



딥러닝 기반의 농산물 가격 예측 시스템에 대한 연구

김진*, 이정일**

A Study on Agricultural Price Prediction System based on Deep Learning

Jin Kim*, Jeong-Il Lee**

이 논문은 중원대학교 교내 학술연구비 지원에 의한 것임(과제관리번호 : 2017-078)

요 약

본 논문은 딥러닝 기반의 농산물 가격 예측 시스템에서 농산물 가격정보를 이용한 강화학습 모델을 제안한다. 한국농수산물유통공사 농산물유통정보(KAMIS)에서 제공하는 정보를 기반으로 기후정보를 적용하기에 부적합한 시설농작물을 대상으로 가격변동 정보를 수집한 후 제안 시스템에 적용하였다. 제안 시스템은 환경, 에이전트, 정책신경망과 정책학습기로 구성되었으며, 정책신경망은 LSTM 신경망을 적용하여 판매와 수급 행위에 대해 수익을 높일 수 있을지 확률을 계산한다. 제안 시스템을 이용하여 대상 농산물에 대하여 시뮬레이션한 결과 농산물 가격 변화와 비슷한 예측 결과를 보였으며 기존의 복잡한 농산물 가격 예측 기법과 비교하여 상당히 개선된 결과를 보였다.

Abstract

This paper proposes a reinforcement learning model using agricultural price information as a deep learning-based agricultural price prediction system. Based on the information provided by KAMIS of the Korea Agro-Fisheries & Food Trade Corporation, information on price changes was collected and applied to the proposed system for the facilities crops that are not suitable for applying climate information. The proposed system consists of the environment, agents, policy neural networks and policy learners, which apply the LSTM neural network to calculate the probability of increasing revenue for sales and supply and demand activities. The simulation of the target agricultural products using the proposed system showed similar prediction results to the change of agricultural prices and showed a significant improvement compared to the existing complex agricultural price prediction techniques.

Keywords

agricultural price prediction system, deep learning, reinforcement learning, KAMIS

* 중원대학교 컴퓨터공학과

- ORCID: <https://orcid.org/0000-0001-8856-879X>

** 한국폴리텍대학 스마트소프트웨어과(교신저자)

- ORCID: <https://orcid.org/0000-0002-4959-6592>

• Received: May 28, 2019, Revised: Jun. 17, 2019, Accepted: Jun. 20, 2019

• Corresponding Author: Jeong-Il Lee

Dept. of smart software, Korea Polytechnics, Korea

Email: ayin4me@hanmail.net

I. 서론

인공지능(AI, artificial intelligence)이란 용어는 2016년 알파벳의 구글 딥마인드에서 개발한 바둑 인공지능 프로그램인 알파고와 이세돌 구단과의 대결에서 알파고가 승리하면서부터 우리 일상에서 흔한 용어가 되었으며 현재는 미래 4차 산업의 핵심 기술로 인식되고 있다. 또한 인공지능 분야의 연구는 다양한 분야에서 응용되고 있으며, 특히 다항식 시간(polynomial time)의 알고리즘이 아직 발견되지 않은 NP-완전(nondeterministic polynomial complete) 문제의 해를 탐색함에 있어 기존의 군집화(clustering) 알고리즘, 백트래킹(backtracking) 알고리즘과 유전자 알고리즘(genetic algorithm) 등의 알고리즘 보다 좋은 성능의 근사해 또는 최적해를 제공하며 학습이 완료된 이후에는 이전 알고리즘의 시간복잡도 보다 빠른 결과 값을 출력한다.

오늘날 한국 농산물 유통의 여건이 대내외적인 환경 변화로 인해 가격 결정 요인이 다변화됨에 따라 과거 농산물 유통 관련 주체들은 제한된 비분석적 정보와 단편적 정보에 의존하여 수급을 결정하는 방식으로 상당한 위험을 내포하고 있다. 이러한 문제를 해결하기 위해 농업부문 시뮬레이션 모형 개발과 각종 통계정보 공급, 정책 수립 및 관련 분야 연구 등과 같이 농산물에 대한 유통 및 가격 예측을 위한 다양한 시도가 이루어졌다. 변동성이 높은 작물인 배추와 양파, 마늘, 무를 대상으로 가락 시장에서 월별 가격 예측에 대한 ARIMA 모형 연구와[1], 청과물 14품목의 가격 계절성을 고려하여 가락시장의 평균가격을 GARCH(generalized autoregressive conditional heteroscedasticity)모형과 VAR(vector autoregression) 모형을 이용하여 예측한 연구가 있다[2]. 인공지능망을 이용하여 중국 정저우 곡물 도매시장에서의 월별 밀 가격을 예측한 연구와[3], 인도 코임바토르 시장의 토마토 가격을 역전과 뉴럴 네트워크(backpropagation neural network) 모형을 이용하여 예측한 연구가 있다[4][5].

