



특집 08

빅데이터 분석을 위한 기업의 Big Analytics 환경 변화



김근태 (SAS Korea)

목 차 »

1. 서 론
2. Big Analytics의 새로운 분석 패러다임
3. Big Analytics를 위한 High Performance Computing
4. 빅데이터와 Data Visualization
5. Text Analytics를 통한 Insight 확장
6. 결 론

1. 서 론

빅데이터는 기업들에게 다양한 측면에서 새로운 변화와 도전을 요구하고 있지만, 이는 곧 새로운 기회이기도 하다. 본 문서에서는 빅데이터를 분석하여 가치를 창출하기 위한 Big Analytics의 새로운 관점들에 대해 살펴보고, 어떻게 기업들이 Big Analytics를 통해 정확하며 적시성 있는 분석 Insight를 만들어 낼 수 있는지에 대해 다루어 보고자 한다.

1.1 빅데이터와 Big Analytics의 개념

빅데이터는 전통적인 환경에서 관리하거나 분석하기 어려운 규모의 데이터를 의미하며, 4V 속성을 통해 그 성격이 규정되고 있다. 빅데이터를 규정하는 4V 속성 및 분석 관점에서의 해결 과제는 <표 1>과 같다.

비즈니스적 관점에서 가장 중요한 부분은 결국 빅데이터로부터 얻어내 수 있는 가치 부분이 될 것이며, 결국 Big Analytics는 빅데이터로부터 의

<표 1> Big Data의 4가지 속성과 분석 관점의 관심 영역

Big Data 속성 (4V)	분석 관점의 해결 과제	비즈니스 적용 예
Volume (데이터의 규모)	현행 환경에서 다루기 힘든 큰 규모의 데이터를 어떻게 분석할 것인가?	전체 데이터 대상의 고객 프로파일링, 리스크, 가격 최적화 모델링등
Velocity (데이터의 주기)	어떻게 실시간에 가깝게 Actionable한 분석 Insight를 제공할 것인가?	실시간 Fraud 탐지, 마케팅 측면의 Real-time Promotion
Variety (데이터의 다양성)	어떻게 Text를 비롯한 비정형 데이터를 분석할 것인가?	Social Media Analysis, VOC Analysis
Value (데이터의 가치)	Big Data로부터 어떤 가치 있는 정보를 얻을 것인가?	가치를 창출하기 위한 비즈니스 예측 및 최적화 주제 선정

〈표 2〉 Big Analytics를 설명하는 다양한 관점들

분석의 속성	Big Analytics의 새로운 관점들
분석 데이터 크기	수십 Giga Bytes → 수십, 수백 Tera Bytes 이상의 데이터
분석 데이터 종류	구조화된 데이터 → 비정형 Text Data로 확장
분석 소요 시간	수 시간 이상 (Batch) → 수 초, 수 분 이내 (Near Real-time)
분석의 복잡도	특정 주제 위주 → 예측 및 최적화 관점의 요건들
수행할 분석의 양	수십개의 모델 → 수천~수만개 이상의 세분화된 모델
분석 사용자 환경	전문가 위주 → Data Visualization, 분석 모델링 자동화

미 있는 가치를 얻어내기 위한 다양한 접근 방법들을 의미한다고 정의할 수 있다. 물론, 빅데이터를 규정하는 개념이 상대적인 만큼 Big Analytics를 정의하는 개념도 절대적이지는 않지만, 전통적인 Analytics와 비교될 수 있는 Big Analytics의 새로운 관점들은 <표 2>와 같다.

1.2 Big Analytics의 비즈니스 적용 사례

2.5 Peta Bytes의 방대한 데이터를 EDW로 운영하고 있는 미국의 Catalina Marketing의 사례를 통해 Big Analytics가 비즈니스에 적용되는 모습을 엿볼 수 있다. 고객이 대형 할인마트의 매장 계산대에서 결제를 하는 순간 고객의 구매이력 데이터가 분석되어 이 고객을 위한 맞춤형 쿠폰이 인쇄된 영수증이 출력된다. 이때, 여러 소비자가 동일 제품을 구매해도 각자의 구매이력에 따라 각각 다른 쿠폰이 출력된다. 고객의 개인별 구매 성향이 실시간으로 반영된 쿠폰 제시를 통해 구매를 유도하는 것이다. 결과적으로 쿠폰 사용을 통한 추가 구매 비율이 기존 10%에서 25%로 크게 향상되었으며, 고객의 마케팅 반응율이 곧 막대한 수익 증가로 이어졌음은 물론이다.

