

16_빅데이터로 진화하는 미래 산업-스마트 팩토리

#1

1. 스마트 팩토리

가. 스마트 팩토리의 개요

제조업, 특히 하이테크 제조 분야에서는 인더스트리 4.0(Industry 4.0)과 스마트 공장(Smart Factory)과 같은 새로운 기술에 발 빠르게 대응합니다. 새로운 방법에 대한 모색을 통하여 끊임없이 변화하는 시장 요구사항에 대처해야 하기 때문입니다. 그렇기 때문에 혁신적인 제품을 더 빨리 제공할 수 있는 방법을 찾는 데 노력을 기울이고 있습니다.

이러한 노력과 더불어 사물인터넷 기반의 센싱(Sensing) 기술이 발전하면서 이전에 수집하지 못하던 다양한 데이터를 측정할 수 있게 되었습니다. 네트워크 기술의 발전을 통해 측정된 데이터를 빠르게 저장소로 이동시킬 수 있게 된 것입니다. 또한 하둡(Hadoop), 스파크(Spark)와 같은 오픈소스 기술의 발전으로 인해 적은 비용으로 대용량 데이터를 저장하고 빠른 속도로 분석이 가능하게 되었습니다.

#2

나. 공장 자동화(Factory Automation)와의 차이점

① 자동화된 공장은 일관된 공정에 따라 제품을 생산합니다. 반면에 스마트 팩토리에서는 각 공정 단계의 설비나 장비 간에 이루어지는 센서 정보의 교환을 통해 모듈별로 제조 공정을 달리할 수 있습니다.

② 기존 공장은 일정하게 규정된 생산 경로를 따라 자재를 운반해야 합니다. 반면에 스마트 팩토리에서는 생산 경로를 유연하게 활용하므로 자재를 운반하는 경로도 이에 맞추어 달라질 수 있습니다.

③ 기존 공장은 중앙집권적인 제어 체계를 가지고 있으나, 스마트 팩토리에서는 각 모듈이 자율적인 프로세스를 진행하는 수평 분산형 체계를 가집니다. 이때 각 설비와 자재의 상태 정보와 위치 정보를 실시간으로 공유합니다.

#3

2. 스마트 팩토리의 활용

데이터 마이닝(Data Mining), 머신러닝(Machine Learning), BI(Business Intelligence) 등의 데이터 분석 기술이 고도화됨에 따라 데이터 분석에 대한 관심이 커지고 있습니다. 기업이나 공장에서는 축적된 데이터를 활용하여 변화의 기회를 만들어 내거나 위험 요소를 줄이며 자산의 효율성을 높입니다.

제조 분야에서는 일찍이 설계 및 생산 과정에 자동화와 정보 시스템 등을 도입했습니다. 이에 따라 데이터를 분석하고 그 결과를 활용하여 제품의 품질과 공정의 효율성을 높이고 있습니다.

#4

가. 데이터 분석을 활용하는 스마트 팩토리의 도입

① 데이터 분석(Analytics)

제조 공정에 대한 예측과 분석을 통해 성과를 내는 기업이 증가하고 있는 추세입니다. 산업의 성숙화에 따른 시장의 포화와 기술의 평준화에 따른 하드웨어적 성장의 한계는 제조 기업에게 한계를 직면하게 했습니다. 이에 차별성 획득과 경쟁력 강화를 위해 데이터 분석의 도입을 고려하게 되었고, 많은 기업들이 성공 사례를 남겼습니다.

#5

② 클라우드(Cloud)

‘제조 속의 인공지능(AI in Production)’이라는 보고서에는 AI의 도입이 공장 운영 시스템에 효과적일 것이라고 설명합니다. 다만, 제조 공장이 인공지능을 도입하여 성과를 내기 위해서는 ICT 분야의 전문가들과 협업하는 것이 중요한데, 대부분의 제조 공정은 전문 인력이 없거나 부족한 상태입니다.

이러한 문제를 해결하기 위해 클라우드 스마트 팩토리의 구현을 선택할 수 있습니다. 클라우드를 통해 서버의 인프라, 플랫폼, 분석 인프라를 제공받을 수 있어 스마트 팩토리 도입의 진입 장벽을 낮출 수 있습니다.

#6

③ 플랫폼(Platform)

제조 기업들은 적정의 생산 계획, 생산성의 개선, 품질 향상 등을 해결하기 위해 산업인터넷이나 빅데이터, 인공지능 등의 기술을 적극 도입하고 있습니다. 이는 플랫폼 경제로의 전환을 암시합니다. 스마트 팩토리 구축의 핵심 열쇠는

플랫폼입니다. 플랫폼의 연결성을 통해 지식과 정보를 공유하며 협업 체계를 이루어 스마트 팩토리를 구축할 수 있습니다.

