

## ABORDAGEM DE *DATA ANALYTICS* PARA DIAGNÓSTICO DE DESEMPENHO DE PROCESSO\*

Matheus Lage Miranda<sup>1</sup>  
Hamilton Guilherme de Oliveira Junior<sup>2</sup>

### Resumo

A Indústria 4.0 tem provocado uma onda de transformação nos modelos de negócio e operacional de muitas empresas, sendo que a disciplina de *Data Analytics* tem sido bastante explorada para alavancar melhorias operacionais dos processos produtivos. Neste sentido, este artigo aborda como esta disciplina pode aprimorar o gerenciamento dos processos produtivos, fazendo com que a tomada de decisão seja baseado em dados. Dessa forma, será apresentado como o *Data Analytics* pode ser usado para realizar diagnóstico de desempenho de processos produtivos que irão permitir mapear ações que promovem uma melhoria operacional. Ao longo da descrição da abordagem metodológica proposta, este artigo também menciona um estudo de caso e respectivos resultados onde tal abordagem foi aplicada.

**Palavras-chave:** Indústria 4.0; *Data Analytics*; Desempenho de Processo Industrial.

### DATA ANALYTICS APPROACH FOR PROCESS PERFORMANCE DIAGNOSIS

#### Abstract

Industry 4.0 has been generating a transformation wave in the business and operational models of many companies, and the Data Analytics discipline has been heavily used to leverage operational improvements in production processes. Based on it, this article discusses how this discipline can improve the management of productive processes, transforming the decision making to be based on data. It will be presented how Data Analytics can be used to perform performance diagnostics of productive processes that will allow mapping actions that support operational improvement. Throughout the description of the proposed methodological approach, this article also mentions a case study and respective outcomes where such an approach was applied.

**Keywords:** Industry 4.0; Data Analytics; Process Performance.

<sup>1</sup> Engenheiro de Controle e Automação com pós-graduação em Automação Industrial e Gestão de Projetos. Gerente de Soluções Industriais, Accenture. Belo Horizonte-MG – Brasil.

<sup>2</sup> Engenheiro de Controle e Automação com especialização em Gestão de Negócios. Gerente de Soluções Industriais, Accenture. Belo Horizonte-MG – Brasil

## 1 INTRODUÇÃO

A recente onda de transformação digital, reconhecida como Indústria 4.0, tem demandado que as empresas repensem seus modelos de negócio, reavaliando e aprimorando sua forma de trabalhar, seus processos internos e, principalmente, a forma com que se relacionam com clientes e fornecedores. Segundo Tagliani [1], essa onda transformacional, intensificada nos últimos 5 anos, visa não somente a aplicação de ferramentas digitais para aumentar sua maturidade tecnológica, mas também objetiva que tais ferramentas gerem valor ao negócio, melhorando o desempenho da empresa e alcançando resultados efetivos na sua cadeia de valor.

No contexto da indústria de recursos naturais, as empresas têm buscado conectar os temas relacionados à Indústria 4.0 ao ambiente industrial, visando promover a melhoria dos processos operacionais internos. Para conseguir tal objetivo, Pessoa[2] aponta que é preciso estabelecer um plano estratégico de transformação que mapeie iniciativas que promovam maior integração entre os setores da empresa, reduzindo os silos existentes, trazendo maior eficiência e agilidade nos processos de comunicação e maior gerenciamento dos processos de tomada de decisão baseado em dados.

Dentre as soluções tecnológicas relevantes para a geração de valor para os processos operacionais, destaca-se o DataAnalytics. Trata-se da ciência que se dedica à análise de grandes volumes de dados que busca extrair informação relevante com o propósito de tirar conclusões e a partir dessas conclusões, desdobrar decisões e ações que impactam no negócio. A Figura 1 ilustra, de forma geral, uma jornada de desenvolvimento de soluções de Data Analytics, destacando os passos necessários para tratamento de um problema de negócio e para a transformação em ações que irão alavancar resultados de negócio.

As empresas de alto desempenho reconhecem que os dados são um ativo estratégico nos dias atuais e têm se esforçado para adotar uma cultura baseada em dados. Em tempos de mudanças em ritmos exponenciais e alta velocidade de acesso à informação, as decisões baseadas em dados levam cada vez mais a resultados comerciais claros, gerando um retorno mensurável do investimento.