이러한 결과를 얻기까지 Catalina Marketing은 DW 어플라이언스에 High Performance Computing 분석 기술을 접목하여 2.5 Peta Bytes의 EDW로부터 고객의 구매 성향을 실시간적으로 예측하였다. 즉, 고객에 대한 맞춤형 쿠폰 발행 과정에서 고객 행동에 대한 프로파일링을 실시간에 가깝게 수행하였으며, 기존에 4.5시간 걸리던 예측 모델에 대한 스코어링 시간도 불과 60초로 단축시켰다. 결론적으로 Catalina Marketing은 빅데이터를 대상으로 “Insight”의 속도 향상을 통해 고객 밀착 서비스에 성공한 사례라고 할 수 있다.

2. Big Analytics의 새로운 분석 패러다임

Big Analytics가 등장한 배경에는 빅데이터를 분석하기 위해 각종 자원의 활용을 최적화하여 적시에 활용 가능한, 정확한 분석 정보를 제공해야 한다는 니즈가 존재한다. 무엇보다 Big Analytics가 제공하는 다양한 측면의 새로운 관점들은 전통적인 분석 방법의 패러다임을 바꾸고 있다.

2.1 분석 가능한 데이터 크기 증가

데이터 저장과 관리의 측면에서 본다면, 어플라이언스 DW와 함께 오픈스스 기반의 분산 데이터 저장 기술인 Hadoop이 기존 DB로 관리하기 힘든 규모와 성격의 데이터를 처리하기 위한 기술로 주목받고 있다.

여기에 데이터 분석의 관점에서 주목해 볼 만한 사실은, 분석의 가치를 극대화 할 수 있는 Advanced Analytics 영역이 High Performance Computing (이하 HPC) 기술과 밀접하게 결합되면서 본격적으로 빅데이터 분석에 대응하고 있다는 점이다.

Advanced Analytics 영역에 HPC 기술이 도입

됨으로써 분석가가 얻을 수 있는 첫 번째 혜택은 분석 가능한 데이터의 한계에 대한 극복이다. 이는 곧, 그동안 모델링 과정에서 Data 크기의 문제를 극복하기 위해 활용되었던 데이터 샘플링 과정을 버릴 수 있게 되었음을 의미한다. 수억건 이상의 행과 수만개 이상의 열(변수)을 가진, 그동안 다루기 어려웠던 데이터를 별다른 변환 없이 그대로 분석 작업에 활용 할 수 있게 되었다는 점은 분석가의 고민을 줄여줄 뿐만 아니라 모델의 정확도를 높이는 주요한 요소로 작용할 수 있다.

2.2 분석 모델의 세분화

Big Analytics의 Big 이라는 용어와 상반되게 여겨질 수 있겠지만, Big Analytics가 제시하는 또 하나의 큰 변화는 분석 모델을 매우 작은 단위로 세분화하여, 보다 정교한 분석 결과를 얻을 수 있도록 한다는데 있다.

일반적으로 예측 모델(Predictive Model)은 전체 데이터를 대상으로 통합되어 적용되는 것 보다 세분화된 차원의 데이터에 적용될 때 훨씬 좋은 효과를 나타낼 수 있다. 각 개인 고객별로 적용되는 고객 이탈 모델, 모든 개별 제품라인에 별도로 적용되는 부품 불량 예측, 각 계좌별로 적용되는 구매 예측 등이 그 좋은 예가 될 것이다. 또한, 최근 들어서는 각각의 세분화된 예측 결과를 실시간적으로 업무 프로세스에 적용해야 하는 비즈니스 니즈가 점차 많아지고 있다.

2.3 분석 라이프 사이클 단축

전통적인 분석 프로세스는 데이터를 준비하고 탐색하고, 모델을 개발하고, 운영 환경에 적용하는데 수일에서 수주 이상의 많은 시간을 소모하는 경우가 많았다. Big Analytics 환경에서는 모

델의 생성에서 적용에 이르는 전체 분석 라이프 사이클에 HPC 기술을 적용함으로써 분석의 효율성을 극대화하게 된다.

