

Redes neurais artificiais aplicadas à moagem de minério de ferro combinadas a modelos empíricos

Artificial neural networks applied to iron ore grinding process combined with empirical models

Redes neuronales artificiales aplicadas a la molienda de hierro combinadas a modelos empíricos

Recebido: 30/06/2022 | Revisado: 19/07/2022 | Aceitado: 22/09/2022 | Publicado: 02/10/2022

Daniel Henrique Cordeiro Silva

ORCID: <https://orcid.org/0000-0001-8981-3953>

Universidade Federal de Ouro Preto, Brasil

E-mail: daniel.hcs@aluno.ufop.edu.br

Vladmir Kronemberger Alves

ORCID: <https://orcid.org/0000-0001-7320-3384>

Universidade Federal de Ouro Preto, Brasil

E-mail: vladmir.alves@ufop.edu.br

Ernandes Savio

ORCID: <https://orcid.org/0000-0002-0455-6570>

Instituto Tecnológico Vale, Brasil

E-mail: ernandes.savio@vale.com

Resumo

A cada dia, com o advento da Indústria 4.0, novas tecnologias são disponibilizadas e aplicadas ao Tratamento de Minérios, impulsionadas pela crescente disponibilidade de dados de chão de fábrica. Algumas das oportunidades a serem exploradas estão atreladas à utilização das ferramentas de Big Data, Advanced Analytics, Machine Learning e Inteligência Artificial. No beneficiamento de minério de ferro, nos processos de moagem, os ganhos potenciais oriundos de ferramentas de aprendizado de máquina tendem a ser ampliados quando combinados com modelos matemáticos consagrados, oriundos do conhecimento do processo, sejam eles empíricos ou fenomenológicos. Este artigo apresenta a aplicação de redes neurais artificiais para a predição da granulometria no produto de uma planta de moagem, principal parâmetro de qualidade, combinadas a modelos empíricos que descrevem o processo, visando estabelecer se tais equações podem contribuir para uma melhor performance dos modelos preditivos, de forma a suportar uma melhor e mais rápida tomada de decisão pelo operador da unidade.

Palavras-chave: Redes neurais artificiais; Aprendizado de máquina; Tratamento de minérios; Moagem; Controle de processos.

Abstract

Step by step, the technologies provided by Industry 4.0 are being inserted in the mining processes and one of the opportunities to be explored is the use of Big Data and Advanced Analytics tools. In iron ore beneficiation plants, in the milling processes, the potential gains from machine learning tools tend to be amplified when combined with mathematical models derived from process knowledge, whether empirical or phenomenological. This article presents the application of artificial neural networks for the prediction of the main product quality parameter of a milling plant, combined with empirical equations that describe the milling process, to establish whether such equations can contribute to a better performance of the predictive models.

Keywords: Artificial neural networks; Machine learning; Mineral processing; Iron ore; Grinding mill; Process control.

Resumen

Cada día, las tecnologías proporcionadas por la Industria 4.0 se están insertando en el Tratamiento de Minerales y una de las oportunidades a explorar es el uso de herramientas de Big Data y Analítica Avanzada. En el beneficio del mineral de hierro, en los procesos de molienda, las ganancias potenciales de las herramientas de aprendizaje automático tienden a amplificarse cuando se combinan con modelos matemáticos derivados del conocimiento del proceso, ya sea empírico o fenomenológico. Este artículo presenta la aplicación de redes neuronales artificiales para la predicción del principal parámetro de calidad del producto de una planta de molienda, combinado con ecuaciones empíricas que describen el proceso de molienda, con el fin de establecer como pueden contribuir a un mejor desempeño de los modelos predictivos.

Palabras clave: Redes neuronales artificiales; Aprendizaje automático; Tratamiento de minerales; Mena de hierro; Molienda; Control de procesos.

1. Introdução

Mesmo com seguidas crises econômicas, a demanda por minério de ferro tem aumentado mundialmente nos últimos anos, tornando vital a necessidade de melhorias significativas no desempenho das instalações de beneficiamento mineral. A otimização dos processos possibilita o aumento de capacidade de produção mais rapidamente do que a implantação de novas instalações, além de requisitar menores investimentos. A melhoria de desempenho reflete positivamente também em questões ambientais, principalmente quando os benefícios resultam em maior recuperação mássica ou metalúrgica das plantas: permite melhor utilização dos recursos naturais já em processo de lavra, reduzindo assim a quantidade de rejeitos dispostos em barragens, e ainda posterga impactos ambientais na abertura de novas minas (Souza, 2014).

