

# 농업 경영전략 개발을 위한 빅데이터 플랫폼 설계



뉴엔 반 퀴엣  
전남대학교  
전자컴퓨터공학부/박사과정  
quyetict@utehy.edu.vn



김 미 선  
전남대학교  
전자컴퓨터공학부/석사과정  
misun\_kim05@hanmail.net



양 형 정  
전남대학교  
전자컴퓨터공학부/교수  
hjang@jnu.ac.kr



김 경 백  
전남대학교  
전자컴퓨터공학부/교수  
kyungbaekkim@jnu.ac.kr

## 1. 서론

최근 들어 세계적으로 자연재해가 빈발하고 농산물 시장개방의 진전으로 가격변동이 심화되면서 농업경영에서 불확실성과 위험이 더욱 증가되고 있다. 농업의 규모화 및 자본화가 진전됨에 따라서 경험과 영농지식이 지배했던 과거의 관행농업을 탈피하여 농업경영 및 농식품의 효과적 마케팅을 위한 빅데이터 기반의 합리적 의사결정 기법 개발이 중요하게 인식되고 있다. 이에 농업 분야에서 생산되는 각종 정부통계, 기상정보, 생산자 소비자패널, 소득조사정보 등 방대한 데이터(정형, 비정형)를 활용하여 미래 농업이슈 발굴, 경영 및 마케팅 의사결정 전략 수립 및 농업현장에 적용하는 방안에 대한 연구의 필요성이 강하게 제기되고 있다. 본 기고에서는 이러한 스마트 농업경영을 위한 전략개발을 위해 필수적인 농업빅데이터 기반의 플랫폼 설계 시, 고려할 사항들에 대해 살펴보고자한다.

## 2. 연구 및 기술의 국내외 현황

농업경영 및 마케팅 분야의 빅데이터 관련 연구는 아직까지 시작단계로 구체적인 연구보다는 그 필요성 및 활용방안에 관한 논의가 이루어지고 있는 실정이다.

일반기업의 빅데이터 활용 전략은 주로 마케팅 전략에 초점이 맞추어져 있다. 예를 들어, 이우재 등(2014)은 KB국민카드의 내부 및 외부의 빅데이터를 결합한 실시간 마케팅과 모니터링을 위한 시스템 구축작업을 소개하고 있다. 카드사용 고객의 라이프 사이클을 분석해서 유사하게 나타나는 패턴을 예측하여 이에 실시간으로 대응하는 빅데이터 활용전략을 제시하고 있다[1].

마찬가지로, 현재까지의 농업 관련 빅데이터 연구는 대부분 식품정보 분석을 통한 마케팅 정보 활용방안에 관한 논의가 대부분이다. 임종태 등(2014)은

빅데이터 분석을 이용한 식품 정보 분석 시스템을 설계, 구현하였다. 이 연구에서 제안한 시스템은 식품에 대한 정보를 분석하여 실생활에서 구매자에게 유용한 서비스를 제공한다[2].

또한, 농업기술분야에서 ICT 농업 및 스마트 팜 구현을 위한 통계체계 구축 방안으로서 빅데이터 분석의 필요성 및 활용법에 대한 논의가 진행되고 있다. 농림축산식품부도 2014년 빅데이터 분석기반 마련에서 2015년 빅데이터 활용 중심으로 본격 전환 추진하려는 계획을 수립하고 있다. 예를 들어, 빅데이터, 클라우드컴퓨팅, 사물인터넷 등 ICT 신기술간 연계 및 융합, 농식품부 스마트농정 통계체계 구축 등 각 부처 빅데이터 사업을 추진하고 있다.