예컨대, 분석가는 모델링 단계에서 매우 빠르게 수천개 이상의 변수로부터 유의 변수를 선정하거나, 모델의 정확도를 높이기 위해 서로 다른 모델링 기술을 즉시 실험적으로 적용해 보고, 기존 모델에 새로운 관점을 반영해 볼 수도 있게 된다. 또한 복잡한 마이닝 프로세스 대신 자동화된 방법으로 예측 모델을 생성함으로써 모델 생성 시간을 몇분 이내로 단축시키거나, 빠르게 생성된 분석 모델을 고성능 어플라이언스 DW에 자동으로 Deploy 하고 즉시 운영업무에 반영 할 수도 있다. 결국 분석 모델의 생성과 적용 전반에 걸쳐, 이전에는 생각하기 어려웠던 실시간 프로세스 구축이 가능해진 것이다.

2.4 비정형 Text Data로 분석 인사이트 확장

인터넷, SNS, Mobile과 같은 개인화 된 환경에서 통해 쉽없이 쏟아져 나오는 Data의 폭발적 증가는 빅데이터라는 화두를 만들어 낸 주요 원인이라고 볼 수 있다. 이와 맞물려 그동안 구조화된 데이터를 대상으로 삼았던 분석 기술도 비정형 Text 데이터 영역으로 급속히 확장되고 있다.

많은 기업들이 그동안 적지 않은 마케팅 비용을 투자하면서도, 정작 고객의 솔직한 생각을 읽는데 많은 어려움을 겪어 왔다. 따라서 기업들이 고객의 숨김없는 생각이 그대로 드러나 있는 인터넷, SNS의 분석에 적극적인 관심을 가지게 된 것은 매우 자연스러운 현상으로 볼 수 있을 것이다.

또한 분석의 측면에서 볼 때 비정형 Text Data 영역으로 대상이 확장되었다는 것은, 사전 정의된 테이블을 대상으로 한 분석이 가졌던 제한적

인 분석 인사이트의 한계를 상당부분 벗어날 수 있게 되었음을 의미한다.

3. Big Analytics를 위한 High Performance Computing

분석 관점의 HPC 기술 중, MPP 기반의 DW 어플라이언스 혹은 In-Memory 기반 OLAP은 전통적인 BI 영역에 대한 성능 문제를 해결해 주는 기술로 볼 수 있다.

여기서는 기업들이 빅데이터로부터 궁극적으로 얻고자 하는 예측 및 최적화 정보를 제공해 주는 Advanced Analytics 영역이 어떻게 다양한 HPC 기술들과 접목되고 있는지 살펴보고자 한다.

3.1 Grid Computing

다수의 컴퓨터를 연계하여 성능을 극대화하는 Grid Computing 기술은 그다지 새로운 기술로 볼 수 없지만, 복잡한 Advanced Analytics 작업이 Grid 환경과 결합하여 운영될 수 있다는 사실에 주목해 볼 필요가 있다.

다양한 과정들이 연계된 큰 분석 프로세스는 여러 대의 분석 서버가 연계된 Grid Computing 환경 내에서 분산되어 처리되거나, 다른 여러 개의 작업 사이에서 효율적으로 로드 밸런싱 될 수 있다. 또한 사용자나 작업의 증가에 따라 Grid 노드의 증설을 통해 안정적인 확장성을 가져갈 수도 있다.

Grid 분석 환경으로부터 체감할 수 있는 가장 기본적인 장점은 분산 환경이 제공하는 분석 성능의 향상이다. 예컨대, 대용량 데이터에 대응하기 위한 목적으로 Grid 분석 환경을 구축한 한 금융 업체는 고객의 재무 상태에 대한 예측 및 시뮬레이션 분석 시간을 96시간에서 4시간 이하로 단

축시킨 바 있다. 그밖에 Grid 아키텍처와 결합된 분석 환경은 다음과 같은 가치를 제공한다.