#7

나. 스마트 팩토리의 추진 이슈

① 커넥티드 팩토리(Connected Factory)의 구축이 요구된다

공정 간의 연계를 제어하기 위해 주요 설비의 공정 데이터를 실시간으로 통합하고 분석해야 합니다. 하지만 설비의 경우에는 제조사별로 포맷이 상이하기 때문에 이를 통합하고 실시간 데이터베이스를 운영하기 위해서는 대규모 투자가 필요합니다.

② 스마트 팩토리의 도입을 위한 이니셔티브를 확립해야 한다

커넥티드 팩토리의 구축이 선행되어야 하지만 막대한 투자비를 정당화할 수 있어야 하기 때문입니다.

③ 통합 미들웨어(Middleware)와 표준화를 위한 협의가 필요하다

설비 제조사별로 데이터 포맷이 달라 데이터 통합을 위한 미들웨어 개발에는 상당한 규모의 투자가 필요합니다.

#8

3. 성공적인 빅데이터 분석의 필요 조건

가. 처리 속도

① 기술의 한계

빅데이터 기술이 급격히 발전함에 따라 기존에는 가능하지 않았던 전체 데이터 분석을 수행할 수 있게 되었습니다. 그럼에도 빅데이터 분석이 확산되지 않은 이유는 증가한 데이터의 양만큼 데이터 분석에 걸리는 시간도 비례하였기 때문입니다.

② 빅데이터의 분석과 처리 성능의 중요성

빅데이터 분석이 실효성을 얻기 위해서는 대용량의 데이터를 저장하거나 단순 조회하는 것뿐만 아니라 예측 분석과 같은 고급 분석에 있어서도 기존처럼 짧은 시간 안에 결과를 얻을 수 있는 고속 분석 성능이 필요합니다. 이는 제조 현장의 불량 분석이나 공정 분석 등에 할애되는 시간이 짧아질수록 원가 절감 및 매출 향상에 막대한 영향을 주기 때문입니다.

#9

나. 상용화 가능성이 있는 기술

① R

오픈소스 R은 데이터 크기에 비례하여 실행시간이 급격하게 증가한다는 단점이 있습니다. 그러한 특성 때문에 대용량 데이터 분석에는 적합하지 않습니다. 이는 오픈소스 R이 실행되는 서버의 메모리 크기에 성능이 종속적이기 때문입니다. 따라서 대용량 제로 데이터의 분석에서는 대용량 데이터를 분석할 수 있는 툴이 필요함을 보여줍니다.

② SDS 브라이틱스(Brightics) AI

스파크(Spark)를 기반으로 병렬 및 분산처리가 가능한 대용량 분석 솔루션입니다. 데이터의 크기에 상관없이 성능을 유지할 수 있는 확장성(Scalability)을 기본적으로 제공합니다.

#10

③ 데이터 마트(Data Mart)

전통적인 분석에서는 분석용 데이터 마트(Data Mart)를 별도로 구축하고 이 데이터 마트에서 데이터를 가져와 분석을 수행합니다. 하지만 일부 현장에서 시시각각 쌓이는 데이터를 매번 데이터 마트로 구축하는 과정이 불필요해지기 시작하였고, 이를 해결하기 위해 각 분석 시스템이 데이터 레이크에서 데이터를 필요할 때 가져와서 분석하려는 시도를 하고 있습니다.

④ 데이터 레이크(Data Lake)

초기 단계인 데이터를 수집하고 저장하여 데이터 레이크(Data Lake)를 구축하는 것은 많은 기업에서 운영하고 있는 방식입니다. 데이터 레이크에서 데이터를 추출하여 각 분석 단계로 데이터를 흘려보낼 때 데이터 추출 속도와 저장 공간의 비효율성은 시급한 개선 사항으로 대두되고 있습니다.

#11

⑤ JDBC(Java Database Connectivity): 데이터베이스에 저장된 데이터를 빅데이터 분석 시스템으로 가져오기 위해 많이 사용되는 표준 인터페이스로는 자바(Java) 환경을 위한 JDBC(Java Database Connectivity)가 있습니다. 하

지만 이는 데이터를 분산하여 저장하고 병렬처리하는 빅데이터 분석 환경에 적합하지 않습니다.