### 3. 농업 빅데이터 플랫폼 설계

산업을 발전함에 따라 매일 새로운 데이터가 쌓이고 있다. 예를 들어 농업 부분에서는 기상데이터, 토양 데이터, 농가데이터, 수확 및 소비 등 다양한 농업 데이터가 쌓이고 있다. 이러한 데이터들을 이용하여 농민과 농업 관련기업에 가치 있고 유용한 정보를 찾아낼 수 있다. 그러나 이러한 데이터들을 사용하기 위해서는 우선 엄청난 양의 데이터들을 수집하고 저장하는 것이 중요하다. 일부 시스템은 저장할 수 있는 데이터 유형이 제한되어 여러 종류의 농업데이터를 가공 및 분석하는데 적합하지 않기 때문에, 수집되는 데이터의 특성을 고려하여 수집 및 저장 구성요소를 잘 설계해야 한다. 또한, 농업 빅데이터 분석 시 자주 사용되는 기능들에 대한 고려를 통해, 보다 효과적인 분석을 수행할 수 있도록 플랫폼의 주요 구성요소들을 설계해야 한다.

### 3.1 설계된 농업 빅데이터 플랫폼 구조

그림 1은 설계된 농업 빅데이터 플랫폼의 주요 구조의 개요를 나타낸다. 맵리듀스(MapReduce) 기반 모듈, 플럼(Flume) 모듈, 스콥(Sqoop) 모듈을 이용하여 다양한 형태의 데이터를 병렬 방식으로 수집한다. 하둡(Hadoop)기반의 데이터 클리너(Data Cleaner) 모듈을 사용하여 중복 데이터를 제거하고 불규칙한 데이터를 저장한다.

하둡 분산 파일 시스템(Hadoop Distributed File System, HDFS)은 하둡 프레임워크를 위해 자바 언어로 작성된 분산 확장 파일 시스템이다. HDFS은 여러 기계에 대용량 파일들을 나눠서 저장을 하게 되고, 데이터들을 여러 서버에 중복해서 저장하여 안정성을 높인다. 즉, 빅데이터를 분산해서 저장하기 때문에 확장성 및 신뢰성이 보장되고 대규모 클러스터 서버를 이용해 방대한 데이터를 신속하게 랜덤 액세스가 가능하다. 설계된 빅데이터 플랫폼에서는 기본 저장소인 HDFS 이외에 빅데이터를 위한 저장 프레임워크인 HBase와 많은 양의 고속 랜덤 액세스

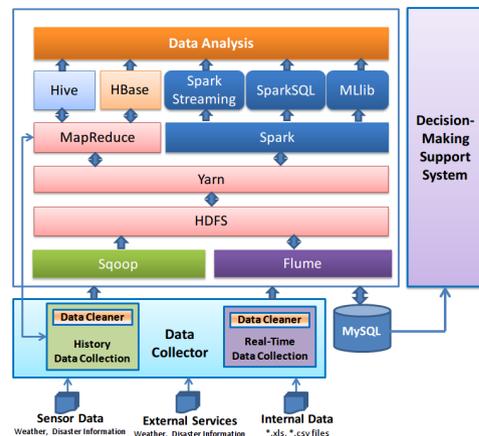


그림 1. 설계된 농업 빅데이터 플랫폼 구조의 개요[3]

스가 가능한 또 다른 저장 프레임 워크인 Hive를 포함하여 데이터 저장을 수행한다.

이러한 저장 프레임워크들은 데이터 웨어하우스 인프라 도구로 구조화된 데이터를 처리하고 큰 데이터를 요약하기 위해 하둡 위에 위치하며, 쉽게 쿼리를 실행할 수 있다. 또한, Sqoop을 사용하여 CSV에서 데이터를 가져오거나, xls 파일을 HDFS, Hive 또는 HBase로 변환할 수 있다.

저장된 데이터들을 활용해 Spark 및 Hadoop 기반의 Data Analysis 모듈을 통해, 클러스터를 만들거나 데이터를 분류하는 알고리즘, K-mean 또는 Naïve Bayes와 같은 다양한 기계학습 및 패턴인식 작업을 수행할 수 있다. 이러한 분석 작업의 결과는 빅데이터 플랫폼 저장소에 저장되어 분석 작업에 재사용되거나, MySQL과 같은 관계형 데이터베이스에 저장하여 사용자들이 편하게 접근하게 할 수 있도록 한다.