- Distributing Enterprise Job Scheduling - 많은 사용자 및 분석에 대한 작업 분산
- Workload Management - 분석 작업에 대한 병렬처리 및 부하 관리에 의한 성능향상
- High Availability - 실시간 비즈니스 연계가 필요한 분석 환경의 안정성 보장

3.2 In-Database Analytics

전통적으로 데이터를 담고 있는 DBMS와 마이닝등을 수행하기 위한 통계 분석 서버는 서로 다른 물리적 환경에서 운영되어 왔다. 그러나 최근 들어 데이터의 크기가 급격히 커지는 과정에서 이러한 아키텍처도 많은 문제를 나타내게 되었다. 즉, DBMS와 분석 서버간 데이터 다운로드에 따른 시간 소요, 네트워크 리소스의 과부하 유발은 물론, DBMS와 분석 서버간 데이터의 과도한 중복성 및 동기화 문제 등이 드러나게 된 것이다.

이러한 문제들을 극복하기 위해 새로운 대안으로 등장한 개념이 바로 In-Database Analytics이다. In-Database Analytics는 MPP 어플라이언스

〈표 3〉 In-Database 분석 기술

Analytic Sandbox	분석가에게 DB 공간을 할당해 주고, 분석에 필요한 데이터 가공이나 탐색 작업을 DW 내에서 처리하도록 함으로써 전반적인 분석 시간을 단축하고 데이터 이동을 줄이는 개념
Analytic Accelerating	SQL과 구분되는 기초 통계는 물론 회귀, 상관 분석등의 통계 분석 기능의 일부 또는 전부를 DW 내부에서 처리함으로써 데이터 이동을 줄이고 성능을 향상시키는 개념
Scoring Accelerating	마이닝 툴에서 생성된 예측 모델을 자동화된 방법으로 DW내에 사용자 할수의 형태로 만든 후, SQL을 통해 DW 내에서 직접 예측 값을 조회(Scoring) 하는 개념

DW가 제공하는 강력한 데이터 처리 성능을 단순한 Query뿐만 아니라 마이닝과 고급통계 분석에 활용하는 기술이다. In-Database Analytics의 관점에서 접근할 수 있는 기술적 요소들은 다음과 같다.

결과적으로 In-Database 분석 기술은 DBMS와 분석 서버간 데이터의 이동을 크게 줄이고 분석의 성능을 크게 향상시킬 수 있으므로 Big Analytics의 유용한 기술적 방안이 될 수 있다.

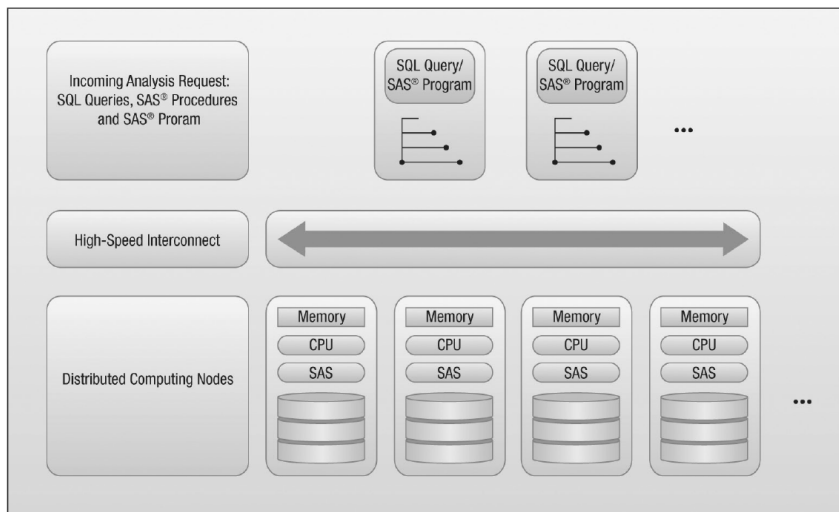
3.3 In-Memory Analytics

In-Memory Analytics는 각종 고급통계 분석 함수 실행 또는 마이닝과 같은 Advanced Analytics 작업들을 강력한 하드웨어 기반인 MPP DW 어플라이언스 내에서 직접 수행되도록 함으로써 분석 성능과 효율성을 극대화 하는 개념이다. 각종 통계분석이 MPP DW의 분산된 노드의 In-Memory에서 직접 수행되며, 이 과정에서 분석 가능한 데이터의 크기 (레코드 및 컬럼 수)의 제약도 해소할 수 있다. 결과적으로 빅데이터의 크기 문제에

대응하고 강력한 성능을 보장할 수 있다는 측면에서 Advanced Analytics를 지원하는 High Performance Computing의 가장 진보된 개념이라고 볼 수 있다. 확장성 있는 아키텍처 구조의 In-Memory Analytics는 다음과 같이 분석 라이프 싸이클 전반에 걸쳐 적용이 가능하다.