Figura 1: Processo de geração de valor com Data Analytics

Fonte: Accenture

## 2 CONTEXTO DE CONTEXTUALIZAÇÃO DO PROBLEMA DE NEGÓCIO

Embora haja uma tendência de consolidação da percepção de valor com o uso de *Data Analytics*, ainda há um grande desafio em definir uma abordagem padronizada diante da existência de muitas plataformas e ferramentas, fornecedores e técnicas analíticas. Além disso, ainda é observada uma visão equivocada de focar em tecnologia em detrimento dos objetivos de negócio.

O primeiro passo e, também o mais importante, para a concepção e o desenho de uma solução de *Data Analytics* é estabelecer claramente a missão do projeto. Ou seja, é preciso estar claro qual problema de negócio pretende-se resolver por meio do uso de uma solução que envolve *data analytics*.

No contexto deste artigo, apresenta-se um estudo de caso de uma planta de extração de manganês que conta com uma usina de beneficiamento de minério destinada à produção de granulado e finos. A planta industrial apresentava uma tendência de redução de sua taxa média de alimentação ao longo de um determinado período de tempo entre os anos de 2016 e 2017. Tal fato comprometia não só a produtividade, como também a eficiência do processo na recuperação do mineral de valor. O cliente extrai o minério de manganês de várias frentes de lavra usando caminhões, escavadeiras e carregadeiras e o transporta para uma pilha de *runof mine* (ROM). Desta pilha, o minério é enviado à planta de beneficiamento, responsável por produzir *pellet feed* e granulado.

Diante disso, definiu-se então que o problema de negócio constitui em se identificar as principais causas-raiz da queda na taxa de alimentação por meio do desenvolvimento de um modelo matemático, construído a partir dos dados históricos do processo. Dessa forma, vislumbrou-se que modelo deveria ser capaz de explicar o comportamento de uma variável alvo com base em um conjunto de variáveis de entrada.

Em um segundo momento, garantida a confiabilidade e assertividade do modelo construído, o modelo matemático poderia ser usado para estimar um comportamento futuro da variável alvo em função dos valores conhecidos para as variáveis de entrada.

Após a definição do problema de negócio, é importante fazer uma avaliação da condição da infraestrutura de dados existente, buscando entender a arquitetura de sistemas de informação, inventário das ferramentas a serem usadas, base de dados e disponibilidade e qualidade dos dados históricos. No estudo de caso citado, o cenário atual da planta industrial considerava a existência de diversos sistemas de automação e informação não integrados, o que dificultava ter uma visão completa e unificada do processo produtivo. Dessa forma, as bases de dados históricos dos sistemas relacionados às áreas de operação de mina, operação da planta industrial, manutenção e laboratório se encontravam isoladas em silos de informação. Além disso, observou-se a ocorrência de entradas manuais em alguns dados históricos, o que impactava diretamente a confiabilidade da informação.

## 3 A BORDAGEM E METODOLOGIA

Apesar de o cliente já ter algumas suspeitas de possíveis causas, as análises se apresentavam fragmentadas e dispersas. Outros trabalhos foram feitos

anteriormente tentando utilizar *machinelearning* para se chegar a uma resposta, mas não houve sucesso. A Accenture e a Avanade trabalharam em parceria para a solução deste problema, onde a abordagem do problema teve como diferencial a combinação do conhecimento do processo produtivo com a adoção das técnicas de modelagem e *machinelearning* na estruturação da metodologia de trabalho. A Figura 2 ilustra, de forma geral, a metodologia aplicada ao trabalho de identificação de causa-raiz do problema de queda de desempenho do processo.



Figura 2: Metodologia para Determinação de Causas-Raiz da Queda de Desempenho do Processo

Fonte: Accenture

A metodologia de trabalho se inicia com a estruturação do problema, buscando agrupar potenciais causadores do referido problema assim como classificar fontes de dados a serem analisadas em categorias e/ou pilares inter-relacionados como forma de não só organizar o racional de investigação do problema como também direcionar as posteriores análises de dados. Dessa forma, foram estabelecidas 4 categorias:

- *Operação de Mina* abordou causas relativas à operação da mina em si, antes do processo de beneficiamento, como incapacidade de prover a alimentação necessária para a planta de beneficiamento e modificação das características do minério devido à operação da mina.
- *Características do minério* foi a dimensão onde foram consideradas as características do minério que poderiam influenciar na produção da planta de beneficiamento, tais como a composição química proveniente das diferentes frentes de lavra utilizadas, a granulometria, sua umidade, etc.
- No pilar de *Manutenção*, foram considerados problemas relacionados à manutenção, tais como paradas por falhas em equipamentos, paradas de processo, mudanças de equipamentos após paradas programadas, dentre outros.
- O pilar de *Eficiência de Operação* refere-se à análise de mudanças na forma de operação do processo produtivo, tais como mudanças de pontos operacionais, limites de equipamentos e procedimentos operacionais.