의사 결정 모듈은 데이터 분석 모듈로부터 결과를 수신하고 사용자가 농업경영을 효율적으로 할 수 있는 유용한 결정을 내리는데 도움을 주거나 데이터 저장소에 저장된 통계 데이터를 제공한다.

### 3.2 데이터 수집 모듈

농업 빅데이터를 수집할 때는 데이터의 특성을 고려하여 수집 작업을 수행해야 한다. 광주 서부 시장의 웹사이트에는 매일 약 500페이지의 농산물의 데이터가 업데이트된다. 반면, 농업관련 웹사이트에는 200만건 이상의 아카이브된 데이터가 존재한다. 이처럼, 매일 실시간으로 업데이트 되는 데이터와 대용량의 데이터를

효과적으로 수집하기 위해서는 데이터의 특성을 고려한 수집 방법이 필요하다.

매일 약 500페이지가 업데이트되는 농업관련 웹사이트의 데이터의 수집에 전통적인 데이터베이스 시스템을 적용하는 것은 적합하지 않다. 이러한 연속적이고 높은 가용성을 가진 데이터는 실시간 수집기능을 가진 모듈을 통해 수집되어야 한다. 이를 위해, 병렬적이고 실시간으로 데이터를 처리할 수 있는 하둡 분산 파일 시스템(Distributed Files System, HDFS)과 Flume을 사용해 실시간으로 데이터를 모을 수 있도록 설계하였다. 시스템 또는 서비스의 크롤링 프로세스를 간소화하고 Flume의 스트리밍 데이터 처리를 위해 실시간 데이터를 효율적으로 수집하고 대량의 데이터를 HDFS로 주기적으로 전송한다. 그림 2는 Flume을 사용하여 데이터를 수집하기 위한 설계를 보여준다. Apache flume은 장애가 나더라도 로그, 이벤트를 유실 없이 전송할 수 있도록 설계되었으며, 수평 확장(Scale-Out)이 가능하여 분산수집이 가능하다. Apache flume은 크게 Flume을 통해 전달되는 데이터의 기본 부분인 Event와 그림 2의 Apache flume 그룹의 Agent 부분(Source, Channel, Sink)으로 구성된다. Source에서 Event를 수집하여 Channel로 전달한다. Channel은 Event를 Source와 Sink로 전달하는 통로로써 세 가지 타입이 존재하는데, 여기서는 간편하고 빠른 고성능을 가진 Memory 타입을 사용한다. 이렇게 전달된 데이터는 Sink에서 HDFS에 저장된다[4].

빅데이터 분석을 위해서는 실시간 데이터는 물론 아카이브 데이터도 필수적으로 수집해야 한다. 현재 농업관련 다양한 웹사이트에 올라와 있는 아카이브 데이터들은 실제 농업 웹 시장과

