



**VI Encontro de Iniciação Científica e
Tecnológica
VI EnICT
ISSN: 2526-6772
IFSP – Câmpus Araraquara
21 e 22 de outubro de 2021**



UTILIZAÇÃO DE BIG DATA PARA GESTÃO DE PROCESSO

MATHEUS C. DE OLIVEIRA¹, ELCIO R. ARANHA²

¹ Graduando em Engenharia de Controle e Automação, IFSP – Campus Cubatão, c.matheus@aluno.ifsp.edu.br.

² Mestre em Engenharia Elétrica, Docente no IFSP – Campus Cubatão, aranha@ifsp.edu.br.

Área de conhecimento (Tabela CNPq): 30405033 Controle de processos eletrônicos, retroalimentação.

RESUMO: Um processo produtivo pode ser observado a partir de propósitos diversos, sem considerar todos os dados e, por conta disso, as conclusões podem ser discrepantes para um mesmo período de análise. Nos diversos segmentos industriais, a recente e crescente disponibilidade de informações dos processos, podem disponibilizar dados diversos dos que as análises tradicionais levavam em conta. Técnicas de apuração e análise de dados podem ser bem-vindas para uma tomada de decisão mais assertiva quando passa a ser fundamentada em um conjunto mais abrangente de informações. Como os demais processos industriais, o processo de mineração, demanda de diversos insumos e, cuidar para o maior rendimento da matéria prima, energia e diminuição da emissão de gases são necessidades recorrentes deste e da maioria dos processos energo-intensivos. A flotação é um processo utilizado para separação de partículas de um material, a coleta desse dado é importante para determinação de impureza do minério, diminuindo a quantidade de rejeito utilizada. Os resultados obtidos demonstram uma análise dos principais dados do processo de flotação, verificando as suas correlações e possíveis tomadas de decisão.

PALAVRAS-CHAVE: Análise de dados; Determinação de impurezas; Flotação; Mineração.

USE OF BIG DATA FOR PROCESS MANAGEMENT

ABSTRACT: A production process can be observed from different purposes, without considering all the data and, therefore, conclusions may differ for the same period of analysis. Across industries, the recent and growing availability of process information can provide data that differs from what traditional analytics have considered. Data verification and analysis techniques can be welcome for more assertive decision making when based on a broader set of information. As with other industrial processes, the mining process, the demand for various inputs and taking care of the highest yield of raw material, energy and reduction of gas emissions are recurrent needs of this and most energy intensive processes. Flotation is a process used to separate particles from a material, collecting this data is important to determine the impurity of the ore, reducing the amount of waste used. The results obtained demonstrate an analysis of the main data from the flotation process, verifying their correlations and possible decision-making.

KEYWORDS: Data analysis; Determination of impurities; Flotation; Mining.

INTRODUÇÃO

Nos últimos anos, o aumento da capacidade de processamento e relação custo-eficácia dos sistemas eletrônicos têm causado um grande impacto nas empresas, influenciando os negócios de modo a gerar um aumento de produção, redução de custos, melhoria da qualidade, segurança, e monitorização remota. O aumento de dados gerados por esses dispositivos trazem informações importantes sobre o processo de produção e gerenciamento.

A Mineração de Dados, *data mining*, é um processo de exploração para grandes volumes de dados para que anomalias, padrões e correlações suportem a tomada de decisão e propiciem vantagens estratégicas. A Mineração de Dados se tornou popular a partir da utilização do termo *Big Data* que torna a mineração de dados como elemento central e responsável pela preparação e análise das massas de dados. A utilização da mineração de dados pode ser ferramenta fundamental para a seleção, agrupamento, classificação de massas de dados (DE CASTRO; FERRARI, 2016) e (KUMAR et al, 2009).

A mineração de dados é determinada por quatro processos principais, onde possui o banco de dados, determinado pelas diversas informações sobre um determinado processo, o depósito de dados local onde se localiza os dados digitais, que a partir das minerações de dados busca correlação entre padrões e reconhecimento, de forma, a determinar os padrões consistentes que melhor represente o processo em análise, gerando ao fim, dados de conhecimento sobre o estudo predefinido.

O termo big data, vem sendo utilizado para representar a interpretação e análise de um grande volume de dados. Uma vez que o armazenamento dos dados pode apresentar uma forma diversa para cada fonte e muitas vezes se apresentando de forma não estruturada, métodos de agrupamento, validação, contextualização e correção das informações são significativos para o sucesso do projeto. Este método será utilizado para analisar o grande volume de dados gerado pelo processo de mineração.

Cahill e Gallachóir (2011) descreveram um método para qualificar, mensurar e acompanhar a eficiência energética em grandes consumidores de energia ao longo dos anos, sendo que um grupo de indicadores de desempenho pode contribuir para a manutenção do foco dos colaboradores envolvidos no projeto entre períodos de análise.