Estabelecido o racional de investigação do problema nessa primeira etapa, inicia-se a etapa dedicada ao Entendimento do Processo, caracterizado por um período de

tempo com intensa interação com equipes e de vivência do dia-a-dia do processo. Foram realizadas diversas entrevistas com os principais *stakeholders* para aprofundar o conhecimento a cerca do funcionamento do processo, suas particularidades e desafios. Além disso, nesta etapa foram mapeadas as fontes de dados históricos e identificadas as principais variáveis que deveriam compor as análises de dados.

Já o terceiro estágio da metodologia se destina à preparação de infraestrutura e coleta de dados. Foi adotada uma infraestrutura tecnológica de software baseada em nuvem, de forma a facilitar o trabalho das equipes remotas de ciência de dados e análise, conforme ilustrado na Figura 3.



Figura 3: Infraestrutura em Nuvem adotada

Fonte: Accenture

Foi utilizada a plataforma do Microsoft Azure nas camadas de integração e de armazenamento de dados na nuvem. Sendo assim, foram coletados vários dados históricos dos principais sistemas de automação e de informação existentes, tais como o Sistema de Gerenciamento de Informações de Processo (*Process Information Management System – PIMS*), de gerenciamento de frota, de paradas de equipamentos, de gerenciamento da produção, de laboratório contendo a qualidade do minério e também de bases meteorológicas locais. Para a camada de *Analytics* propriamente dita, foi adotada a ferramenta estatística R, utilizada em boa parte das análises e na construção de modelos, enquanto que o Microsoft Power BI foi ferramenta escolhida para visualização dos dados e resultados do modelo.

Na sequência da metodologia, no quarto estágio, se inicia o processo de transformação e integração dos dados, onde são corrigidos e/ou eliminados eventuais erros nos dados e derivados cálculos adicionais e/ou gerados novas métricas. Além disso, destaca-se nesta etapa a estruturação de um modelo de arquitetura de dados onde são estabelecidos as hierarquias e os relacionamentos entre variáveis.

A etapa seguinte, de análise exploratória, como o próprio indica se destina a exploração dos dados visando avaliar a distribuição estatística de cada variável e faixas de operação assim como estabelecer correlações das variáveis de entrada

com uma variável alvo. A análise exploratória é, na verdade, uma etapa interativa com a etapa anterior de transformação dos dados, uma vez que é natural que a partir da análise exploratória se perceba a necessidade de estruturar, limpar e/ou relacionar os dados de uma forma diferente para o processo de análise produza um resultado adequado.

Na sequência, inicia-se a etapa de seleção das variáveis e modelagem, onde foram realizadas análises baseadas, principalmente, em duas diferentes técnicas analíticas. A primeira delas se trata do PCA – *Principal Component Analysis*, com a qual buscou-se isolar as variáveis que apresentaram as maiores mudanças de variabilidade entre os períodos de melhor desempenho e de desempenho indesejado. A segunda técnica foi correlação cruzada, na qual as variáveis foram selecionadas de acordo com a sua correlação (índice  $R^2$ ) direta com a variável alvo do problema de negócio. A correlação foi importante por habilitar a análise das variáveis pré-selecionadas dentro de um contexto de causa e efeito. Além da correlação matemática, dado pelas técnicas analíticas utilizadas, a etapa de seleção de variáveis contou também o aporte de conhecimento técnico do processo produtivo. Como exemplo, buscou-se entender as correlações entre o tipo do minério e a produção total do processo, mas não a correlação entre a corrente do motor da correia onde era medida a produção final e a própria produção. Apesar de já ser conhecido que a correlação entre a corrente do motor e a produção na correia é extremamente forte, tal correlação é óbvia e não indicaria a causa-raiz da queda na produção. Por último, utilizou-se as variáveis selecionadas para a construção dos modelos do processo e quantificação das causas de baixas de produção

Por fim, utilizou-se as variáveis selecionadas para a construção dos modelos matemáticos do processo e quantificação das causas de baixo desempenho. Para um modelo de análise inicial, adotou-se como técnica a regressão linear múltipla. As variáveis selecionadas foram testadas no modelo respeitando-se princípios estatísticos, tais como os resíduos do modelo, a análise de variância e o número máximo de variáveis a serem consideradas dado os números de pontos do histórico de dados. Foi também construído um modelo em rede neural para demonstrar a capacidade de predição desta técnica.