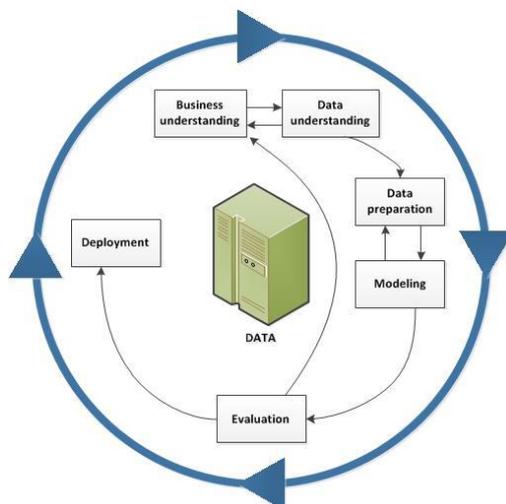
Na análise e simulação de processos de beneficiamento mineral, os modelos matemáticos mais utilizados são os empíricos, que normalmente são conjuntos de equações obtidas experimentalmente, e os fenomenológicos, que são equações algébricas e diferenciais que utilizam princípios físicos, químicos ou de engenharia e requerem calibração para sua utilização (Souza, 2014). O artigo apresenta um modelo desenvolvido a partir das novas tecnologias disponíveis, somadas a equações consagradas. Mais especificamente, uma combinação de técnicas de aprendizado de máquina com modelos que definem o consumo energético dos processos de moagem.

O modelo obtido é aplicado a uma planta de beneficiamento de minério de ferro, localizada no chamado Quadrilátero Ferrífero, em Minas Gerais, Brasil, e será utilizado na predição do percentual retido em malha de 0,15mm no produto de uma planta de moagem. O trabalho em questão faz parte da pesquisa de mestrado do autor, a cuja dissertação será defendida no início de 2023 e a escolha do processo de moagem ocorreu devido ao fato de ser uma das fases mais dispendiosas dentre as operações de beneficiamento mineral. Como descrito por Delboni (2007), os custos de operação e de capital nos processos de britagem e moagem são determinados pela energia necessária neste processo e os esforços mecânicos aos quais os equipamentos são submetidos.

2. Metodologia

De acordo com Mariscal (2010), o Processo Padrão Inter-Indústrias para Mineração de Dados, ou CRISP-DM (acrônimo para Cross-Industry Standard Process for Data Mining) é a metodologia mais difundida para modelagem em ciência de dados e é apresentada na Figura 1 abaixo.

Figura 1 - Processo CRISP-DM.

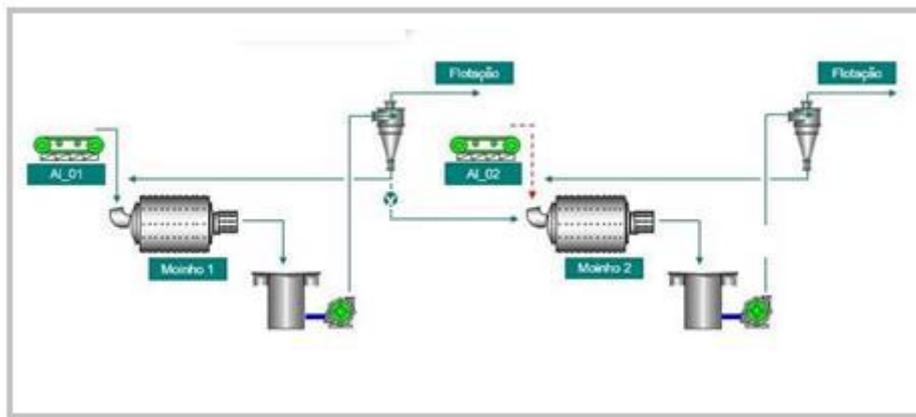


Fonte: IBM.

2.1 Entendimento do Negócio

Na fase inicial, busca-se entender o contexto do problema a ser abordado e definir os objetivos do projeto. Neste caso, foi analisado o fluxo do processo de moagem em questão. Conforme apresentado na Figura 2, a planta estudada possui dois moinhos, onde o primeiro recebe a alimentação “nova” do alimentador AL-01, executa a primeira fase de moagem e tem o seu produto classificado por uma primeira bateria de ciclones, que envia o material grosseiro (underflow) para o segundo moinho, cujo produto é novamente classificado por uma segunda bateria de ciclones. O material fino classificado em ambos os ciclones é encaminhado como produto da moagem.

Figura 2 - Processo de Moagem.



Fonte: Autores.

Alternativamente, é possível trabalhar com os dois moinhos em paralelo, direcionando o underflow do primeiro ciclone de volta ao primeiro moinho e fazendo o segundo moinho trabalhar de forma independente e com alimentação “nova”, oriunda do Alimentador AL-02.

2.2 Entendimento dos Dados

Esta fase consiste em verificar o universo de informações disponíveis, sua origem e formato, bem como coletar, organizar e documentar todos os dados que se encontram disponíveis para realizar a análise exploratória.

No problema em questão, os dados são obtidos do sistema historiador de dados de processo PIMS (Plant Information Management System). A Tabela 1 apresenta o conjunto das informações de processo disponíveis para a primeira análise. A base de dados disponibilizada possui as informações de média horária para as 34 variáveis de processo listadas (ou *features*), para um período total de quase 4 anos, totalizando 34559 amostras.

Tabela 1 - Features planta de moagem.