May et al. (2013), apresentam em seu trabalho um formato para uma visão gerencial da eficiência energética na produção e as relações entre as diversas frações do processo que será tomado como balizador para as premissas do aparato de avaliação (SCHMIDT et al., 2016).

A flotação é um processo que consiste na separação de misturas de elementos que não possuem dissolução entre si, aplicadas em misturas heterogêneas. O processo de flotação consiste em dois tipos básicos tanto por ar dissolvido, quanto por ar disperso, dependente da relação de ar/sólidos e tamanho de bolha.

O processo de mineração possui diversas etapas, entre elas, o processo de flotação que consiste na seleção e separação de partículas do material, esses dados são de suma importância para determinação do grau de impureza que existe na concentração do minério. Com a antecipação desses dados é possível obter uma tomada de decisão e possíveis ações corretivas como reduzir impurezas, além de ajudar o meio ambiente com a diminuição da quantidade de rejeito à medida que se reduz a sílica na concentração do minério (KAGGLE, 2017).

Objetivos Gerais:

- Desenvolvimento de métodos para análise de dados, a partir de um software estatístico;
- Desenvolvimento de métodos qualitativos de predição utilizadas para tomadas de decisão;
- Utilização de métodos matemáticos para correlação das variáveis principais no processo de flotação.

MATERIAL E MÉTODOS

A pesquisa se baseou primeiramente na busca de dados, a partir de um problema real, para isto foi utilizado a plataforma Kaggle que tem como intuito o aprendizado sobre análise de dados, possuindo um grande acervo de processos reais para análise. O processo escolhido consiste na análise da flotação, uma etapa utilizada na mineração.

A primeira etapa do projeto foi a exportação do banco de dados para a plataforma R, software de programação baseado em análises estatísticas. A tabela abaixo demonstra algumas colunas e linhas da tabela original, por conta de sua dimensão é demonstrado alguns dados de análise.

	%Iron Feed	%Silica Feed	Starch Flow	Amina Flow	Ore Pulp Ph	Ore Pulp Density	...	%Iron Concentrate	%Silica Concentrate
0	55,2	16,98	3019,53	557,434	10,0664	1,74		66,91	1,31
1	55,2	16,98	3024,41	563,965	10,0672	1,74		66,91	1,31
2	55,2	16,98	3043,46	568,054	10,068	1,74		66,91	1,31
3	55,2	16,98	3047,36	568,665	10,0689	1,74		66,91	1,31
4	55,2	16,98	3033,69	558,167	10,0697	1,74		66,91	1,31
...	...	16,98
737453	49,75	23,2	2710,94	441,052	9,62129	1,65365		64,27	1,71

Tabela 1 – Dados utilizados para análise.

Foi realizado a curva de correlação das variáveis em análise, por meio de um Corrplot. O pacote corrplot R fornece uma ferramenta de exploração visual na matriz de correlação, que suporta a reordenação automática de variáveis para ajudar a detectar padrões ocultos.

