeza quando diz que a bomba está em estado normal de operação, e que quase não haverá falhas não detectadas. No caso do *recall*, ele permanece acima de 0,9990, o que mostra um alto nível, independente das horas nas quais se faz a previsão.

Para o estado de falha incipiente, ainda as duas métricas apresentam valores acima de 0,9980. Para as falhas degradadas, salvo o *precision* quando a previsão é feita em 3 e 24 horas, as métricas atingem o valor de 1,00. Este resultado mostra que o algoritmo consegue distinguir com muita facilidade a física por trás da falha degradada. Para o estado de falha crítica, o *precision* sofre uma queda considerável com relação aos demais estados. Isto é devido a menor presença de instâncias com essa falha e dos diversos mecanismos que podem levar a ela. No caso da previsão em 72 horas, acredita-se que o *precision* teria melhor resultado se a rede fosse treinada com janelas de tempo maiores do que 6 horas. Para a previsão feita em 3 horas, provavelmente o modelo está “sobreajustado”, o que poderia ser solucionado treinando com janelas de tempo menores. O *recall* atinge o valor máximo quando a previsão é em 48 horas, o que faz sentido ao reparar que o modelo foi ajustado com o objetivo de prever falhas em 48 horas. O valor de 0,9996 dá um elevado grau de confiança e a certeza que uma falha crítica não passará despercebida.

5. CONCLUSÃO

Neste trabalho se apresenta um modelo de previsão do estado de funcionamento de uma bomba centrífuga de injeção de água num poço petrolífero. O modelo é composto por uma rede CNN-LSTM treinada para utilizar a informação de 6 horas de dados para prever o estado de funcionamento em 48 horas.

O modelo mostrou ter a capacidade de discernir entre os quatro estados de operação (normal, falha incipiente, falha degradada e falha crítica) com níveis confiáveis de desempenho. Para 48 horas, o *precision* atinge valores acima de 0,9980, enquanto que o *recall* fica sempre acima de 0,9990. Inclusive, as duas métricas atingem o valor de 1,00 para a detecção de falhas degradadas.

O desempenho do modelo é comparado com situações em que a previsão é feita em tempos futuros diferentes. Descobre-se que, mesmo o modelo sendo ajustado para obter o melhor desempenho para previsões em 48 horas, seus resultados atingem elevado nível de acerto para previsões feitas em 3, 24 e 72 horas.

Os resultados mostram que os algoritmos de *Deep Learning* podem ser utilizados para prevenção de falhas. Eles integram informação de diversas variáveis e são capazes de atingir resultados que dão confiança para sua utilização. Os próximos passos da pesquisa são: desenvolver um prognóstico de vida útil remanescente (RUL); e incluir a informação sobre os modos de falhas nos algoritmos, pois essa informação é relevante e ainda não está sendo aproveitada.

6. REFERÊNCIAS

- [1] J. Lee, R. Abujamra, A. K. S. Jardine, D. Lin, and D. Banjevic, “An integrated platform for diagnostics, prognostics and maintenance optimization,” in *Proceedings of the Intelligent Maintenance Systems*, 2004, pp. 15–27.
- [2] J. Ben Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, “Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network,” *Mech. Syst. Signal Process.*, vol. 56, pp. 150–172, 2015.
- [3] H. Hu, B. Tang, X. Gong, W. Wei, and H. Wang, “Intelligent Fault Diagnosis of the High-Speed Train With Big Data Based on Deep Neural Networks,” *IEEE Trans. Ind. Informatics*, vol. 13, no. 4, pp. 2106–2116, 2017.
- [4] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, 27-May-2015.
- [5] Z. Q. Chen, C. Li, and R. V. Sanchez, “Gearbox Fault Identification and Classification with Convolutional Neural Networks,” *Shock Vib.*, vol. 2015, 2015.
- [6] M. Yuan, Y. Wu, and L. Lin, “Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network,” *AUS 2016 - 2016 IEEE/CSAA Int. Conf. Aircr. Util. Syst.*, pp. 135–140,

- 2016.
- [7] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, 2009.