이에 본 연구에서는 딥러닝 기반의 농산물 가격 예측 시스템으로 농산물 가격정보를 이용한 강화학습 모델을 제안한다.

II. 관련 연구

2.1 강화학습

강화학습은 학습주체가 주변환경과 상호작용하며 주어진 목적을 이루기 위해 최선의 행동을 선택하는 기계학습(machine learning)의 한 종류이다. 지도 학습(supervised learning)은 입력과 입력에 대하여 구해진 결과를 학습하지만, 강화학습은 학습주체가 주어진 환경에 대하여 임의의 행동(action)을 취하고, 이 행동에 대한 보상(reward)을 받음으로써, 주어진 상태(state)에서 어떻게 행동해야 보상을 극대화하는지 학습하는 방법이다[6]. 여기서 학습주체를 에이전트(agent)라 하고, 에이전트와 상호작용하는 외부 전체를 환경(environment)이라고 한다. 또한 에이전트의 행동에 대한 가치평가를 스칼라 값으로 표현한 것이 보상(reward)이다. 환경과 에이전트 간의 상호작용은 그림 1과 같이 매 시간 t 마다 에이전트는 현재 상태 s_t 에 대한 보상 r_t 를 받고, 학습된 규칙을 통해 행동 a_t 를 수행한다. 환경은 에이전트의 행동에 대한 변화된 결과로 에이전트의 상태를 s_{t+1} 가 되게 하고, 새로운 보상값 r_{t+1} 을 받게 한다. 강화학습에서 에이전트의 행동은 정책함수(policy function) $\pi(s_t)$ 에 의해서 결정되는데, 강화학습의 목표는 최적의 정책함수 π^* 를 찾음으로써 미래에 받을 보상의 합을 최대화하는 것이다.

$$G = r_{t+1} + r_{t+1} + \dots = \sum_{k=1}^{\infty} r_{t+k} \quad (1)$$

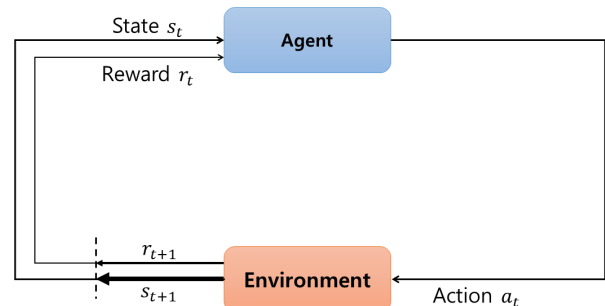


그림 1. 에이전트와 환경 간의 반복적 상호작용
Fig. 1. Iterative interaction between agent and environment

2.2 마르코프 의사결정 과정

불확실한 상황에서 어떠한 의사결정을 하기 위해서는 "확률"에 기초하여 분석을 한다. 어떤 사건에 대한 확률이 시간에 따라 변하는 과정을 확률적 과정(stochastic process)이라고 하며, 일반적으로 강화 학습이론은 마르코프 의사결정 과정(Markov decision process, MDP)을 기반으로 한다. 마르코프 의사결정 과정은 모든 환경이 어떤 상태가 다음 결정을 내리기에 필요한 모든 정보를 가지고 있음을 의미하는 마르코프 속성(Markov property)을 갖는다는 가정을 한다. 즉, MDP는 식 2와 같이 상태집합 S , 행동집합 A , 상태 전이 확률 행렬 P , 보상 함수 R , 할인 요인 γ 로 구성되어 있다.

$$MDP = (S, A, P, R, \gamma) \quad (2)$$

상태 집합은 MDP에서 가질 수 있는 모든 상태의 집합 $\{s_1, s_2, \dots, s_{|S|}\}$ 이며, 식 3과 같이 어떠한 시점에서의 상태 x_t 는 상태 집합 S 에 포함된 특정 상태가 된다.

$$x_t = s, s \in S \quad (3)$$

행동 집합은 행동 주체인 에이전트가 할 수 있는 모든 행동들의 집합 $\{a_1, a_2, \dots, a_{|A|}\}$ 이며, 에이전트는 어떠한 시점에서 행동 a 를 취하는 것이다. 또한 MDP에서의 상태 전이 확률은 마르코프 과정의 상태 전이 확률보다 조금 더 복잡하며, MDP의 상태 전이 확률은 식 (4)와 같다.

$$P^{a_{s_i, s_j}} = P(x_{i+1} = s_j | x_t = s_i, A_t = a) \quad (4)$$

P_{s_i, s_j}^a 는 에이전트가 어떠한 상태 s_i 에서 행동 a 를 취했을 때 상태 s_j 로 변할 확률이며, 보상 함수는 에이전트가 어떠한 상태에서 취한 행동에 대한 보상을 내리기 위한 함수이며 식 (5)와 같다.

$$R^a: s, a \rightarrow \mathbb{R} \quad (5)$$

보상 함수 R_s^a 는 상태 s 에서 행동 a 를 했을 때의 보상을 수치로 반환한다. 또한 할인 요인은 과거의 행동들을 얼마나 반영할지를 정하는 값으로 0에서 1 사이의 값으로 결정한다.

