빅데이터란?

빅데이터(big data)란 대규모, 빠른 생성 속도, 다양한 형태라는 특징을 가진, 기존의 정보 관리 기술로는 저장/관리/분석하기 어려울 정도의 큰 규모의 정형 또는 비정형 데이터를 의미한다. 또한 그러한 데이터를 저장/관리/분석하는 기술도 빅데이터라는 개념에 포괄된다. 빅데이터 현상은 기업들의 고객 데이터 수집 활동 및 멀티미디어 콘텐츠의 폭발적 증가와 스마트폰 보급, SNS 활성화 및 사물 통신망의 저변 확대로 빠르게 확산되고 있다.

특히 기업들은 온라인/오프라인 사용자 정보, 소비자 행태에 대한 정보 수집, 적극적 고객 관련 정보 수집의 증가로 더 많은 데이터 스토리지와 정교한 분석 능력을 필요로 한다. 예를 들면 영국의 대형 유통 업체인 테스코(Tesco)는 매달 15억 건 이상의 (고객) 데이터를 수집하고, 고객의 소비 행태에 대한 대량의 데이터를 분석해 고객 맞춤형 서비스를 제공한다. 고객은 이제 무엇을 살지 고민하지 않아도 된다. 유통업체가 고객이 필요한 것, 좋아하는 패턴, 냉장고 속의 상황을 판단해서 모든 것을 알려주는 것이다.

4가지 데이터 분석 모델

• Descriptive Analytics

과거나 현재에 어떤 일이 일어났는지 파악하는 분석이다. 즉, “What happened?”를 보며 상황을 모니터링한다. 예를 들어, 지난 1년간 지역별로 무슨 제품이 많이 팔렸는지, 언제 어떤 제품의 판매량이 높았는지, 고객의 불만 내용이 무엇인지 파악한다. 단지 소비자의 긍정적/부정적 반응 뿐만 아니라, 무엇을 얼마나 좋아하는지 상황을 정확히 묘사한다.

• Diagnostic Analytics

원인을 밝히기 위한 분석이다. 즉, “Why did it happen?”을 보며, 데이터 간의 인과관계를 찾는다. A 제품이 왜 이 지역에서 잘 팔렸는지, 분기별로 매출 차이가 발생하는 이유는 무엇인지, 고객 불만이 최근 1년간 증가했다면 그 이유는 무엇인지를 원인을 이해한다.

• Predictive Analytics

미래에 어떤일이 어느 정도의 확율로 일어날지를 예측한다. 기계학습 모델을 사용하여, “What will happen?”을 분석한다. 예로, 기업이 사용자들의 단어 사용패턴을 분석하여 특정 고객의 이탈 가능성을 예측하기도 하고, 특정 제품의 검색 또는 구매 이력을 토대로 구매가능성이 높은 제품을 추천하기도 한다. 이를 통해 기업은 데이터 기반의 성장을 할 수 있게 된다.

• Prescriptive Analytics

무엇을 해야할지 솔루션을 도출하기 위한 분석이다. 최적의 성과를 낼 수 있도록 방법을 처방한다. 예를 들어, 어떤 시점에서 어떤 액션을 취해야 고객이탈을 막을 수 있는지, 어떤 채널에 어느정도 예산을 배분해야 광고 효과를 최대화 할 수 있는지, 어떤 종목에 투자해야 높은 수익률을 기대할 수 있는지가 해당된다.