

의사결정나무를 이용한 토마토 스마트팜 전문가시스템

남윤만¹ · 이인용² · 백운보^{1*}

¹동의대학교 · ²(주)오토이노텍

Expert System for Tomato Smart Farm Using Decision Tree

Youn-man Nam¹ · In-yong Lee² · Woon-Bo Baek^{1*}

¹DongEui University · ²AutoInnoTech Co. Ltd.

E-mail : kedin13579@naver.com / samjintech1@hanmail.net / wbbak@deu.ac.kr

요 약

의사결정나무를 이용한 토마토 스마트 팜 용 전문가시스템을 설계하여 각 하우스의 주변 환경에 따라 달라지는 요소들에 의해 생성되는 데이터를 사용하여 보다 농민의 결정과 비슷한 의사결정구조를 가진 제어시스템을 구축한다. 현재 스마트 팜의 제어시스템은 지금까지 농민들이 해온 방식과 유사하게 스스로 제어하지 못하였기에 스마트 팜 제어시스템의 의존율은 여전히 높지 못하다. 스마트 팜 내의 센서 값 등 주변 환경을 기준으로 한 환경제어에 농민들의 직접적인 개입이 필수불가결 하다. 그래서 보다 농민의 의사결정과 비슷한 시스템을 만들기 위하여 전문가시스템에 의사결정나무를 접목시킨 제어를 설계하는 것을 목표로 하였다. 하우스 내의 장비를 제어하기에 앞서 여러 환경요소 중에 가장 직접적인 영향을 미치는 것을 자동으로 선정한 후, 농민들의 의사결정 선정기준을 포함시켜 복합적인제어를 위하여 전문가시스템을 구축한다. 이번 연구는 무거운 툴을 사용하지 않고 데이터를 이용하여 결과를 도출하는 것에 초점을 맞추어 진행하였다. 현재 많은 농장에서 스마트 팜을 이용한 데이터들이 쏟아져 나오고 있고 이것을 농민들이 손쉽게 빠르게 접근하여 직접적인 개입을 줄일 수 있는 방법론에 대한 표준이 될 수 있을 것으로 예상 한다.

ABSTRACT

We design an expert system for tomato smart farm using decision trees and construct a control system with decision structure similar to that of farmers by using the data generated by factors that vary depending on the surrounding environment of each house. At present, Smart farm's control system does not control itself like the way farmers have done so far. Therefore, the dependency of smart farm control system is still not high. Direct intervention by farmers is indispensable for environmental control based on surrounding environment such as sensor value in smart farm. Therefore, we aimed to design a controller that incorporates decision trees into the expert system to make a system similar to the decision making of farmers. Prior to controlling the equipment in the house, it automatically selects the most direct effect among the various environmental factors, and then builds an expert system for complex control by including criteria for decision making by farmers. This study focused on deriving results using data without using heavy tools. Data is coming out of many smart farms at present. We expect this to be a standard for a methodology that allows farmers to access quickly and easily and reduce direct intervention.

키워드

decision tree, expert system, id3, smart farm, tomato farm

* corresponding author

I. 서 론

지금까지의 스마트 팜의 제어시스템은 스스로 판단하고 제어하지 못하고 스마트 팜 내.외부의 센서 값 등 주변 환경을 기준으로 환경제어에 농민들이 직접적으로 개입하여 판단 및 결정을 하였다. 그래서 전문가 시스템에 의사결정나무를 접목시킨 제어시스템을 설계하여 보다 농민들의 의사결정과 비슷하며, 나아가서 다년간의 데이터를 축적하여 더욱 정확한 결정을 내릴 수 있다고 기대한다.

하우스 내의 장비를 제어하기에 앞서 ID3 알고리즘을 이용하여 가장 직접적인 영향을 미치는 환경요소를 선정한 후, 농민들의 의사결정 선정기준을 포함시켜 복합적인 제어를 위한 전문가시스템을 구현한다. 이번 연구는 무거운 툴을 사용하지 않고 데이터를 이용하여 결과를 도출하는 것을 목표로 하여 진행하였다.

II. 토마토 스마트 팜

2.1 토마토 팜

현재 4차 산업 혁명이라는 혁신적인 변화 속에서 스마트 팜이라는 분야는 국가차원에서도 많은 관심을 가지고 지원하고 있는 분야이다. 실제로 몇몇 농장에서는 스마트 팜을 이용하고 있으며, 그 중 토마토는 대표적인 하우스 작물이다. 토마토 스마트 팜에서의 빅 데이터를 이용한 많은 연구가 활발하게 이루어지고 있지만 실제 농민들이 농장에 적용하기에는 몇 가지 문제점이 있다.