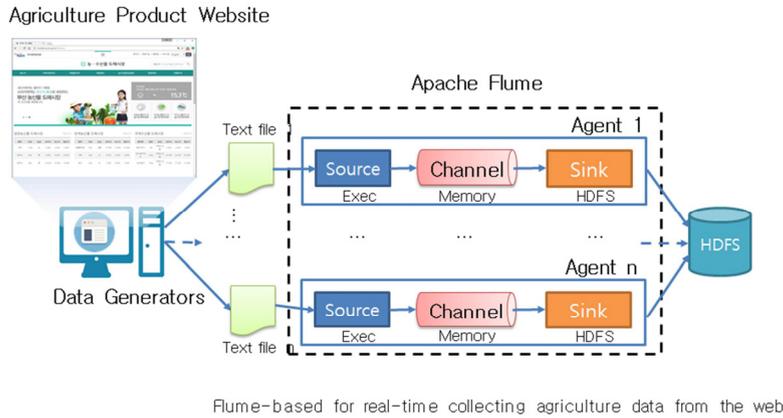


그림 2. Apache flume을 이용한 실시간 데이터 수집 및 저장

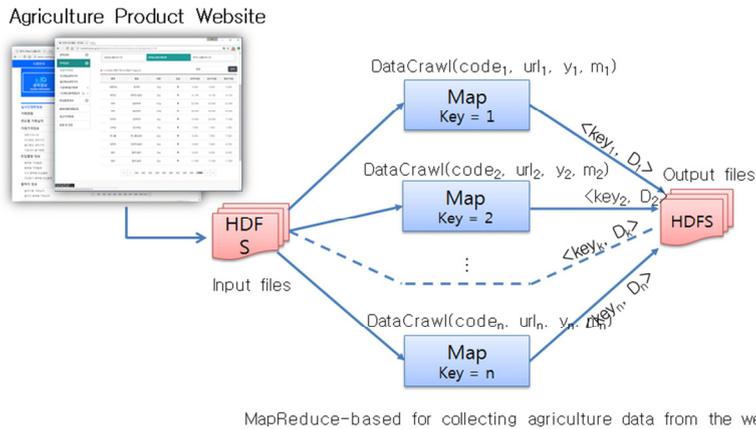


그림 3. 허둠 맵리듀스를 이용한 데이터 수집 모듈

유사하다. 약 2백만 건에 달하는 거대한 데이터를 효율적으로 수집하기 위해 허둠 프레임워크(Hadoop Framework)를 사용한다. 그림 3은 아카이브 데이터를 수집하는 맵리듀스 모델을 나타낸다. 맵리듀스 프레임워크는 대용량 데이터를 분산처리하기 위한 목적으로 개발된 프로그래밍 모델이다. 맵리듀스는 임의의 순서로 정렬된 데이터를 분산처리(Map)하고 이를 다시 합쳐서(Reduce) 처리하는 과정을 거친다. 웹사이트에서 수집된 대용량 입력 파일을 분리(Split)하고

Map 함수에 적용한다. Map 단계와 Reduce 단계 모두 입/출력으로 키-값(Key-Value) 쌍을 가진다. 작업이 완료된 후에는 HDFS에 저장한다[4].

### 3.3 데이터 저장 모듈

설계된 플랫폼에서는 분산파일 시스템인 HDFS를 근간으로 사용해 Hbase, Hive를 데이터웨어하우스 프레임워크로 활용한다. 이때, 데이터 저장 모듈을 데이터 수집 및 분석 작업의 성과와

밀접한 연관이 있다. 즉, 빅데이터 플랫폼의 성능을 위해서는 실시간으로 연속적인 데이터를 수집하기 위한 저장 모듈과 분석 작업을 위해 다양한 데이터를 무작위적으로 접근하기 위한 저장 모듈이 구분되어 활용되어야 한다. 설계된 플랫폼의 Apache Flume은 데이터 수집을 단순화하여 실시간으로 여러 소스에서 데이터를 크롤링해서 HDFS로 데이터를 저장한다. 하둡은 여러 웹에서 거대한 데이터 세트와 같은 아카이브 데이터를 수집하여 HDFS로 저장한다. Jsoup API를 사용하여 집계 데이터 속도를 높이고, Sqoop은 오프라인 데이터를 HDFS, HBase 및 Hive로 가져온다[3]. HDFS는 기본적인 분산 파일 시스템으로 연속적인 데이터 접근에 효과적인 반면, HBase는 무작위 데이터 액세스에 효과적인 저장 모듈이다[5]. 또한 Hive는 빅데이터 플랫폼 상에서 간단한 데이터 질의를 SQL을 통해 수행할 수 있도록 지원한다.