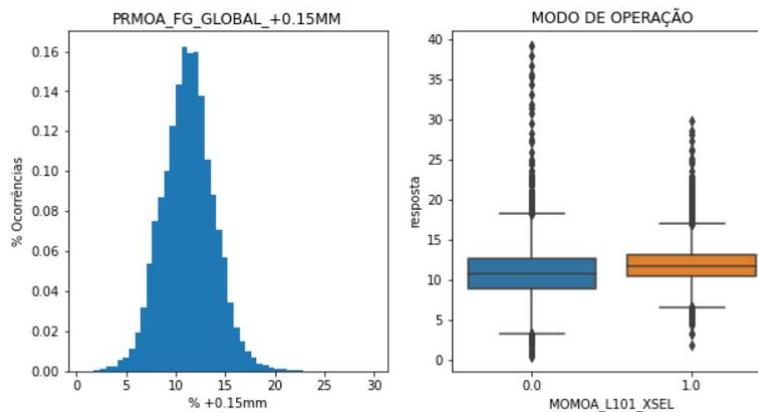
Feature	Descrição
ALMOA_L101_WI	Alimentador de correia_Vazão mássica (BIMOA_L101/ WitMOA_L1501)
ALMOA_L102_WI	Alimentador de correia_Vazão mássica (BIMOA_L102/ WitMOA_L1502)
BPMOA_L101_DI	Bomba de polpa_Densidade no recalque (DIT-MOA_L1-001)
BPMOA_L101_FI	Bomba de polpa_Vazão no recalque (FIT-MOA_L1-001)
BPMOA_L101_WI	Bomba de polpa_Vazão mássica no recalque (FY-MOA_L1-001B)
BPMOA_L102_DI	Bomba de polpa_Densidade no recalque (DIT-MOA_L1-002)
BPMOA_L102_FI	Bomba de polpa_Vazão no recalque (FIT-MOA_L1-002)
BPMOA_L102_WI	Bomba de polpa_Vazão mássica no recalque (FY-MOA_L1-002B)
BPMOA_L107_DI	Bomba de polpa_Densidade no recalque (DIT-MOA_L1-007)
CIMOA_L101_CONT1	Hidrociclone_Número Válvulas com Estado Aberto
CIMOA_L101_PI	Hidrociclone_Pressão (PIT-MOA_L1-501)
CIMOA_L102_CONT1	Hidrociclone_Número Válvulas com Estado Aberto
CIMOA_L102_PI	Hidrociclone_Pressão (PIT-MOA_L1-502)
CIMOA_L107_PI	Hidrociclone_Pressão (PIT-MOA_L1-507)
CXMOA_L101_LI	Caixa de polpa_Nível (LIT-MOA_L1-009)
CXMOA_L110_FI	Usina_Deslamagem Linha 1_Caixa de polpa_Vazão de soda (FIT-MOA_L1-018)
CXMOA_L110_LI	Usina_Deslamagem Linha 1_Caixa de polpa_Nível (LIT-MOA_L1-018)
MOMOA_L101_XSEL	Seleção Modo Linha 01 como Circuito Série (1 -Série; 0- Paralelo)
MOMOA_L101_FI9	Moinho (FIT-MOA_L1-011)_Vazão de água
MOMOA_L101_JAI1	Motor principal do Moinho_Potência
MOMOA_L101_JSP1	Moinho_Setpoint proporção água/minério
MOMOA_L102_FI9	Moinho_Vazão de água (FIT-MOA_L1-012)
MOMOA_L102_JAI1	Moinho_Motor Principal_Potência
MOMOA_L102_JSP1	Moinho_Setpoint proporção água/minério
MOMOA_L101_KDONDA	K de Donda para MO-MOA_L1-01 pelo APC
MOMOA_L102_KDONDA	K de Donda para MO-MOA_L1-02 pelo APC
ALMOAFG_GLOBAL_+0.15MM	+0.15MM do Aliment. Moagem
ALMOAFG_GLOBAL_+12.5MM	+12.5MM do Aliment. Moagem
ALMOAFG_GLOBAL_+6.3MM	+6.3MM do Aliment. Moagem
ALMOAFG_GLOBAL_-0.045MM	-0.045MM do Aliment. Moagem
ALMOAFG_GLOBAL_-1+0.15MM	-1+0.15MM do Aliment. Moagem
ALMOAQQ_GLOBAL_SIO2	Sílica do Aliment. Moagem
PRMOA_FG_GLOBAL_+0.15MM	+0.15MM Produto Moagem L1

Fonte: Autores.

Nesta fase exploratória, as diversas variáveis foram analisadas estatisticamente e visualmente através de histogramas e gráficos *box plot*. Tal análise é a base do entendimento da distribuição e qualidade dos dados, conhecimento fundamental para a fase de preparação dos dados, onde os dados serão tratados e são selecionadas as *features* que realmente serão utilizadas.

Conforme mostra a Figura 3, pode-se observar como a variável “PRMOA_FG_GLOBAL_+0.15MM” se distribui para a situação de operação com moinhos em série (moagem e remoagem), “MOMOA_L101_XSEL” = 1, a para a situação de moinhos em paralelo (duas moagens simples).

Figura 3 - Processo de Moagem.



Fonte: Autores.

Pode-se observar, em meio a diversas variáveis de processo, como vazões, pressões, densidades, análises de laboratório, etc., a presença da constante K de Donda, cujo cálculo é apresentado na revisão bibliográfica a seguir. A presença desta variável nos remete ao uso de um modelo consagrado, que será combinado com métodos de Machine Learning, conforme objetiva este artigo.