A última etapa da metodologia consiste na validação da análise e dos modelos matemáticos e na determinação de causas-raiz. Sendo assim, os modelos construídos serviram como base para identificar quais variáveis apresentaram as principais contribuições às mudanças de comportamento do processo entre os períodos de melhor desempenho e de desempenho indesejado do processo. Além disso, a validação dos modelos foi feita de forma interativa, permitindo atingir uma melhor e mais aprofundada compreensão das causas-raiz elencadas e particularidades do processo produtivo, o que garantiu a precisão da análise. Por fim, foram construídos relatórios utilizando o Power BI, apresentando os resultados dos modelos, as correlações encontradas, principais causas-raiz e particularidades da fase de análise exploratória dos dados.

#### **4 RESULTADOS**

As análises foram constantemente discutidas, alinhadas e inclusive desafiadas pelo cliente, mostrando-se consistentes e condizentes com a realidade. Ao final do trabalho, os resultados obtidos com as indicações dos modelos permitiram estabelecer e disparar um plano de ações corretivas, que por sua vez, promoveram

uma retomada do desempenho do processo a patamares elevados, inclusive maiores que os observados anteriormente.

As principais fontes de problemas encontradas foram relativas à:

- **Falhas de equipamentos:** maior quantitativo de falhas em bombas de polpa, principalmente causadas devido à alteração no percentual de finos do minério, resultando em redução parcial da produção.
- **Mudança nas características do minério:** maior quantidade de argila no minério, que, potencializada pelas chuvas, promovia produção lenta e forçava a redução da produção, principalmente pela maior exploração de uma determinada frente de lavra. Além disso, foi detectado aumento no seu percentual de finos, influenciado por determinadas frentes de lavra e pelo método atual de empilhamento de minérios.

Ressalta-se ainda que, de forma qualitativa, os modelos também ajudaram a proporcionar um entendimento melhor do funcionamento do processo e a manter o foco nos problemas mais relevantes, melhorando sua eficiência operacional e planejamento da produção.

## 5 CONSIDERAÇÕES FINAIS

Este tipo de trabalho de *Data Analytics* pode ser considerado de alta complexidade, visto que, normalmente, existe pouco tempo para a execução – 10 semanas no caso desse trabalho – e a necessidade de envolvimento de profissionais com perfis diferentes, tais como cientista de dados, engenheiros de processo, engenheiros de produção, engenheiros de controle de processo e automação, profissionais de tecnologia da informação e de manutenção de processos.

Além disso, este trabalho compreendeu um estudo de **todo** o processo produtivo com o objetivo de se obter um modelo. A delimitação de um problema de *Data Analytics* é sempre desafiadora, e deve ser feita com cautela para conseguir manter foco e encontrar as soluções esperadas.

Neste trabalho, foram adotadas, em conjunto, técnicas de análise estatística descritiva e de *machine learning* dentro de uma moderna estrutura em nuvem, o que facilitou o trabalho sincronizado das equipes com diferentes perfis.

Diante dos resultados positivos obtidos, o trabalho de *Data Analytics* possibilita a geração de novas oportunidades de melhoria do processo como uma evolução natural do trabalho entregue. Como exemplo, cabe mencionar que é possível considerar a construção de modelos analíticos em tempo real para apoiar rapidamente a programação da produção, e/ou de modelos preditivos para aumentar a eficiência da operação da unidade. Para tanto, deve ser sempre levada em consideração a confiabilidade dos dados atuais, podendo estar vinculada à alteração da atual instrumentação de campo, deixando os modelos e análises mais precisos.

## REFERÊNCIAS

- 1 TAGLIANI, Federico; “Quatro Pilares fundamentais para a Transformação Digital na America Latina”. 2016. Disponível em:

<http://computerworld.com.br/quatro-pilares-fundamentais-para-transformacao-digital-na-america-latina>. Acesso em 25/04/2018.

- 2 Pessoa, Leandro. “Os Três Pilares da Transformação Digital”. White paper Overdub. Publicado em 05/12/2016..