설계된 플랫폼의 실시간 데이터 수집 및 저장 성능 확인을 위해, Flume을 마스터에 배치한 후 하나의 Flume 에이전트가 광주 서부 시장 웹사이트에서 매일 생성되는 데이터를 실시간으로 수집하여 설계된 플랫폼에서 제공하는 분석 서비스에서 활용될 수 있도록 하는 응용 프로그램을 구현했다. 해당 수집 모듈을 통해, 제품 가격 데이터가 실시간 방식으로 수집되어 광주 서부 시장의 웹사이트에서 생성되는 모든 텍스트를 2분마다 새로운 파일로 HDFS로 저장할 수 있음을 확인하였다.

또한 아카이브 데이터 수집 및 저장 성능 확인을 위해, 광주 서부 시장 및 부산 업공 시장의 2016년 1월부터 2016년 6월까지 6개월 동안의 아카이브 데이터를 수집하였다. 하둡에서 두 개

의 맵(Map)작업으로 구성된 하둡 어플리케이션을 실행해 788분 동안 약 144,000 웹 페이지에서 제공되는 농산물 가격 데이터를 수집했다. 즉, 각 페이지의 가격 데이터는 약 330 밀리 초 내에 수집되었다. 반면에 8 개의 CPU 코어가 있는 단일 시스템에서는 비슷한 양의 아카이브 데이터를 수집하는 데 총 8000 분 이상 소요되었다. 이에 따라, 설계된 플랫폼을 통해 실시간 데이터 뿐만 아니라 아카이브 데이터를 효과적으로 수집하고 저장할 수 있음을 확인하였다[3].

### 3.4 데이터 처리 모듈

농업 빅데이터 플랫폼에서는 다양한 형태의 데이터들을 이용해 여러 가지 분석 작업들을 수행할 수 있어야 한다. 설계된 플랫폼에서는 데이터 처리에 필요한 일회성 데이터 조작 작업 및 기계학습 기반의 반복 작업을 포함하는 분석 작업뿐만 아니라, 농업관련 이미지 분석 및 농업 데이터 상관관계 분석을 위한 연관 룰 분석 작업을 고려한다.

#### (1) 일회성 데이터 처리

일회성 데이터 처리란 저장된 데이터를 활용하여 한번 실행하여 원하는 결과를 얻는 처리 작업을 말한다. 예를 들어, 각 농장 또는 생산물의 전체 필드의 평균 면적 계산과 같이 반복 작업이 필요 없는 일회성 작업을 일회성 데이터 처리라 할 수 있다. 이러한 처리 작업에서는 1,100 만 개가 넘는 레코드가 있는 농가 및 관련 정보를 포함한 농업 생산과 관련된 2GB의 대량의 데이터와 같은 대규모 데이터를 이용한 단순 계산 및 데이터 클리닝이 필요하기 때문에, 맵리듀스 및

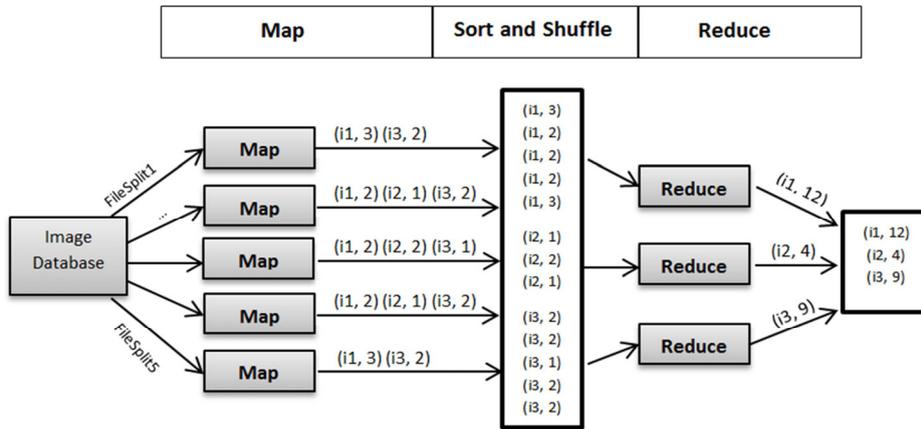


그림 4. 맵리듀스를 이용한 이미지 히스토그램 계산

Spark를 활용한 분석 모듈을 설계해야 한다.