2.3 Preparação dos Dados

É nesta fase em que é criada a base de dados final. Ou seja, é feita a escolha dos dados que serão trabalhados, de como eles serão cruzados para resolver o problema da empresa, a preparação das databases e a definição do formato que será necessário para a análise.

Aqui é realizada a mineração de dados, considerando fatores como a relevância, a qualidade e as restrições técnicas, como limites no volume ou tipo de dados. Além disso, é preciso definir métodos de avaliação de desempenho para que seja realizada nas próximas fases.

Neste trabalho, foram seguidos os seguintes passos nesta fase:

- Tratamento de dados faltantes: Dados faltantes são representados pelo `NaN`. Quando isso ocorreu para uma dada *features* durante esta preparação, o referido registro teve o valor da *feature* substituído pela média. *Features* com grande número de dados faltantes foram excluídas
- Outliers individuais são substituídos pelo terceiro quartil da distribuição normal equivalente.
- Higienização pela qualidade: retiradas *features* com baixa variação ou excessivo número de outliers.
- Seleção de variáveis linearmente independentes: através do uso de uma da matriz de correlação cruzada, foram excluídas *features* com alta correlação com outras variáveis
- Normalização dos dados pelo desvio padrão.

Ao final deste processo, a base a ser utilizada para treinamento do modelo possui agora apenas 17 *features* e pouco mais de 28 mil amostras. Neste momento, os dados são então divididos:

- Entradas, ou variáveis independentes (as *features* propriamente ditas)
- Saída, ou resposta, que é a variável o *target* do nosso modelo, “PRMOA_FG_GLOBAL_+0.15MM”

Na sequência, os dados são novamente divididos, desta vez em base de treinamento, com 75% das amostras, e base de teste, com 25%.

2.4 Modelagem

Aqui são selecionadas e aplicadas as técnicas de *Data Mining* mais apropriadas, com base nos objetivos identificados na primeira fase. Assim como feito por Oliveira (2000) e Serpa (2019), dentre outros, neste trabalho serão utilizados modelos baseados em redes neurais artificiais, conceitos introduzidos por McCulloch (1943) e Minsky e Papert (1969), conforme será apresentado mais adiante neste artigo.

A abordagem se baseará em três fases: treinamento, validação e teste. Cada etapa, contando com um subconjunto de dados de cada um dos três principais conjuntos de dados. Este processo é descrito na revisão bibliográfica a ser apresentada no Item 3.2 deste artigo.

Em geral, na primeira etapa, os parâmetros do modelo são ajustados usando um algoritmo de otimização que visa minimizar a medição de erros, usando a diferença entre a produção estimada e a esperada, também chamada de técnica de aprendizado supervisionado.

O segundo passo é aplicado para ajustar os hiperparâmetros, a fim de superar comportamentos tendenciosos ou superajustados. É seguido por uma comparação entre os resultados de cada tentativa e uma seleção dos que melhor se ajustam aos dados. Como alternativa, é possível mesclar os subconjuntos de treinamento e validação, chamados *K-Folds*, para usar a técnica de validação cruzada, escolhendo os pesos ideais entre os modelos treinados.

Por fim, o desempenho do modelo treinado é avaliado com dados invisíveis, que simulam um contexto real com um conjunto de dados menor que segue a mesma distribuição de probabilidade da fase de treinamento.

Os detalhes da construção do modelo com redes neurais são apresentados na revisão bibliográfica a seguir.

2.5 Avaliação dos Modelos

Nessa fase, é preciso avaliar os resultados e levantar todas as possibilidades de variações que os dados possam ter, analisando se existe algum fator que tenha sido negligenciado e até que ponto o modelo atende aos objetivos de negócios. Caso o modelo não apresente um resultado satisfatório, devemos retomar a primeira etapa para entender melhor o negócio e os dados, ciclo que se repetiu algumas vezes neste trabalho.

Foi considerada como métrica de performance dos modelos a raiz do erro médio quadrático, também conhecida como RMSE (*Root Mean Square Error*), conforme apresentado no guia de Machine Learning do Google (2022). Conforme citado por Pessoa (2021), o erro quadrático médio é importante porque é uma maneira melhor de verificar a precisão do modelo.

2.6 Implementação

Após validar e testar um modelo, caso as métricas de desempenho atinjam patamares que realmente tenham potencial negócio, são implementadas soluções para que as previsões sejam executadas de forma contínua para que os modelos possam realmente agregar valor ao processo produtivo.

No caso deste trabalho, esta solução poderia ser uma arquitetura simples que faria a leitura dos dados de entrada, diretamente do sistema PIMS, executaria o script Python com o modelo, e retornaria a previsão em forma de escrita no próprio históricos. Todavia, como se trata de uma prova de conceito, a solução, no momento da escrita deste artigo, ainda não fora implementada.

3. Revisão Bibliográfica

3.1 Processo de Moagem

Durante muitos anos, os estudos sobre os processos de cominuição priorizavam a ótica da energia consumida, relacionando esta grandeza ao grau de redução proporcionado. Geralmente, o grau de redução é expresso pela curva de

distribuição granulométrica e observou-se experimentalmente que a relação entre a variação de diâmetro e a energia consumida é inversamente proporcional, ou seja:

$$dE = -K dx/x^n, \quad (1)$$

E : representa a energia aplicada a uma massa unitária de minério.

x : representa o diâmetro das partículas.

