

6ZD-05

C4.5 アルゴリズムを用いた環境コントロールシステムの開発

松本圭一[†] 山下勇輝[†] 山崎裕貴[‡] 堀部典子[‡] 青木振一[†]

崇城大学工学研究科[†] 崇城大学情報学部情報学科[‡]

1. はじめに

近年、日本の農業で、農業就業者の高齢化と後継者不足の問題が指摘されている。農林水産省が公開している農林水産省「農林業センサス」のデータでは、平成12年から平成22年の10年間で農業就業人口が33%減少したことが報告されている。この主要な原因として、高齢化による離農が増えてきている。また、平成22年における農業就業人口は、平均年齢が65.8歳であり、65歳以上の農業就業者の人口が6割、75歳以上の農業就業者の割合が3割となっている。そのため、多くの農家では、必要な作業を行うための十分な人手を確保することが困難な状況となっている。これらの問題を解決するために、効率的・効果的な農作業計画の立案、農作業の自動化、及びそれにかかるコストの削減が重要な課題となっている。そこで、本研究では、データマイニングの手法を用いて、農地で収集される栽培環境データから、効果的な環境制御のための作業内容を導き出すための仕組みの開発を目的とする。本研究本では、農作物の栽培実験と連携し、学習アルゴリズム「C4.5」[1]を用いて環境制御規則を生成する実験と、その有効性を検証する。

表1. サンプルデータ

番号	天気	気温	散水	結果
1	晴れ	高い	×	×
2	晴れ	高い	○	×
3	曇り	高い	×	○
4	雨	普通	×	○
5	雨	低い	×	○
6	雨	低い	○	×
7	曇り	低い	○	○
8	晴れ	普通	×	×
9	晴れ	低い	×	○
10	雨	普通	×	○

2. ID3 アルゴリズム

決定木を生成するため、まず、C4.5の基盤となっているID3アルゴリズムを用いての実験を行う。表1に、アルゴリズムへの入力となるサンプルデータを示す。サンプルデータは、天気

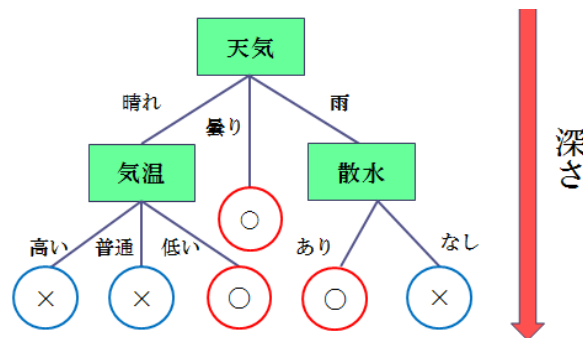


図1. 決定木

と気温と散水の有無から、その結果、農作物や土の状態がどうであったかを「○」と「×」で示すものとなっている。このサンプルデータにID3を適用すると、図1に示す決定木を生成することができる。以下で決定木の生成方法について解説する。

まず、属性「天気」、「気温」、「散水」に対して、識別力を求める。識別力はエントロピーを用いて表される。以下にエントロピーの式を示す。単位はビットである。

$$I(P_1, \dots, P_n) = -\sum_{i=1}^n P_i \log_2 \frac{1}{P_i} = -\sum_{i=1}^n P_i \log_2 P_i \quad (1)$$

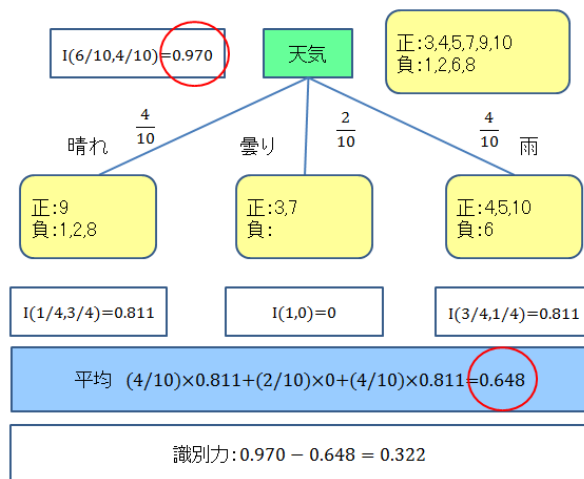


図2. 属性「天気」の識別力

「天気」の属性に着目して、識別力を求めていくと図2の0.322が求められる。同様に、「気温」と「散水」についても識別力を求めていくと図3に示す値が得られる。これら3つの属性の識別力を比べると、「天気」の値が最も大きいので、この決定木の根のラベルは「天気」となる。

Development of Environmental Control System based on Machine Learning Algorithm “C4.5”

[†] Graduate School of Engineering, Sojo University.