첫째, 적용이 어렵다. 아무리 많은 연구가 이루어지고 있고 수익성이 보장된다고 하지만 수차례 교육을 받고 전문가가 되지 못한다면 실제 농장에 적용 하기는 불가능하다. 수년간 여러 농장에서는 많은 데이터가 모여 빅 데이터를 이루었지만 이제 막 스마트 팜에 입문한 농민들은 데이터 없이 시작을 해야 한다.

이러한 문제점이 있다고 판단하여 표준 제어 시스템을 제시하고, 각기 다른 농장의 환경에 맞추어 전문가시스템을 구축한다. 그 후 매년 쌓이는 데이터를 의사결정나무를 이용하여 보다 좋은 결정을 내릴 수 있으며, 데이터를 활용하여 스스로 학습할 수 있는 시스템을 제안한다.[1]

2.2 의사결정나무

여러 가지 분류 알고리즘들이 있지만 그중 가장 직관적이며 많이 사용하는 방법인 의사결정 트리를 이용한다. 신경망이나 로지스틱 회귀 분석의 경우에는 의사결정기준이 명확하게 보이지 않기 때문에 전문가 시스템의 지식을 최신화하기에 적절하지 못하다. 토마토의 경우 생장의 종류가 생식생장과 영양생장으로 나누어지며, 각 생장에

따라 알맞은 온도와 습도, 지온, CO2농도, 일사량 등의 환경 요소들이 표 3을 참고하여 정확한 값으로 나누어 저있기 때문에 연속적인 값을 나누기에 적합하지 않은 ID3 알고리즘의 보완할 수 있다.[2] 또한 실제 데이터는 일관적이지 못하며 의사결정나무의 특성상 오차가 있을 수 있다. 하지만 이러한 문제점은 전문가시스템을 이용하여 보완한다.

III. 시뮬레이션

3.1 데이터 생성

농촌 진흥청에서 발부한 ‘ICT 융복합 1세대 스마트 팜 운영 가이드’[3]의 자료 중 표 1 ‘토마토 종합 생장 조절 조건표’를 참고하여 토마토의 생식생장과 영양생장에 따른 조절항목을 ‘농촌 진흥청 농사로 농업인’[4] 홈페이지의 표 3 ‘토마토 생육단계별, 계절별 시설환경 관리 표’와 비교하여 사계절 중 생식생장과 영양생장에 가장 적합한 계절을 선정하였다.

실제 토마토 비닐하우스의 환경제어 센서 값을 받아 올수 없어서 가상데이터를 생성하였다. 먼저 ‘기상 자료 개방 포털’[5]에서 받아온 외부온도 센서, 강우센서, 풍속 풍향 센서 등 9가지 센서 값과 표 2를 바탕으로 하우스 재배에 필요한 데이터인 시설 내 온습도, 잔존CO2, 난방관 온도, 지온 값 등 총 14가지 데이터를 2017년 9월부터 2018년 8월까지 매일 2시간 단위로 총 4,380개를 만들었다.

표 1. 토마토 종합 생장 조절 조건표

종합 생장 조절 조건표				
조절 항목 (환경 및 재배)	현 상태			
	생식생장		영양 생장	
생장 종류	초세 약함	초세 강함	초세 약함	초세 강함
생장 속도	낮음		높음	높음
광합성 속도	높음		높음	높음
광			많이	많이
24시간 평균 온도	낮음	높음	낮음	높음
온도 편차	작게	작게	중간	크게
이른아침 온도상승	느리게	빨리	느리게	빨리
늦은오후 온도하강	빨리	느리게	빨리	느리게
낮 온도	낮게			높게
야간 온도		높게	낮게	
공기 습도	다습	중간	중간	건조
CO2 시비	높음		높음	
축지 제거	안함	제거	일찍	제거
유인	느슨히	중간	중간	지주
착과량	낮음	높음	낮음	높음
조절항목(영양)	낮음			높음
EC	높음			낮음
함수량	높음			낮음
1회 관수량	적게	중간	중간	많게
함수편차	작게	중간	중간	크게
급액지속시간	짧게	중간	중간	길게
급액빈도	자주	중간	중간	드문
급액개시 시각	일찍	중간	중간	늦게
급액미감 시간	늦게	중간	중간	일찍
질소	높음		높음	낮음