K, n : representam constantes dependentes do material.

Desenvolvida em 1867, a chamada “Primeira Lei de Cominuição”, ou Lei de Rittinger, sugere que a energia útil consumida é diretamente proporcional à nova superfície produzida por britagem ou por moagem diretamente proporcional à energia útil consumida. Por outro lado, a superfície específica dos materiais granulares é inversamente proporcional ao diâmetro das partículas, do que se origina. (Napier-Munn et al., 1999).

$$E = K (S_f - S_0) \quad (2)$$

E : representa a energia específica.

K : representa o fator de proporcionalidade.

S_f : representa a área do produto.

S_0 : representa a área inicial.

A segunda lei da cominuição foi formulada por F. Kick, o trabalho requerido é proporcional à redução em volume das partículas envolvidas, a expressão da lei é:

$$E_k = C_k (\ln X_1 / X_2) \quad (3)$$

E_k : representa a energia específica para um dado material.

C_k : representa uma constante.

X_1 : representa o diâmetro inicial.

X_2 : representa o diâmetro final.

Como os postulados de Rittinger e Kick não satisfaziam a todos os casos encontrados na prática e como se necessitava, na indústria, de algumas regras para classificar os materiais segundo as respostas à fragmentação, F.C. Bond postulou uma lei empírica muitas vezes chamada de "3a Lei de Fragmentação" – ou, simplesmente, Lei de Bond. “A energia consumida para reduzir o tamanho de um material é inversamente proporcional à raiz quadrada do tamanho”. Onde, por tamanho, entende-se a abertura da peneira pela qual passam 80% do material (Figueira, 2004)

Após uma intensa campanha de ensaios de laboratório e correlações industriais, em 1952, foi preconizado que a energia requerida na cominuição é proporcional ao comprimento das fissuras iniciais que se desenvolvem no fraturamento das partículas. Após a esta formulação e, devido ao fato de levar a resultados ajustados aos reais, a Lei de Bond passou a ser amplamente utilizada na seguinte expressão:

$$W = WI \left(\frac{10}{\sqrt{P}} - \frac{10}{\sqrt{F}} \right) \quad (4)$$

W : representa a energia aplicada em kWh/ton curta.

WI : representa o Work-index, em kWh/ton curta.

P, F : representa o diâmetro em microns, em que passam 80% da massa do produto e da alimentação, respectivamente

O Work Index, segundo Bond, é uma constante do material representando a energia necessária para cominuir de uma granulometria representada por um diâmetro infinito a uma representada por 80% passante em 100 micrometros.

Hukki (1961) verificou que as leis da cominuição tinham aplicação a certos intervalos de granulometria e propôs a seguinte expressão:

$$dE = -K \left[\frac{d_x}{x^{f(x)}} \right] \quad (5)$$

Donda (1999) apresentou estudos de uma metodologia desenvolvida para previsão de consumo específico de energia para moagem que vinha sendo estudada desde 1995. O objetivo de Donda era desenvolver um ensaio simples, de rápida execução e que fosse necessárias pequenas quantidades de amostras com intuito de possibilitar estudos geometalúrgicos. Este método desenvolvido apresentava resultados com baixa variação média em relação aos resultados industriais. As variações oscilavam entre 2 e 10%, evidenciando-se boa acurácia (Donda et al., 1999).

Savio (2003) adaptou o método desenvolvido para moagem primária para moagem secundária através de técnicas estatísticas de planejamento fatorial de experimentos. Com isso estabeleceu as melhores condições de operação para o ensaio desenvolvido (DONDA, 2003). De forma empírica, a partir da obtenção das curvas, determina-se que a energia específica necessária para a redução de tamanho dos minérios do Quadrilátero Ferrífero, no eixo pinhão de um moinho de bola. De acordo com Donda e Rosa (2014) pode ser expressa da seguinte forma:

$$R_p = R_f e^{-KE} \quad (6)$$

E : Energia específica em kWh/t, em moinhos tubulares, no eixo pinhão do moinho.

K : parâmetro característico do minério.

R_f : porcentagem retida de partículas na malha de interesse na alimentação do circuito.

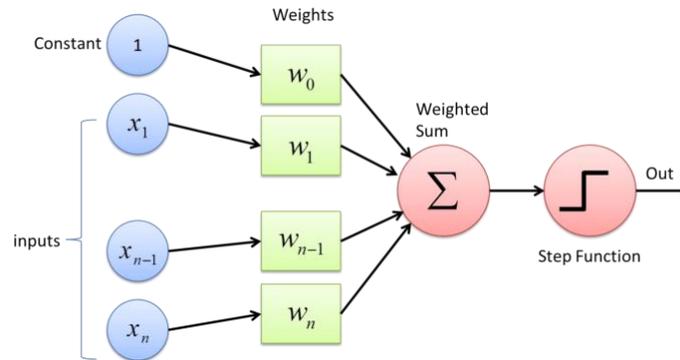
R_p : porcentagem retida de partículas na malha de interesse no produto do circuito.

Onde K é expresso em t/kWh e é dependente das condições de moagem. Este parâmetro é também conhecido como “K de Donda” e é a forma como vamos combinar este modelo empírico com os modelos de Machine Learning.