- High-Performance Data Exploration (기술 통계, 데이터 탐색 및 요약)
- High-Performance Model Development (일반 모델링 기법 및 변수 선택)
- High-Performance Model Deployment (모델 비교 및 스코어링)

실제로 High-Performance Analytics 환경을 구축한 금융회사는 예측 모델을 생성하고 기존 SMP 환경에 적용하기까지 167시간 걸리던 작업 시간을 불과 84초로 단축시킨 바 있는데, 이는 곧 비즈니스 예측 모델의 생성이나 적용 과정을 비즈니스 운영 프로세스의 일부에 실시간적으로 적용할 수 있게 되었음을 의미한다.



(그림 1) MPP 어플라이언스 기반 In-Memory 분석 아키텍처

4. 빅데이터와 Data Visualization

빅데이터 개념이 등장하면서 대용량의 수많은 속성들을 가진 데이터를 분석하기 위한 기술들이 새롭게 주목받고 있다. 사용자 관점에서 볼 때, 직관적인 화면에서 빅데이터를 가장 쉽고 빠르게 이해할 수 있는 방법이 바로 Data Visualization이다.

4.1 Data Visualization

시장조사 기관인 가트너에서는 현재 BI 툴을 이용하는 기업 사용자가 전체 잠재 사용자의 30%에 불과하다는 수치를 제시하고 있다. BI 툴의 활용도가 낮은 현상에는 여러 가지 원인이 있겠지만 사용하기 너무 어렵고, 대응이 늦거나, 관련성이 떨어지는 콘텐츠를 제공하는 등의 이유를 들 수 있다.

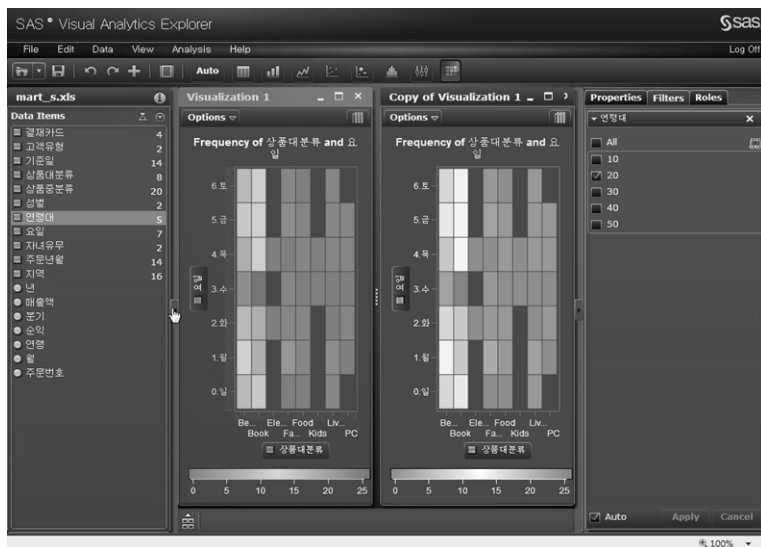
이에 대한 대안적 측면에서, 사용자 지향적인 Data Visualization 툴은 사용자들이 데이터를 편안하게 해석할 수 있는 시각적 체험을 제공한다.

즉, 사용자가 원하는 항목들을 선택하는 것만으로 거기에 알맞은 시각화된 결과를 보여줌으로써 데이터에 숨겨진 패턴을 파악할 수 있도록 해 준다. 이런 방식은 비전문적인 사용자가 이러한 툴을 매우 직관적으로 자연스럽게 사용할 수 있는 것으로 받아들여도도록 해준다.