## (2) 기계학습 기반 반복 데이터 처리

플랫폼에서는 다양한 기계학습 기반의 데이터 처리가 필요할 수 있고, 대부분의 기계학습 기반 데이터 처리는 반복적인 작업을 필요로 한다. 예를 들어, 잘 알려진 클러스터링 방법인 K-means 알고리즘의 경우도 무수한 반복 작업이 필요하다. 예를 들어 이미지를 분할하는 작업을 진행한다 할 때, 이미지의 영역을 동질의 시각적 영역 단위로 나누는 작업을 진행한다. 이 작업에서는 3차원의 데이터(RGB format : Red, Green, Blue)를 각각의 색상의 범위를 정규화하여 [0,1] 사이의 값만 가지도록 변환한다. 보통 한 색상을 0~255로 표현한다[6].

K-means 알고리즘은 다음 단계를 수행한다.

- 1단계 : 각 데이터 집합으로 사용할 데이터 집합에서 k 개의 데이터 포인트를 무작위로 선택한다.
- 2단계 : 각 클러스터 중심까지의 거리에 따라

클러스터에 데이터 포인트 지정한다.

- 3단계 : 각 클러스터에 대해 새로 할당된 클러스터 구성원을 사용하여 클러스터 중심을 다시 계산한다.

K-means 알고리즘에서 계산 비용은 주로 거리를 계산하기 위해 발생하는데 각 (임의의 중심점 개수 \* 클러스터 개수) 배의 거리 계산이 필요하다. 따라서 알고리즘의 성능을 향상시키기 위해 반복하는 부분을 주 알고리즘과 분리해서 계산한다. 이 단계에서 거리계산하는 부분을 분리하기 위해 맵리듀스 함수를 구현한다. 그림 4는 맵리듀스를 이용한 이미지 히스토그램 계산 작업의 개요를 나타낸다. 이미지 데이터를 입력 받았을 때, 크기가 너무 크면 맵리듀스를 실행하는데 오류가 생길 수 있기 때문에 먼저 파일을 나눠서 각각 Map함수에 적용한다. Mapping과정과 Reduce과정을 진행한다. 클러스터링 부분은 거리 계산의 최종 결과에 영향을 미치지 않는다. 특히 Spark의 경우, K-means 알고리즘의 매개 변수와 호출 함수를 Spark MLlib 라이브

리리를 기반으로 맵리듀스 및 Spark의 반복 작업을 지원한다.

### (3) 이미지 처리

제안된 플랫폼에서는 텍스트 데이터뿐만 아니라 이미지 데이터도 처리될 수 있도록 설계되었다. 그림 5는 이미지프로세싱을 위한 플랫폼 구조를 보여준다. Sensor 등을 통해서 수집된 이미지 데이터는 Clusters를 통해 저장된다. 이미지를 분류하는 등의 데이터 처리를 위해서는 이미지의 가장자리의 경계를 찾는 Low-level 단계의 처리와 기계 학습 영역에 가까운 이미지 분류와 같은 High-level 단계의 작업을 해야 한다. 그리고 Image Processing Algorithms에서 이미지를 구분하거나 분류하는 과정을 거친 후, 하둡에서 맵리듀스를 이용해 결과 값을

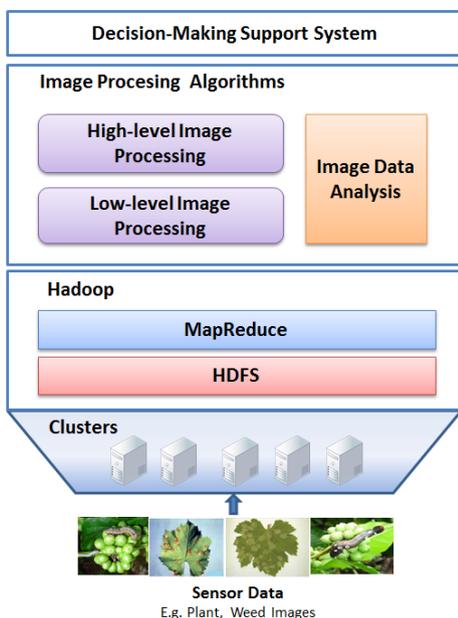


그림 5. 이미지프로세싱을 위한 빅데이터 플랫폼기

HDFS에 저장한다.