3.2 Redes Neurais Artificiais

O Perceptron foi introduzido por McCulloch (1943) como uma versão computacional de um neurônio biológico, que pode ser entendida como uma soma ponderada das entradas usadas em uma função de ativação para calcular uma saída, como mostra a Figura 4.

Figura 4 - Neurônio Artificial (Perceptron).

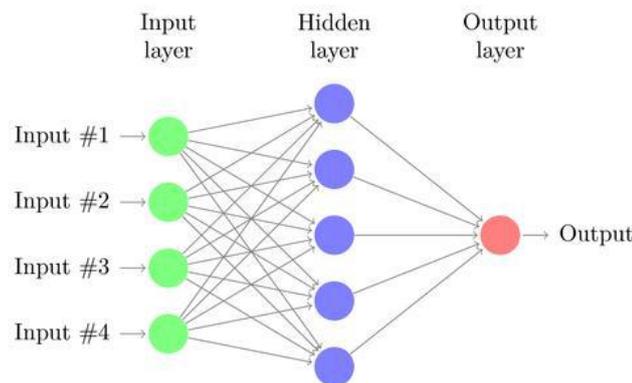


Fonte: https://medium.com/@tech_savvy/ml-scientist-day-3-notes-neural-networks-4f63dfe5be15

Com base nessa abstração, Minsky e Papert (1969) propuseram perceptrons organizados em camadas interconectadas e confirmaram as Redes Neurais Artificiais (RNA) com mais de uma camada oculta como uma ferramenta poderosa capaz de resolver problemas de separação não linear.

Uma rede Perceptron Multicamadas (*MLP, Multi Layer Perceptron*), por sua vez, possui uma camada de entrada para receber os sinais, camadas escondidas e a camada de saída. Cada camada escondida possui um dado número de neurônios, composto por peso, bias e função de ativação. Uma rede MLP é então inicializada com um vetor de entrada, que é multiplicado pelos pesos de cada camada e reavaliado pela função de ativação, propagando para a camada seguinte até que a camada de saída seja atingida. A Figura 5 ilustra um exemplo de MLP, onde a camada de entrada possui quatro elementos, há somente uma camada escondida com cinco neurônios e a camada de saída é única. A escolha da topologia da rede está relacionada à complexidade do problema e por isso é um desafio definir tais parâmetros. Neste trabalho, foi selecionada uma arquitetura com duas camadas escondidas, com 128 e 64 unidades (neurônios), utilizando-se função de ativação sigmoid.

Figura 5 - Multi-Layer Perceptron.



Fonte: https://medium.com/@tech_savvy/ml-scientist-day-3-notes-neural-networks-4f63dfe5be15

Como resumiu Serpa (2019), durante um treinamento da rede neural os pesos sinápticos são modificados após a iteração anterior ter extraído necessidade de novos parâmetros. Para a atualização é comum que se use o erro quadrático médio (RMSE) para minimizar a diferença entre a saída da rede e a saída desejada, e o algoritmo chamado de backpropagation é o método utilizado para definir pesos corretos para que a rede passe a mapear corretamente entradas e saídas, responsável pela evolução do treinamento de redes múltiplas camadas (Goodfellow, 2016). Como critério de parada do treinamento, pode-se limitar o número de iterações (ou épocas) ou o alcance de determinado erro proposto.

Considerou-se uma taxa de dropout de 10%, que, segundo (Srivastava, 2014), é uma técnica de regularização fácil de ser implementada, mas que aumenta o tempo de treinamento em cerca de 2 a 3 vezes. Os ganhos em usá-la estão relacionados à redução de probabilidade de acontecer overfitting e a diferenciação de arquiteturas da rede (Labach, 2019).

Neste trabalho, a preparação dos dados e a construção dos modelos com redes neurais foram feitas utilizando-se da linguagem de programação Python e as bibliotecas SkitLearn (Pedregosa, 2011) e Keras (Chollet, 2015).

4. Resultados e Discussão

A avaliação dos modelos foi feita dando continuidade aos passos seguintes da metodologia CRISP-DM. Após vários ciclos iterativos da metodologia apresentada, chegou-se a uma nova base de dados com 17 features de entrada, apresentadas na Tabela 2, mais a variável resposta, que representa o percentual de material com granulometria acima de 0,15mm. Diversas features foram excluídas por terem alta correlação com outras, mostrando que a base original não era linearmente independente.

Tabela 2 - Features selecionadas para modelagem.