2.3 Q 러닝 강화학습

Q 러닝 강화학습은 Watkins and Dayan[7]에 의해 제안된 동적 프로그래밍 기법(Dynamic Programming)을 기반으로 한 알고리즘으로, 행동가치함수 Q^π 를 이용하여 최적정책 π^* 를 찾는 기법이다. 행동가치함수 $Q^\pi(s, a)$ 는 가치함수 V^π 와는 다르게 최적의 행동가치함수 $Q^*(s, a)$ 가 만들어지면 환경 정보 없이 최적의 정책을 추출할 수 있다는 장점이 있다. Q 러닝 강화학습의 또 다른 특징은 off 정책 학습 기법으로 최적 정책 Q^* 을 도출하기 위해 특정 정책을 항상 따르지 않아도 주어진 환경을 균일하게 방문하기만 하면 최적정책으로 수렴한다는 것이다. Q 러닝 강화학습의 행동가치함수 $Q(s_t, a_t)$ 의 갱신은 식 (6)과 같다.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \cdot \max_{a_{\max}} Q(s_{t+1}, a_{\max}) - Q(s_t, a_t)] \quad (6)$$

Q 러닝 강화학습이 최적정책으로 수렴하기 위해서는 상태-행동 공간 내에 있는 모든 상태-행동 쌍을 일정하게 탐험(exploration)하는 것이 필요하다. 미지의 영역에 대한 행동을 선택하는 탐험과 현재 정책에서 최선의 행동을 선택하는 탐사(exploitation)가 있으며, 이 둘을 적절하게 사용하면 Q 러닝 강화학습의 정책을 최적정책으로 수렴시킬 수 있다 [8].

III. 딥러닝 기반의 농산물 가격예측 시스템

딥러닝 기반의 농산물 가격예측 모델은 기존 농산물 가격 결정에 주요 변수인 기후정보와 생산정보, 유통정보, 소비자 의향지수 등을 고려한 빅데이터 분석과 시계열 데이터 학습 기반의 인공지능 기술이 주로 사용되었으며, 이러한 예측 모델의 경우

다양한 변수를 사용함에 따라 필연적으로 많은 데이터를 정제하고 분석해야 하며 학습 데이터의 상관관계를 어떻게 정의하는가에 따라 예측 결과에 대한 신뢰도에도 많은 영향을 미친다[9]. 이에 본 논문에서 제안하는 딥러닝 기반의 농산물 가격 예측 시스템에서는 다양한 변수를 고려한 학습 모델의 문제인 데이터 분석과 예측 신뢰도 저하를 효과적으로 개선하기 위해 농산물 가격정보를 이용한 강화학습 모델을 제안하며 제안 모델의 핵심 모듈 구성은 그림 2와 같다.

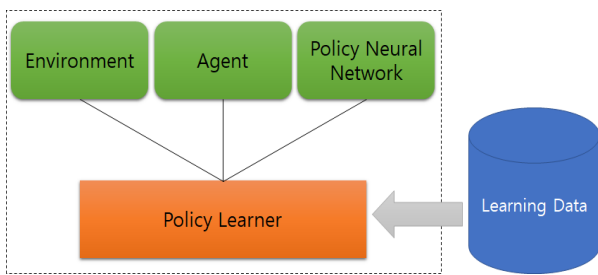


그림 2. 모듈 구성도
Fig. 2. Module configuration diagram

제안 시스템의 핵심 모듈은 환경, 에이전트, 정책 신경망과 정책학습기 및 학습 데이터로 구분되며 각 모듈의 기능 명세는 다음과 같다.

- 환경 모듈은 예측에 필요한 농산물 가격변동에 대한 데이터를 관리하며 에이전트에게 과거 시점에서 최근 시점까지 가격변동 데이터를 순차적으로 제공한다.
- 에이전트 모듈은 예측 대상 농산물에 대한 판매와 수급을 결정하는 모듈로서 농산물 재고량과 초기 설정 금액, 현금잔고라는 상태가 있고 농산물 가치평가 기준을 기준 가치(BV, Base Value)라 정한다.

$$BV = \text{재고량(판매단위)} \times \text{현재가격} + \text{현금잔고} \quad (7)$$

BV가 초기 설정 금액보다 높으면 이익이 발생하는 것이고 낮으면 손실이 발생한 것으로 에이전트의 전략은 BV를 높여 나가는 것이다.