표 2. 센서값 활용을 위한 제어 항목 및 내용

설치 장비 제어 항목	활용 센서	제어 내용
천창, 이중창, 축장모터개폐	실내온도/실내습도/실외온도/일사/강우/풍향/풍속센서	환기제어
차광, 보온, 측 커튼 모터 개폐	실내온도/실내 습도/실외온도/일사센서	차광, 보온, 측 커튼제어
CO2 공급밸브	CO2 센서/실내습도/일사센서	CO2공급량 제어
유동팬	실내온도/실내습도 센서	공기유동제어
보광등	일사 센서	보광등 제어
훈증기	타이머	훈증기 제어
스프링클러	실내온도/실내습도 센서	지붕 SP제어
배기팬	실내온도/실내습도 센서	공기배출 제어
난방 순환펌프	실내온도/실내습도 센서	난방 제어
난방 3-WAY 밸브	실내온도/난방수온도 센서	난방 제어

표 3. 토마토 생육단계별, 계절별 시설환경 관리

계절, 생장	일사량 (J)	내부 온도 (°C)	습도 (%)	잔존 CO ₂ (ppm)	흡수량 (ml)
봄	1,500~2,000	18.5 초과	78.0 미만	395~474	3,928 초과
여름	2,000 초과	18.5 초과	78.0~84.2	395~474	3,928 초과
가을	900~1,500	18.5 초과	78.0~84.2	395~474	3,928 초과
겨울	900 미만	18.5 초과	84.2 초과	474 초과	2,579 ~ 3,928
생식	900~1,500	20.3~21.6	81.0~85.6	399 초과	3,350 ~ 4,699
영양	2,000 초과	20.5~23.4	66.8~77.5	348 초과	5,713 초과

각 데이터를 기준으로 표 2 ‘센서값 활용을 위한 제어 항목 및 내용’을 기준으로 천창, 이중창, 축장 모터, 커튼 모터, CO2 공급 밸브 등 14가지 설치 장비 제어 유무를 선정한다.

3.2 결과 및 고찰

생성한 데이터와 표3을 비교하여 각 생장에 따른 센서 값이 적절한지 높은지 낮은지 확인하고, 각 환경 센서 값에 따라 상황에 맞는 설치장비 제어 유무를 정하여 전문가시스템에 의한 가상 데이터를 생성한다.

그 후 매트랩에서 데이터를 이용하여 정밀트리 모델만을 이용하여 각 계절별로 2940개의 트레이닝 데이터를 이용하여 학습 시키고, 나머지 각 계절별 1440개의 테스트 데이터를 이용하여 확인해보니 트레이닝 데이터의 학습 정확도는 96.5%이

며, 테스트 결과는 81.2% 로 확인되었다.

그림 1은 데이터를 기반으로 생성한 전문가시스템의 천창개폐 유무만을 응답변수로 설정하고 매트랩 툴 박스 중 ‘Statistics and Machine Learning Toolbox’의 의사 결정 나무 중 조밀트리를 모델과 비교한 결과이다. 조밀트리의 정확도는 전문가 시스템과 동일한 결과가 나와서 100%의 정확도를 보였으며, 전문가 시스템의 결과와 전문가 시스템을 생성하기 위하여 사용한 변수만을 이용하여 조밀트리 모델로 비교 했을 때의 결과는 97.4%의 정확도를 보였다.

1 ☆ 트리	정확도: 100.0%
마지막 변경: 조밀 트리	특징 45/45
2 ☆ 트리	정확도: 97.4%
마지막 변경: 39개 특징이 제거됨	특징 5/45

그림 1. 의사결정나무 학습 정확도

결과에서 볼 수 있듯이 스마트 팜 내.외부 센서를 통하여 들어오는 일관적이지 않은 데이터들을 시설의 상황과 환경에 맞추어 정확한 기준으로 나누고, 이 값을 이용한 전문가시스템을 사용한다면 의사결정나무의 오차를 상당부분 줄일 수 있다는 것을 확인할 수 있다.

그림 2와 표 3은 동일한 데이터를 사용하였으며, 그림 2는 매트랩의 ‘Compact tree’를 사용하여 어떤 변수를 Root node에 선정하였는지 확인한 결과이고, 표 3은 엑셀에서 ID3 알고리즘을 이용하여 정보 이득을 계산 하였을 때 가장 큰 값이 동일한 변수인지 확인하여 결과에 직접적인 영향을 미치는 변수를 ID3 알고리즘만으로도 도출할 수 있는지를 확인한 결과이다. 매트랩에서는 시설 내 온도 23.45°C를 기준으로 하였고 엑셀을 이용한 ID3 알고리즘에서도 (시설 내)온도의 정보 이득이 약 0.336으로 2번째 값 일사량과 큰 차이를 보이며 가장 높게 나타났다.