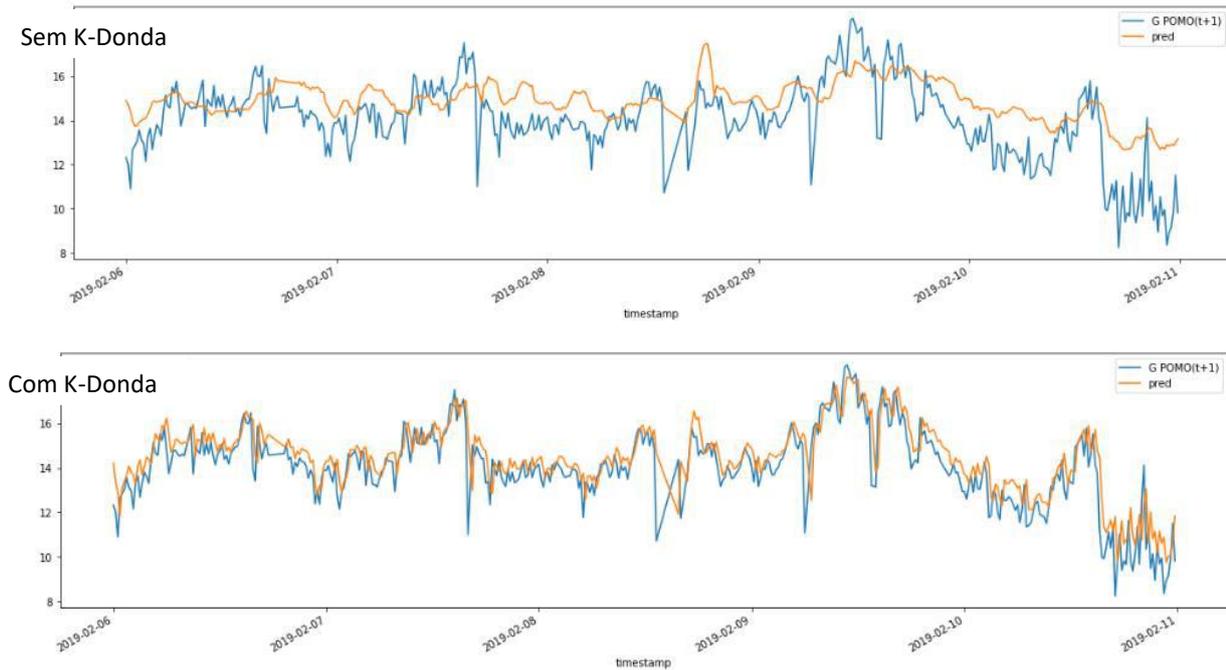
TAG	FEATURE
ALMOA_L101_WI	Usina_Moagem Linha 1_Alimentador de correia_Vazão mássica (BI1490Cn01/Wit1490Cn501)
BP1000CN04_JAI	Usina_Utilidades_Sistema de drenagem_Bomba de polpa_Potência Ativa Calculada
BPMOA_L101_DI	Usina_Moagem Linha 1_Bomba de polpa_Densidade no recalque (DIT-1490CN-001)
BPMOA_L101_FI	Usina_Moagem Linha 1_Bomba de polpa_Vazão no recalque (FIT-1490CN-001)
BPMOA_L102_DI	Usina_Moagem Linha 1_Bomba de polpa_Densidade no recalque (DIT-1490CN-002)
CIMOA_L102_CONT1	Usina_Moagem Linha 1_Hidrociclone_Número Válvulas com Estado Aberto
CIMOA_L102_PI	Usina_Moagem Linha 1_Hidrociclone_Pressão (PIT-1490CN-502)
CXMOA_L101_LI	Usina_Moagem Linha 1_Caixa de polpa_Nível (LIT-1490CN-009)
CXMOA_L110_LI	Usina_Deslamagem Linha 1_Caixa de polpa_Nível (LIT-1490CN-018)
MOMOA_L101_JAI1	Usina_Moagem Linha 1_Motor principal do Moinho_Potência
MOMOA_L101_KDONDA	Usina_Moagem Linha 1_Controlo Avançado_Constante K de Donda para MO-1490CN-01 pelo APC
MOMOA_L102_JAI1	Usina_Moagem Linha 01_Auxiliar seleção da linha 01 para circuito série
ALMOAFG_GLOBAL_+0.15MM	Usina_Moagem Linha 1_Moinho_Motor Principal_Potência
ALMOAFG_GLOBAL_+12.5MM	Itabiritos_+0.15MM do ALMOACN
ALMOAFG_GLOBAL_-0.045MM	Itabiritos_+12.5MM do ALMOACN
ALMOAQQ_GLOBAL_SIO2	Itabiritos_-0.045MM do ALMOACN
MOMOA_L101_XSEL	Itabiritos_Sílica do ALMOACN

Fonte: Autores.

Com a obtenção deste novo conjunto de dados, foram feitos testes para construção de redes neurais incluindo a utilização da constante de Donda (“MOMOAL101_KDONDA”) e também sem a inclusão da mesma, visando-se comparar os resultados e identificar o quanto o K-Donda pode contribuir com a melhoria da performance da rede neural.

Conforme pode ser observado na Figura 6, a aderência do modelo que considerou o K-Donda é visualmente melhor, fato confirmado pela Tabela 2, onde se pode verificar um RMSE quase 25% menor quando a referida constante é utilizada.

Figura 6 - Resultados Validação dos Modelos Com e Sem K-Donda.



Fonte: Autores.

Tabela 3 - RMSE para Modelo com e sem o K-Donda.

	RMSE (% +0,15)
SEM K DONDA	1,22
COM K DONDA	0,92

Fonte: Autores.