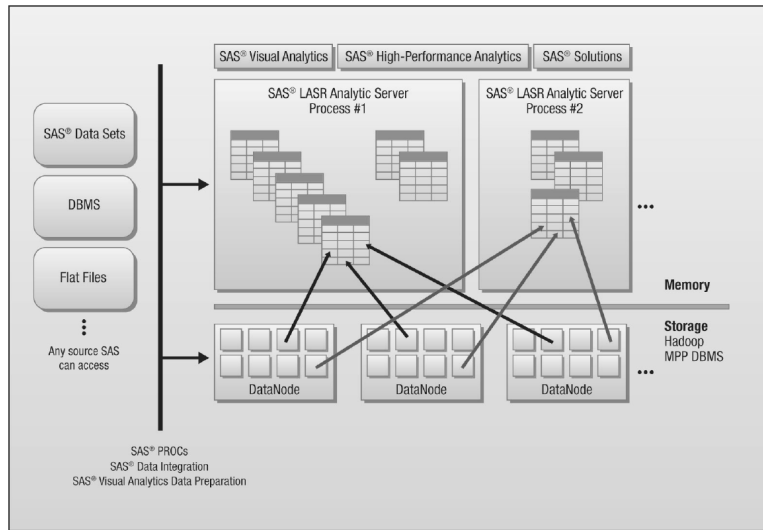
4.2 빅데이터에 대응하기 위한 Data Visualization 툴 요건

Data Visualization 개념이 자유로운 분석의 관점에서 효과적으로 빅데이터 분석에 적용되기 위해서는 아래와 같은 요건들이 만족되어야 한다.

- Speed Free : 수억건 이상의 데이터를 인터랙티브하게 분석할 수 있는 In-Memory 기반 성능 제공
- Ad-Hoc Free : 다양한 시각화 기법을 혼용한 직관적인 비정형 분석
- Dimension Free : 사전 정의가 불필요한 다차원 분석 및 드릴다운 분석



(그림 2) 사용자 지향적인 Data Visualization 분석



(그림 3) Big Data 분석을 위한 In-Memory 분산 아키텍처

- Analytics Free : 회귀, 상관분석 등 Advanced Analytics 영역의 시각화
- Source Free : Hadoop을 포함한 DBMS 및 PC 파일에 대응

Data Visualization 기능이 빅데이터를 지원하는데 가장 핵심적인 문제는 Speed Free의 관점이라고 할 수 있다. 반복적인 분석을 하는 과정에서 원하는 시각화 결과가 Memory 기반에서 수초 이내에 나오지 않으면 사용자는 더 이상 Visualization 툴을 사용하지 않을 것이기 때문이다. 따라서 빅데이터 환경에서 Data Visualization 툴이 자리매김 하기 위해서는 데이터의 폭발적인 증가에 유연하게 대응하여 실시간 응답속도를 보장할 수 있는, 확장성 있는 In-Memory 기반의 분산 Architecture가 꼭 필요하다고 할 수 있다.

5. Text Analytics를 통한 Insight 확장

Text Analytics는 문서 기반의 정보를 추출하고 이를 해석, 마이닝 및 구조화하여 문서간 패턴,

감정 및 관계를 밝혀냄으로써 텍스트의 정보 가치를 극대화할 수 있는 기술들을 의미한다. 최근 들어서는 한글 문서에 대한 분석의 정확도가 크게 향상되고 SNS가 주목 받으면서, 국내에서도 여러 분야에서 Text Analytics가 활발히 검토되거나 적용되고 있다. 효율적인 Text Analytics 환경을 구성하기 위해서는 소스 문서를 추출 및 클리닝을 자동화하는 사전 단계가 선행되어야 하며, Text를 분석하여 활용하고자 하는 목적에 따라 다양한 분석 기법이 고려되어야 한다.

5.1 Text Mining

Text Mining은 문서를 파싱하여 분류하거나 문서 간의 연관성을 파악하고 카테고리 별로 문서에 대한 클러스터링을 수행하는 개념이다. 예컨대, 문장 안에서 자주 활용되는 단어들 간의 연관성을 시각적인 그래프를 통해 분석함으로써 전체적인 문맥을 분석하거나 글쓴이의 의도를 시각적으로 파악해 볼 수 있다. Text Mining의 또다른 중요한 가치는 Text 분석의 결과 데이터

를 마이닝 예측 모형을 위한 인자로 활용하고, 궁극적으로 모델의 정확도를 향상시킬 수 있다는 점에 있다.