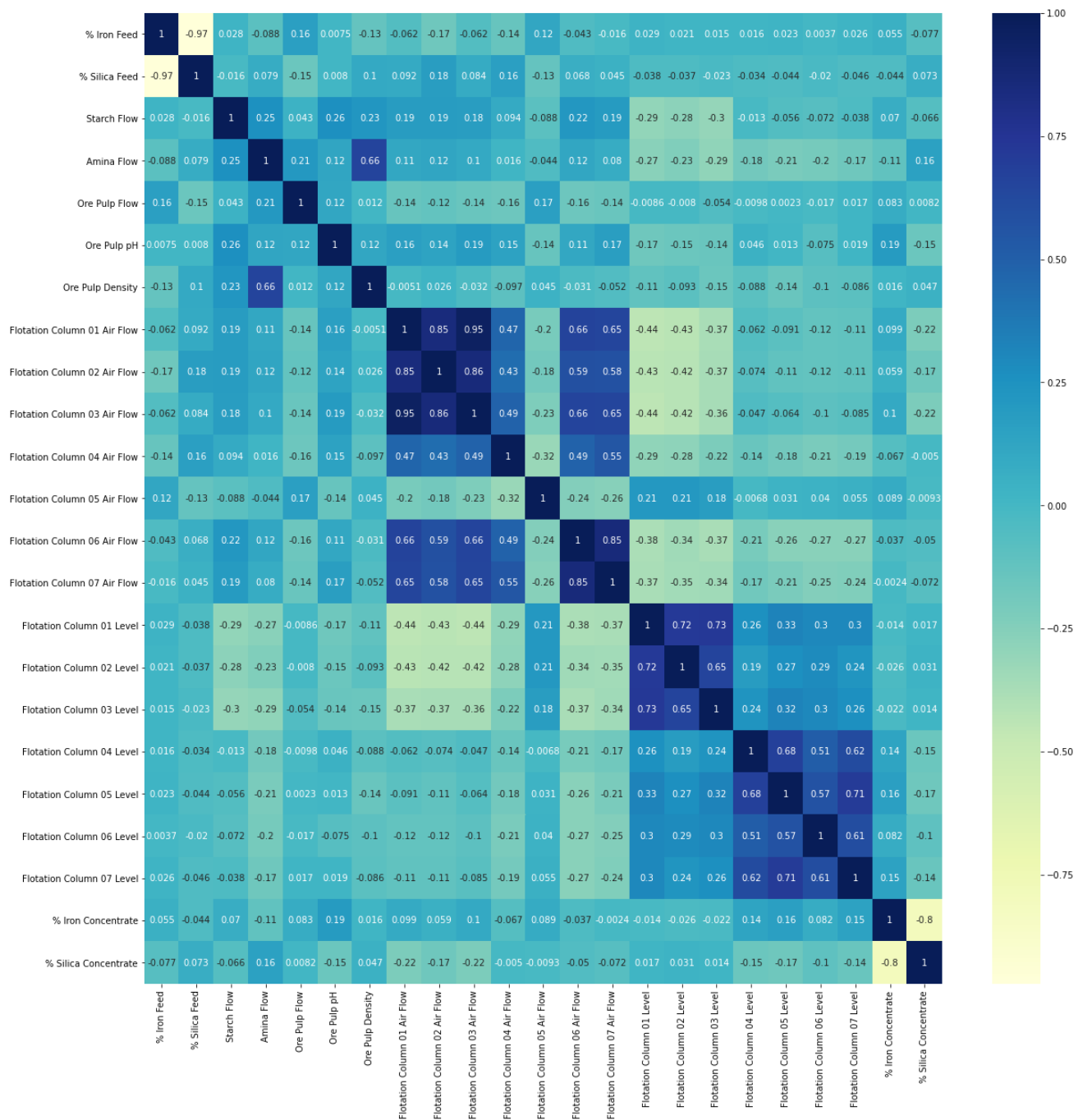


Figura 1 – Corrplot das variáveis em análise.

Observe que na diagonal principal, obtemos o maior valor de correlação, por conta de ser relacionando o mesmo conjunto de variável. É possível observar a correlação das principais variáveis do processo, o que pode ser utilizado para uma tomada de decisão de acordo com o critério do operador da planta.

Na próxima etapa foi analisado as características mais importantes do processo de flotação a % Iron Feed e % Silica Feed, bem como o nível de PH do minério, parâmetros substanciais de controle, Figura 2.

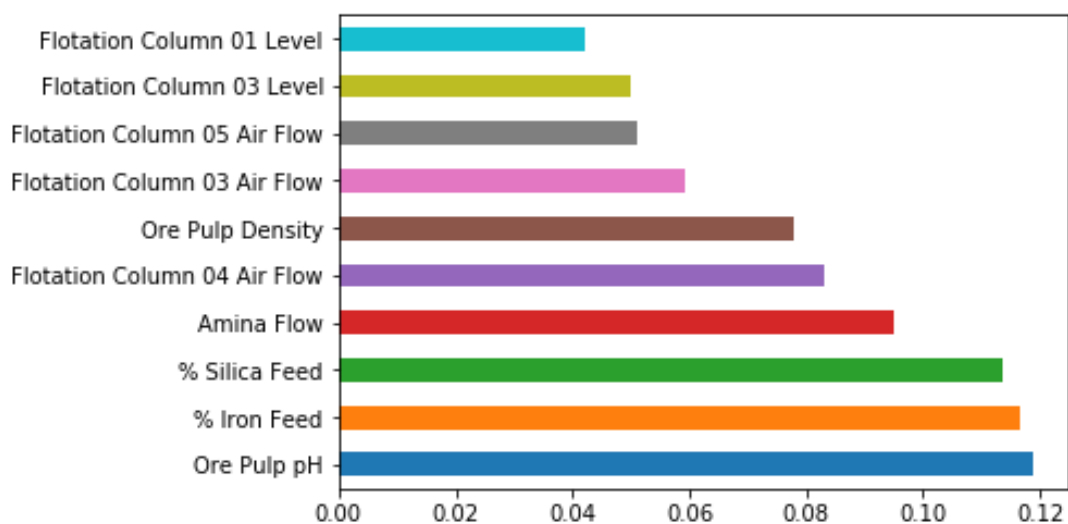


Figura 2 – Parâmetros de controle.

A Figura 2 demonstra a %Iron Feed e %Silica Feed citado anteriormente, além de outros elementos para análise, sendo estes dois principais por conta de representar a concentração do minério de ferro e a sílica presente, parâmetros substanciais para determinação e previsão da porcentagem de impureza de sílica a cada minuto do processo.