⑥ 다중쿼리 방식

JDBC 방식의 문제점을 해결하기 위해 많은 소프트웨어들이 다중쿼리 방식을 활용합니다. 그러나 다중쿼리 방식 역시 쿼리에 대한 응답속도 보장에 한계가 있고, 복잡한 쿼리일수록 데이터베이스의 부담이 급격하게 증가하게 되는 이슈가 존재합니다. 또 서버의 수만큼 연결이 생성되기 때문에 데이터베이스의 가용성을 떨어뜨리는 점, 전체 데이터의 분산처리 능력을 높이거나 낮추기 어렵다는 점, 모든 서버가 데이터베이스와 연결이 가능해야 한다는 점 등이 문제점으로 지적되고 있습니다.

#12

다. 한계점 극복

제조 공장의 빅데이터 분석의 경우 원시데이터를 추출하는 쿼리를 조금씩 바꿔가며 특정 분석 모델을 수행하는 과정을 반복합니다. 이를 통해 의미 있는 분석 결과를 얻어내는 경우가 많습니다. 이때 원시데이터를 가져오는 추출기의 성능이 분석 모델을 반복해서 수행 할 수 있는 횟수와 직결되기 때문에 그 성능이 매우 중요한 요소입니다.

① 새로운 추출 기능

기존 데이터 처리의 문제점들을 개선하기 위하여 대용량 데이터를 위한 새로운 추출 기능이 필요합니다. 이 추출 기능이 제공하는 핵심 기술은 마스터 서버가 데이터베이스와 JDBC 연결을 생성합니다. 그리고 이들 서버 간에 주고받는 패킷데이터를 기반으로 가상의 데이터베이스 접속 환경을 제공하여, 슬레이브 서버(Slave Server)들에게 균등하게 데이터가 분산되어야 합니다.

② 로컬리티(locality)의 보장

마스터서버의 실제 쿼리 수행이 끝난 시점에 이미 데이터가 슬레이브 서버들에게 균등하게 분배되어 재균등화(rebalancing 및 repartitioning)와 같은 후처리 작업이 필요하지 않도록 데이터의 로컬리티(locality)가 보장되어야 합니다. 이는 추출 후 수행하는 분석 작업의 병렬처리 속도를 비약적으로 향상시키기 때문에, 전체적인 분석 모델의 수행 시간이 크게 감소하게 됩니다.

#13

4. 인공지능과 접목한 제조업

가. 제조업에서 인공지능 도입의 도전

제조 데이터는 직관적인 해석이 어려워 기술의 도입에 제약이 많습니다. 제조업 특성상 수율과 안전이 중요하여 인공지능 기술의 운영이 쉽지 않습니다. 인공지능 기술의 성공적인 도입 사례는 B2C 산업의 선례가 대표적입니다. 예를 들어, 구글 포토는 인공지능 딥러닝 기술을 이용하여 사진의 분류를 스스로 할 수 있게 되었습니다. 하지만 제조업과 B2C가 사용하는 데이터와 인공지능 기술의 도입 목적과 응용 분야는 다릅니다. B2C에서는 인공지능을 인간을 대체할 수 있는 작업을 대신하는 용도로 많이 활용합니다.

#14

나. 제조업에서 데이터 분석이 어려운 이유

제조업 데이터는 직관적인 분석이 어렵습니다. 제조업에서 흔히 사용되는 센서 값 그래프의 경우에는 해석하는 데 전문 지식이 요구됩니다. 일반 사람들은 데이터 그래프에서 정상 패턴과 다른 이상값을 찾아내기가 어렵다는 것을 뜻합니다. 데이터를 분석하기 위해 시간과 노력이 소요되고 관련 산업의 전문 지식을 갖추어야 하며, 데이터를 전처리해 분석하기 쉽게 형태를 변형할 수 있어야 합니다.

제조업의 특성상 공정과 제품의 복잡도가 지속적으로 증가하기 때문에 분석하고자 하는 문제의 난이도는 점점 더 높아질 것이며, 제조업의 데이터를 분석할 분석가도 부족한 실정입니다.

#15

질문자: 데이터 노이즈란 무엇인가요?

전문가: 부정확한 데이터가 섞이는 것을 데이터 노이즈라고 하며, 데이터 노이즈는 분석하는 데 어려움을 초래하게 됩니다. 제조업의 데이터는 장비에 달려 있는 다양한 센서에서 생성되며, 하루에도 수 기가바이트 분량의 데이터가 생성됩니다. 여기서의 문제점은 데이터가 항상 정확하게 생성되지 않는다는 것입니다. 데이터의 정확도를 떨어뜨리는 문제에는 통신 네트워크의 오류, 센서에 영향을 줄 정도의 높거나 낮은 온도, 센서 자체의 오류 등이 있습니다.