



FUNDAMENTOS DE BIG DATA

Atividade Somativa 1



Nome: Vanessa Milani Ratusznei

Curso: Análise e Desenvolvimento de Sistemas

RESENHA

Resenha de “Como Big Data Pode Auxiliar as Empresas a Gerar Vantagem Competitiva: Estudo de Caso Único em Empresa da Indústria Agrícola” por Felipe Renz para MBA em Administração de Tecnologia da Informação em Universidade do Vale do Rio dos Sinos (UNISINOS) – São Leopoldo, 2014 <<http://www.repositorio.jesuita.org.br/bitstream/handle/UNISINOS/6708/Felipe%20Renz.pdf?sequence=1>>

O estudo de caso apresentado por Felipe Renz em sua dissertação "Como Big Data Pode Auxiliar as Empresas a Gerar Vantagem Competitiva: Estudo de Caso Único em Empresa da Indústria Agrícola" foca em uma organização do setor de máquinas agrícolas. O objetivo central da pesquisa é investigar a implementação de uma solução de Big Data associada à análise preditiva para prever falhas em máquinas agrícolas, evitando paradas não programadas que podem causar prejuízos e gerar uma vantagem competitiva para a empresa.

O estudo utiliza uma abordagem metodológica que inclui revisão de literaturas existentes sobre Big Data e suas aplicações, e entrevistas com gestores e colaboradores da empresa. Com base nessa coleta de informações, foi desenvolvido uma arquitetura para a adoção de Big Data, compatível às necessidades e estrutura organizacional da empresa.

Dentre os principais objetivos do estudo, se destaca o de prever falhas nas máquinas agrícolas, permitindo a realização de manutenções preventivas e evitando interrupções não programadas. A manutenção preventiva é uma estratégia que pode reduzir significativamente os custos operacionais. Ao prever quando uma máquina agrícola está prestes a falhar, a empresa pode agendar manutenções em momentos que minimizem a interrupção das operações, evitando paradas não programadas que podem ser extremamente dispendiosas. Essa estratégia proporciona um serviço mais confiável e eficiente, permitindo que a empresa se destaque perante à concorrência.

Esta capacidade de previsão de falhas é alcançada através do processamento e análise de grandes volumes de dados operacionais coletados das máquinas em uso. Dessa forma, o estudo explora detalhadamente a aplicação dos 5Vs do Big Data (Volume, Velocidade, Variedade, Veracidade e Valor) para o desenvolvimento da arquitetura proposta, garantindo uma



compreensão completa de como cada um desses aspectos pode contribuir para a eficácia do sistema.

Outro objetivo importante é o de melhoria contínua dos processos produtivos. A capacidade de analisar dados históricos e em tempo real, possibilita a identificação de padrões e tendências, permitindo a criação de estratégias de otimização (desempenho de máquinas) e de processos de produção. Por exemplo, uma análise pode revelar certas condições operacionais que levam a um desgaste mais rápido das peças, dessa forma, a empresa pode ajustar suas práticas de manutenção para prolongar a vida útil dos equipamentos.

Portanto a geração de vantagem competitiva é alcançada através da capacidade de oferecer um serviço superior de qualidade. A previsibilidade e a confiabilidade proporcionadas por uma solução de Big Data se tornam um diferencial importante no mercado agrícola. Agricultores que buscam que suas máquinas tenham mais tempo de produtividade escolherão fornecedores que oferecem essa garantia. Dessa forma, empresas capazes de oferecer esse serviço aumentarão sua participação no mercado.

Para o desenvolvimento da solução de Big Data, o estudo utiliza diversos recursos e ferramentas tecnológicas. A plataforma Hadoop é a base da arquitetura proposta devido à sua capacidade de processar grandes volumes de dados de forma distribuída e eficiente, característica essencial para lidar com os enormes volumes de dados gerados pelo maquinário agrícola. O Hadoop é uma plataforma open-source que permite o armazenamento e processamento distribuído de grandes conjuntos de dados usando clusters de computadores comuns. Sua arquitetura de software foi projetada para escalar de servidores únicos para milhares de máquinas, com um alto grau de tolerância a falhas, o que o torna ideal para grandes volumes de dados.

Ferramentas de análise preditiva são integradas para processar os dados coletados e prever falhas nas máquinas. Essas ferramentas utilizam algoritmos avançados de aprendizado de máquina para identificar padrões e antecipar problemas. Entre essas ferramentas, destacam-se o Apache Mahout e o Apache Spark MLlib, ambos projetos open-source que oferecem uma variedade de algoritmos de aprendizado de máquina escaláveis. O Apache Mahout, por exemplo, é usado para criar algoritmos de recomendação, classificação e clustering, enquanto o Apache Spark MLlib é uma biblioteca de aprendizado de máquina que fornece algoritmos comuns como regressão, classificação, clustering e filtragem colaborativa.

Além das ferramentas de análise preditiva, o estudo destaca a importância de ferramentas de processamento de dados em tempo real, como o Apache Storm e o Apache Kafka. O Apache Storm é um sistema de computação em tempo real que torna fácil processar fluxos de dados de forma confiável e escalável. Ele é especialmente útil para tarefas como análise em tempo real, aprendizado de máquina, detecção de fraudes e monitoramento de log. O Apache Kafka, por



outro lado, é uma plataforma de streaming distribuída que permite a construção de pipelines de dados em tempo real, processando e reprocessando fluxos de dados à medida que ocorrem.

Conforme apontado anteriormente, a coleta de dados foi realizada através de entrevistas com gestores e colaboradores, que possuem conhecimentos profundos sobre os sistemas atuais e as necessidades específicas da empresa. Também contou com revisões de literaturas, utilizando referências de diversos autores que exploraram e aplicaram o conceito de Big Data em diferentes contextos. Com essas bases teóricas e práticas, foi desenvolvida uma arquitetura específica para a implementação do Big Data na organização, desenhada para ser compatível com a estrutura existente e maximizar os benefícios da solução proposta.

A arquitetura sugerida incluiu componentes essenciais para a coleta, armazenamento, processamento e análise dos dados. Entre esses componentes, destacam-se os sistemas de gerenciamento de banco de dados NoSQL, como o HBase, que permite o armazenamento de grandes volumes de dados não estruturados e semi-estruturados. O HBase é usado para suportar operações de leitura e escrita rápidas e é integrado ao Hadoop para fornecer um armazenamento de dados distribuído e altamente disponível.

O estudo de caso apresentado por Felipe Renz demonstrou como a utilização estratégica de Big Data, através de plataformas como o Hadoop e análise preditiva, pode gerar vantagem competitiva significativa para empresas do setor agrícola. A implementação de uma arquitetura bem planejada de Big Data permite não só a prevenção de falhas e a minimização de prejuízos, mas também a oferta de um serviço mais confiável e eficiente aos clientes, aumentando a participação da empresa no mercado. A capacidade de processar e analisar grandes volumes de dados em tempo real permite que a empresa tome decisões mais informadas e estratégicas, otimizando seus processos operacionais e aumentando sua eficiência geral.