[‡] Faculty of Computer and Information Science, Sojo University.



図 3. 属性毎の識別力

本研究では、この決定木で環境制御規則を表現し、水耕栽培実験と連携して、観測された実データへの適用可能性について調査する。

3. 水耕栽培を用いた環境データ収集実験

これまでに行われた太陽光と人工光を比較する実験では、太陽光を利用するよりも完全人工光で育成と安定性があり、大量生産に向いているということが分かっている[2]。そこで、本研究では、データマイニングアルゴリズムの有効性検証を行うための事前実験として、結果が予測できる LED 光源利用の水耕栽培を採用する。本実験では、自然光と LED 光源での植物の成長の差異を見出す実験を行う。水耕栽培キットでレタス系の植物の育成を行い、片方を LED 光源、もう片方を自然光に当てて育成実験を行った。最終的な結果は図 4 のようになった。



図 4. 第一段階実験結果

LED 光源利用のレタスの方が、大きさ・形の点で優れており、安定して成長した。この結果を踏まえて第二段階以降の実験では、LED 光源を用いて、育成実験を行っていく。

次に、第二段階の実験として LED 光源の照射



図 5. 第二段階実験 5 日目

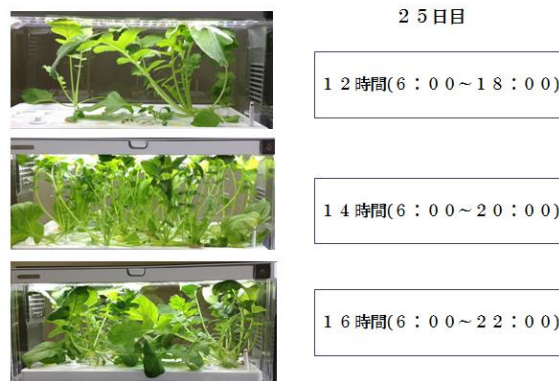


図 6. 第二段階実験 25 日目

時間の違いによって、植物の成長度の違いを見出す実験を行う。ここでは、三台の水耕栽培キットを用いて、それぞれの LED 光源の照射時間を 12 時間、14 時間、16 時間として植物の成長を比較する。図 5 のように、LED 光源の照射時間を 14 時間に設定した場合が発芽している種が多く。また、図 6 においても発芽から食べられる大きさまで成長した種が多かったのはどちらも 14 時間に設定したものだ。

4. おわりに

環境コントロールの自動化のために、仮説モデルとして決定木を用いた。仮説モデルを作る際に、ID3 アルゴリズムの拡張版にあたる C4.5 アルゴリズムを用いて、決定木の生成を行った。

実際に水耕栽培を行って、決定木を生成する際に必要な環境データの収集を行った。その第一段階目の実験として、LED 光源が植物の成長に対して有効であるかの確認実験を行い、LED 光源の有効性を確認した。次に、第二段階目の実験として、植物の成長が安定して行われる LED 光源の照射時間を見出す実験を行い、LED 光源を 14 時間照射と植物が安定して成長することが分かった。今後の課題として、第一段階と第二段階で分かった結果を踏まえて、決定木生成の際に必要なさまざまなパターンを入力データを得るために、さまざまな条件下での水耕栽培実験を行う。さらに、得られたデータから決定木を生成してその決定木の有効性を確認する育成実験を行う。

参考文献

- [1] J. R. Quinlan: Induction of Decision Trees, Machine Learning 1 (1986)
- [2] 植草新一郎 勝俣裕 池田敬 三浦登: 植物工場用光源の研究, 西松建設技報 Vol. 36 (2013).