그림 6은 K-means를 이용해 구현한 이미지 처리 모듈의 성능 평가를 한 결과이다. 이 실험을 위해 평균 약 1.5MB인 이미지를 수집해서 각 연산 모듈에서 0.5GB에서 2.0GB까지의 다양한 크기의 데이터로 실험했다. 실행한 결과 하둡보다 스파크가 약 5.5 배 빠르다는 것을 확인하였다[7].

### (4) 연관 룰 분석

수집된 데이터를 가지고 연관 룰(Association Rules)을 이용해 재배되는 작물들 간의 관계를 구할 수 있다. 연관 룰 분석은 대표적인 무감독 학습 방식의 분석 방법으로 장바구니 분석이라고 불린다. 소비자들의 구매이력 데이터를 이용해 “A 아이템을 구매하는 고객들은 Y 아이템을 함께 구매할 가능성이 높다”의 결론을 내는 알고리즘이다.

위 설명 중 ‘~를 구매한다면’에 해당하는 부분을 조건절(Antecedent)라고 한다. 그리고 그 뒷부분을 결과절(Consequent)라고 한다. 그리고 조건절과 결과절을 구성하는 아이템들의 집합을 아이템 집합(Item Set)이라 한다. 예를 들

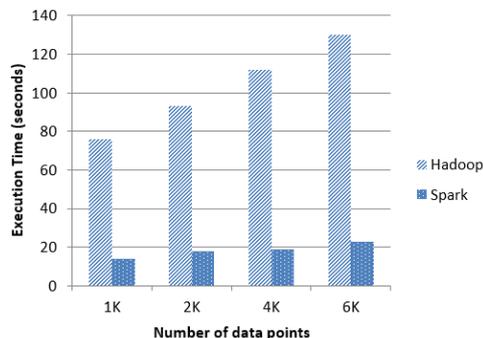


그림 6. 데이터 포인트 개수에 따른 K-mean 알고리즘 수행시간

어 ‘햄버거를 구매하는 사람들은 콜라도 함께 산다’ 라면 ‘햄버거 구매’가 조건절, ‘콜라를 산다’가 결과절이며, 아이템 집합은 ‘햄버거’, ‘콜라’가 된다. 이런 식으로 아이템들 사이에서 많은 규칙들이 생성될 수 있는데 규칙의 효용성을 나타내는 지표는 크게 지지도(Support)와 신뢰도(Confidence), 향상도(Lift)가 있다.

식 1 For the rule  $A \rightarrow B$ ,

$$support(A) = P(A)$$

식 2  $confidence(A \rightarrow B) = \frac{P(A, B)}{P(A)}$

식 3  $lift(A \rightarrow B) = \frac{P(A, B)}{P(A) \cdot P(B)}$

식 1은 빈발 아이템 집합을 판별하는데 쓰이는 지지도(Support)로 ‘조건절이 일어날 확률’로 정의된다. 식 2는 아이템 집합 간의 연관성 강도를 측정하는데 쓰이는 신뢰도이다. 조건절이 주어졌을 때 결과절이 일어날 조건부확률로 정의된다. 식 3은 생성된 규칙이 실제 효용 가치가 있는지를 판별하는데 사용되는 향상도이다. 조건절과 결과절이 서로 독립일 때, 비교해서 두 사건이 동시에 얼마나 발생하는지 비율로 나타낸다. 만약 향상도가 1이라면 조건절과 결과절은 서로 독립이며, 규칙 사이에 유의미한 연관성이 없다는 것을 나타낸다. 규칙의 효용성은 위의 세 가지를 모두 반영해서 평가한다. 세 지표가 모두 클 경우에만 효과적인 규칙이라는 결론을 내릴 수 있다.