표 4. ID3 알고리즘을 이용한 정보 이득 결과

선정항목	정보이득
온도	0.336319493
일사량	0.271155123
지온	0.264760802
습도	0.262715044
CO ₂	0.262480586

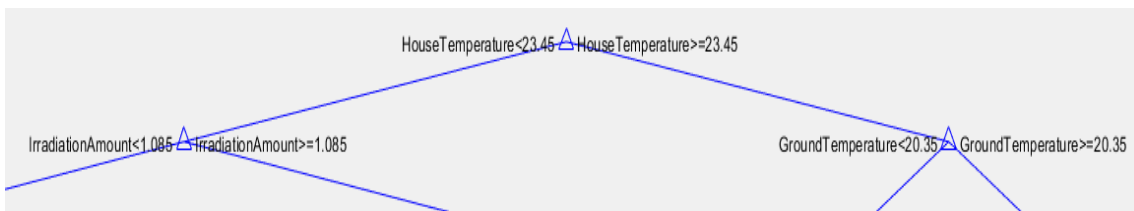


그림 2. 매트랩 Compact tree 결과 그래프

그림 2를 보면 매트랩에서 일사량과, 지온을 그 다음 분류기준으로 잡았으며 표 4에서는 2등과 3등이 일사량과 지온인 것을 확인할 수 있다.

이번 연구 결과를 바탕으로 더 나아가서 ID3 알고리즘을 이용하여 변수들의 가중치를 선정하고 퍼지이론을 이용하여 정확하지 못하고 일관되지 못한 센서 값들을 안정적인 값이 될 수 있도록 하여 오차를 줄이고[6], 결과를 도출하는 것뿐만 아니라 실제 토마토 스마트 팜에서 얻어지는 실측값을 이용하여 실제 제어 값과 시스템에서 도출되어진 제어 값이 어느 정도 일치 하는지 확인할 예정이다.

IV. 결 론

‘빅 데이터 응용 머신 러닝 기반 스마트팜 시스템의 설계 및 구현’[7]이라는 논문에서는 빅 데이터만을 이용하여 시스템을 구성하고 모델을 결정, 팜을 관리하는 시스템을 구축하지만, 이번 연구에서는 농부들의 경험과 판단을 담고 있는 전문가 시스템을 기반으로 하고 있기 때문에 데이터에만 의존하지 않고 여러 기관이나 타 농장에서 더 좋은 사례가 발견 되었을 때 좀 더 빠르고 예민하게 바뀐 정보를 반영 할 수 있으며, 또한 빅 데이터를 사용하기 때문에 각 농장에 최적화된 결정을 반영 할 수 있다.

연구의 목표처럼 무거운 돌이 아닌 액셀을 이용하여 ID3알고리즘을 구현하고, 전문가 시스템을 구축하였으며, 매트랩의 의사결정 나무와 비교하여 의미 있는 결과를 도출하는 것까지 완수를 하였다. 연구를 좀 더 진행하여 액셀이 아닌 좀 더 손쉽게 시스템에 새로운 데이터를 적용 시킬 수 있는 개발환경을 마련 할 것이다.

이 연구가 완료 된다면, 현재 많은 농장에서 쏟아져 나오고 있는 스마트 팜의 데이터들을 농민들이 손쉽게 빠르게 접근하여 직접적인 개입을 줄일 수 있는 방법론에 대한 표준이 될 수 있을 것으로 예상 한다.

References

- [1] G. Oguz, “Decision tree learning for drools,” Master Thesis, Ecole Polytechnique Federale de Lausanne, 2008. [Internet]. Available: http://infoscience.epfl.ch/record/126292/files/oguz-thesis_final.pdf.
- [2] Sonal Patil. et al, (IJCSIT), “Efficient Processing of Decision Tree Using ID3 & improved C4.5 Algorithm” International Journal of Computer Science and Information Technologies, Vol. 6, No. 2, pp. 1956-1961, 2015.
- [3] Rural Development Administration, ICT utilization 1st generation smart farm operation guide, Dec. 2017.
- [4] Rural Development Administration nongsaro. Seasonal plant environment management chart by tomato growth stage [Internet]. Available: <http://www.nongsaro.go.kr/portal/ps/psa/psab/psabm/prdctvImprvntModelSvcList.ps?menuId=PS04313> (2018.09.15.).
- [5] Weather data portal Synoptic weather observation Search [Internet]. Available: <https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>(2018.09.15)
- [6] Renuka D. Suryawanshi, D. M. Thakore, “Decision Tree Classification Implementation with Fuzzy Logic”, IJCSNS International Journal of Computer Science and Network Security, Vol. 12 No. 10, Oct. 2012.
- [7] Oh, Jung Won, “Design and implementation of big data application machine learning-based smart farm system (doctoral degree)”, Graduate School of Catholic University of Daegu, Aug. 2018.