5. Conclusão

A rede neural treinada com os dados em questão resultou em um erro médio quadrático de 0,92% (onde “%” é a própria unidade, não se tratando de percentual de erro). Este valor pode ser considerado satisfatório, principalmente se considerarmos que, no caso de uma simulação preditiva, os operadores desta planta não precisaram aguardar os resultados de laboratório, podendo assim tomar decisões de forma antecipada, trazendo grande benefício ao controle de processo. Tal informação, além de alimentar a tomada de decisão do operador, pode também ser utilizada como entrada em sistemas especialistas e de controle avançado, com potencial de grande retorno em consumo energético para a referida moagem.

De qualquer forma, as métricas de RMSE observadas mostram que o modelo, com ou sem a inclusão do K-Donda, conseguiu diferenciar as duas configurações de circuito citadas, com os moinhos em série (moagem e remoagem) e com os moinhos independentes em paralelo.

Como trabalho futuro, além da implementação do modelo, que seria a última fase da metodologia CRISP-DM, é proposto também construir e avaliar modelos de classificação, ao invés de regressão, discretizando o valor do +0,15mm no produto, o que ampliaria o leque de possibilidades e poderia levar a melhores resultados.

Vale ressaltar que atrasos na entrada de dados oriundos das análises de laboratório podem ser prejudiciais na implementação dos modelos online e devem ser devidamente tratados.

Referências

- Mariscal, G., Marban, O., & Fernandez, C. (2010). *A survey of data mining and knowledge discovery process models and methodologies*. The Knowledge Engineering Review, 25(2), 137-166.
- IBM (2022) *Introduction to CRISP-DM*. <<https://www.ibm.com/docs/en/spss-modeler/18.2.0?topic=guide-introduction-crisp-dm>>
- Shearer, C. (2000). *The CRISP-DM Model: The New Blueprint for Data Mining*. *Journal of Data Warehousing*, 5, 13-22.
- Google (2022) *Machine Learning Crash Course*. <<https://developers.google.com/machine-learning/crash-course/>>
- Pessoa, A. D., Sousa, G. C. L., Araujo, R. C., & Anjos, G. J. M. (2021). *Modelo de rede neural artificial para previsão da capacidade de carga de estacas cravadas*. Research, Society and Development, 10(1), e12210111526. <https://doi.org/10.33448/rsd-v10i1.11526>
- Donda, J. D., Galinari C. M., & Rabelo, P. J. B. (1999). *O Controle da Eficiência Energética nos Circuitos de Pré-moagem e Moagem Primária da Samarco Mineração*. In: II Simpósio Brasileiro de Minério de Ferro - ABM, Ouro Preto, p. 144-150.
- Donda, J. D. (2003). *Um método para prever o consumo específico de energia na (re)moagem de concentrados de minério de ferro em moinhos de bolas*. Belo Horizonte, Tese (doutorado), CPGEM, Universidade Federal de Minas Gerais.
- Donda, J. D., & Rosa, A. C. (2014). *A lei de Moagem – Comprovação para minério de ferro*. Livraria e Editora Graphar. Brasil.
- Figueira, H. V. O., Almeida, S. L. M., & Luz, A. B. (2004). *Tratamento de Minérios, Rio de Janeiro*: Centro de Tecnologia Mineral CETEM. Editora CETEM. Brasil.
- Delboni, H. J. (2007). *Tendências Tecnológicas Brasil 2015*, Rio de Janeiro: Centro de Tecnologia Mineral CETEM. Editora CETEM. Brasil.
- Napier-Munn, T. J., Morrel, S., Morrison, R. D., & Kojovic, T. (1999). *Mineral Comminution Circuits, Their Operation and Optimisation*. Julius Kruttschnitt Mineral Research Center. Austrália.
- Serpa, A. (2019). *Modelagem do Teor de Sílica no Produto da Flotação de Minério de Ferro baseada em Redes Neurais*. Trabalho de conclusão de curso de graduação. Curso de Graduação em Engenharia de Controle e Automação, Escola de Engenharia da UFMG. Belo Horizonte, Brasil:
- Souza, E. S. (2014). *Controle Automático de Circuitos de Moagem*. Tese de Doutorado em Engenharia de Produção, Porto Alegre: Escola de Engenharia da UFRGS.
- Oliveira K P. S. (2000). *Apliação das técnicas de redes neurais e de análise de componentes principais na modelagem de uma lagoa aerada da Ripasa SA*. Tese de Mestrado, Faculdade de Engenharia Química, Unicamp.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press. EUA.
- Chollet, François and others (2022). *Keras API* <<https://keras.io/api/>>
- Labach, A., Salehinejad, H. & Valaee, S. (2019). *Survey of dropout methods for deep neural networks*. <<https://doi.org/10.48550/arXiv.1904.13310>>
- Hinton, G., Krizhevsky, A., Srivastava, N., et al. (2014). *Dropout: A simple way to prevent neural networks from overfitting*. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.
- Pedregosa, et al.,(2011). *Scikit-learn: Machine Learning in Python*, *Journal of Machine Learning Research*, *Journal of Machine Learning Research*, 12, pp. 2825-2830.
- Minsky, M., & Papert, S. (1969). *Perceptrons*. M.I.T. Press. EUA.
- McCulloch, W. S., & Pitts, W. (1943). *A logical calculus of the ideas immanent in nervous activity*. *Bulletin of Mathematical Biophysics* 5, 115–133.