제안된 플랫폼을 이용해 콩을 재배하는 농가의 관련 재배 작물을  $Support(\text{지지도}) = 0.001$ ,  $confidence(\text{신뢰도}) = 0.4$ 의 설정을 가지고 분

석 한 결과, 콩을 재배하는 농가의 많은 수가 벼를 함께 재배하고, 그 뒤로 콩, 벼, 엽경채류를 함께 재배한다는 결과를 얻을 수 있었다.

#### 4. 결과

빅데이터에 기반을 둔 농업경영 및 농식품 마케팅 전략 개발은 농업분야 빅데이터를 수집·저장·처리하는 플랫폼 구축 연구와 병행되어야 농업 경영체 및 농식품 산업계가 필요한 정보를 실질적으로 제공할 수 있다. 본 기고에서는 농업분야 빅데이터를 보다 효과적으로 수집·저장·처리하기 위한 플랫폼의 설계에 대해 소개하였다. 다양한 데이터의 특성과 여러 가지 분석 기법을 함께 고려한 플랫폼을 설계함으로써, 보다 효과적이고 효율적으로 농업 경영 전략 개발을 지원할 수 있게 된다. 이를 통해, 기후변화, 병해충 발생 등으로 농업생산의 불확실성이 증대되고 있는 현시점에서, 다양한 농업데이터 분석에 기반을 둔 작황 및 수확량 예측, 적정 출하시기 및 판매전략 등의 합리적인 의사결정을 지원하여 농가소득 향상에 기여할 것이다. 또한, 빅데이터를 활용한 인구구조 및 소비트렌드 변화 등의 농식품 소비 관련 이슈 분석은 농업경영 및 농식품 산업계에 합리적인 의사결정의 기초자료를 제공할 것이며, 농업연구기관 및 정부의 미래성장 연구개발 분야 탐색 및 정책방안 마련에 도움을 줄 것이다.

#### 사 사

본 성과물은(논문, 산업재산권, 품종보호권 등)은 농촌진흥청 연구사업(세부과제번호: PJ011

823022017)의 지원에 의해 이루어진 것임

## 참고문헌

1. 이유재, 이신형, 이종세, “KB국민카드의 마케팅 활동과 빅데이터 활용”, 경영교육연구 Vol.18 No.1, pp. 145-176, 2014.
2. 임종태, 류은경, “빅데이터 기반의 식품 정보 분석 시스템의 설계 및 구현”, 한국콘텐츠학회, pp. 295-296, 2014.
3. 뉘엔 반 퀴엣, 뉘엔 신 녹, 김경백, “농업 빅데이터 수집 및 분석을 위한 플랫폼 설계”, Journal of Digital Contents Society, Vol. 18, No. 1, pp. 149-158, 2017.
4. 뉘엔 반 퀴엣, 뉘엔 신 녹, 부독 티엡, 김경백, “하둡과 플럼 기반 농산품 가격 수집 시스템 설계와 구현”, 한국스마트미디어학회 2016 추계학술대회 논문집, pp. 244-247, 2016.
5. 뉘엔 신 녹, 뉘엔 반 퀴엣, 김경백, “Hbase와 HDFS에서 데이터 검색 시간 비교”, 한국 스마트미디어학회 2016 추계학술대회 논문집, pp. 101-104, 2016.
6. Ray, Siddheswar, and Rose H. Turi. “Determination of number of clusters in k-means clustering and application in colour image segmentation,” Proceedings of the 4th international conference on advances in pattern recognition and digital techniques, pp. 137-143, 1999.
7. 뉘엔 반 퀴엣, 뉘엔 신 녹, 김경백, “농업 이미지 처리를 위한 빅데이터 플랫폼 설계 및 구현”, 한국정보처리학회 2016년 추계학술대회 논문집, pp.50-53, 2016.