Na última etapa foi realizado um teste para verificar como cada recurso se correlaciona com os outros. É possível verificar que os recursos que se fundem no lado direito estão próximo uns dos outros, o que sugere que possivelmente que apenas um dos dois seria suficiente treinar o modelo. Essas variáveis de seleção podem ser determinadas a partir do P-Valor, descartando possíveis dados não substâncias para o modelo e realizando o treinamento.

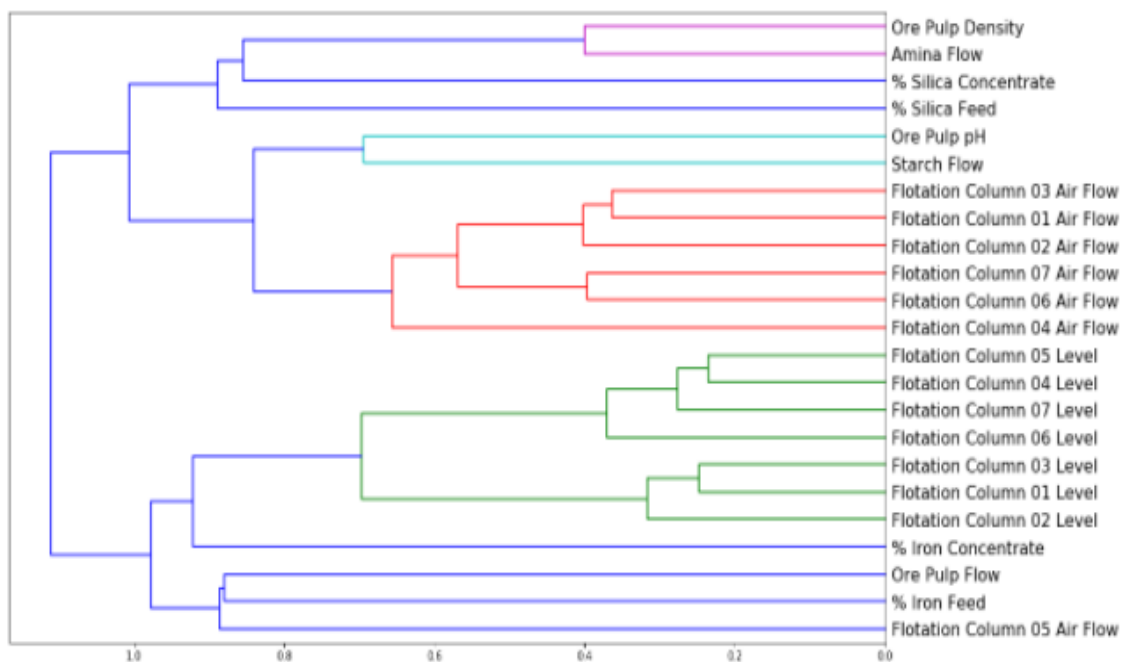


Figura 3 – Validação da correlação do modelo.

Com a validação dos dados é possível verificar possíveis ideias para o tratamento de flotação, entretanto, algumas variáveis finais precisam ser analisadas para determinar todas as soluções para as questões apresentadas nos objetivos.

CONCLUSÕES

Os processos industriais possuem uma gama ampla de dados para análise, método qualitativos de análise numérica se torna válida para uma maior tomada de decisão e diminuição do descarte de recursos naturais. O processo de mineração possui inúmeras variáveis que necessitam de um controle, para predição e validação de resultados. O processo de flotação foi utilizado como análise no projeto, foi verificado a correlação das variáveis que envolvem esse tipo de processo gerando possíveis estratégias de produção, entretanto, se faz necessário ampliar o alcance na análise dos dados, onde será possível até o fim do projeto, com o aporte de um maior número de dados para análise.

REFERÊNCIAS

DE CASTRO, L. N.; FERRARI, G. D.: Introdução A Mineração De Dados - Conceitos: Básicos, algoritmos e aplicações. São Paulo, Saraiva Editora, 1. ed. p.376, 2016.

CAHILL, C. J.; GALLACHÓIR, B. P. Quantifying the savings of an industry energy efficiency programme, Energy Efficiency, v. 5, p. 211 – 224, mai. de 2011.

MAGALHÃES, Eduardo. Quality Prediction in a Mining Process. Disponível em: <https://www.kaggle.com/edumagalhaes/quality-prediction-in-a-mining-process>. Acesso em: 24 ago. 2021.

MAY, G.; TAISCH, M.; PRABHU, V. V.; BARLETTA, I. Energy Related Key Performance Indicators – State of the Art, Gaps and Industrial Needs. Advances in Production Management Systems. Sustainable Production and Service Supply Chains, v. 1, p. 257-267, set. de 2013.