5.2 Content Categorization

자동화된 Content Categorization 기능을 통해 문서 형태로 된 정보들을 재구성하거나 액세스 및 검색을 수행하는데 활용할 수 있다. 또한 매우 긴 문장의 텍스트를 요약하거나 중복 문서를 확인할 수도 있다. 사용자가 의도하는 관점으로 문서를 분류하기 위해서 사전 모델링 단계를 거칠 수도 있으며, 이를 통해 분류의 정확도를 높일 수 있다.

실제로 신문사에서 기자들이 송고한 기사를 정치, 경제, 사회, 문화 섹션으로 자동으로 분류한 다거나, 선거 과정에서 SNS에 올라온 유권자들의 수많은 의견들을 다양한 관점에서 분석하여 선거 판세를 예측하는데 Content Categorization 기능이 활용되고 있다.

5.3 Sentiment Analysis

Sentiment Analysis는 인터넷, SNS 및 다양한 문서에서 자동으로 글쓴이의 감정(Sentiment)을 파악 및 추출하는 기술들을 제공한다. 즉, 글쓴이의 의도를 어떤 대상에 대한 긍정, 중립, 부정의 측면으로 분리하여 파악할 수 있으며, 좀 더 세분화된 결과 활용을 위해 감성에 대한 스코어링 기법이 적용될 수도 있다.

특히 Sentiment Analysis 기술은 요즘 많은 관심을 받고 있는 Social Media Analysis의 기반 기술로 활용되고 있는데, 이를 통해 기업들은 다양한 관점에서 브랜드 이미지를 모니터링 한다든지 상품에 대한 고객의 피드백으로부터 개선 포인트

를 찾아 낼 수 있으며, 시계열적 측면에서 고객 의견의 변화를 추적 할 수 있다.

6. 결론

이상으로 빅데이터를 효과적으로 분석하고 가치를 얻어내기 위한 Big Analytics의 다양한 관점들에 대해 살펴보았다. 마지막으로, Big Analytics를 도입하여 경쟁우위를 갖고자 하는 기업이 종합적으로 고려해야 할 요소들은 다음과 같다.

- 빅데이터에 대응하기 위한 분석 인프라 (HPC, Text Analytics, Visualization)
- 분석모델의 운영이 배치에서 리얼타임으로 변경됨으로써 분석 활용의 패러다임 변경
- 보다 다양한 측면의 의사결정 지원 - 장기적 전략과 전술, 그리고 운영의 측면
- 다양한 환경 지원 - 분석가의 실험적인 작업을 위한 Sandbox, 목적 지원용 Data Mart, 대용량 데이터 저장 및 복잡한 분석을 위한 어플라이언스 등
- 분석 리더십 확보 - 분석 전문 CoE 조직, 분석-IT-비즈니스를 연계할 수 있는 Data Scientist 발굴
- 전사 규모로 확장 - 사용자 및 데이터 규모 확산, 다양한 비즈니스 시나리오 분석 대응
- 새로운 기술 접목 - In-Memory, Hadoop, MapReduce

데이터 그대로의 빅데이터, 또는 빅데이터 관리 기술로 주목받고 있는 DW 어플라이언스나 Hadoop과 같은 도구만으로는 가치 있는 정보를 창출하는데 명확한 한계를 가질 수 밖에 없다. 데이터로부터 어떤 가치 있는 것을 발견하고자 하

는지, 그리고 어떻게 데이터를 처리해야 그러한 정보를 얻을 수 있는지에 대한 고민과 함께, Big Analytics 적용을 위한 종합적인 접근이 체계적으로 병행되어야 할 것이다.

저 자 약 력

참 고 문 헌

- [1] McKinsey ; Big data: The next frontier for innovation, competition and productivity, May 2011.
- [2] IDC :New Analytics Strategies in the 'Big Data' Era, Sep., 2011.
- [3] IDC : Big Data Analytics : Future Architectures, Skills and Roadmaps for the CIO, Sep., 2011.
- [4] TDWI :Big Data Analytics, Oct 2011.
- [5] SAS :In-Memory Analytics for Big Data, Jan., 2012.

김 근 태

이메일 : keun-tae.kim@sas.com

- 1997년-2006년 Oracle Korea, DBMS, BI/DW, CRM 담당 / 컨설턴트
- 2006년-현재 SAS Korea, Business Analytics Framework 담당 / 수석 컨설턴트
- 관심분야: Business Analytics, Big Data & Big Analytics