- 정책신경망 모듈은 특정 시점의 농산물 가격 데이터를 기반으로 판매할지 수급할지를 결정하는 에이전트의 판단 근거를 제공하는 역할을 수행하며 이러한 정책신경망은 LSTM(Long Short-term Memory) 신경망 모델로 구성되며 판매와 수급 행위에 대해 BV를 높일 수 있을지의 확률을 계산한다.

▪ (도매가격 - 품목별) 채소류 / 상추 / 적, 등급:상품, 단위:4kg

(단위: 원)

↓ 데이터 저장

구분	당일 03/11	3일전 03/08	4일전 03/07	5일전 03/06	6일전 03/05	7일전 03/04	1개월전	1년전	평년
평균	12,000	11,600	12,200	12,800	12,000	11,400	11,560	11,150	11,673
최고값	13,000	13,000	14,000	14,000	13,000	13,000	12,400	12,750	12,633
최저값	11,000	10,000	11,000	12,000	11,000	10,000	10,800	9,000	11,000
등락률	-	3.4	-1.6	-6.3	0.0	5.3	3.8	7.6	2.8
서울	적	12,000	12,000	12,000	12,000	12,000	12,200	12,750	11,733
부산	적	11,000	11,000	11,000	12,000	11,000	10,000	11,500	11,500
대구	적	13,000	13,000	14,000	14,000	13,000	13,000	11,000	11,500
광주	적	13,000	12,000	13,000	14,000	13,000	12,400	11,500	12,633
대전	적	11,000	10,000	11,000	12,000	11,000	10,800	9,000	11,000

그림 3. 농산물 가격변동 기초 데이터(출처 : KAMIS)
Fig. 3. Basic data on agricultural price fluctuation (Source: KAMIS)

- 정책학습기 모듈은 에이전트, 환경, 정책신경망을 포함하며 강화학습 수행의 핵심 모듈이다. 정책학습기는 학습 데이터를 가지고 있고 보상이 결정되었을 때 학습 데이터로 정책신경망을 학습한다.
- 학습 데이터는 농산물의 가격변동 데이터를 담은 2차원의 차트 데이터로서 그림 3과 같이 국내 농산물유통정보를 제공하는 KAMIS 시스템에서 수집한 정보이다.

KAMIS 시스템에서 제공하는 정보를 기반으로 예측 대상 농산물에 대한 평년(기준연도 1년 이전 연도에서 과거 5년간 해당 일에 대한 최고값과 최소값을 제외한 평균값) 가격변동 정보를 수집한 후 제안 시스템에 적합하게 전처리하여 농산물 가격변동 차트 데이터를 획득하였다. 이외에도 제안 시스템의 결과 분석을 위해 Matplotlib 라이브러리를 이용한 가시화 도구를 설계에 반영하였다.

강화학습은 경험을 학습하는 것으로서 경험을 충분히 쌓기 위해 많은 반복을 시행하여 학습 대상 데이터 전부를 대상으로 한 차례 반복한 과정을 에포크(epoch)라 한다. 한 에포크에서 경험을 얻기 위해서 무작위로 행동하는 것을 탐험(exploration)이라 하는데, 일반적으로 강화학습 초반에는 탐험을 많이 하고 후반으로 갈수록 탐험을 적게 한다. 또한 탐험 비율을 엡실론(epsilon)이라 하며, 1만 번을 반복한다고 가정했을 때 1 에포크에서는 엡실론을 30%로 정하고 점점 엡실론을 줄여 1만 번째 에포크에서 0%이 되게 하면 에포크가 증가할수록 무작위로 하

는 행동이 줄어들게 된다. 즉 경험이 쌓일수록 무작위로 하는 행동을 줄이는 것이다.

무작위로 행동을 하지 않을 때는 정책신경망으로 행동을 결정하며 정책신경망의 출력값이 높은 행동을 선택한다. 이렇게 무작위 결정이든 정책신경망의 결정이든 에이전트는 결정을 기반으로 행동을 수행한다. 수급의 경우 에이전트는 농산물을 매입할 수 있는 금액을 확인하고 수급을 진행하거나 그렇지 않으면 대기하며 수급을 실행한 후 매수금액만큼 차감하고 재고 수량을 늘려준다. 판매의 경우 재고량을 줄이고 현금보유액을 더하는 방식으로 처리한다. 이와 같은 반복적 방식으로 학습을 진행하면 정책신경망의 가중치들이 업데이트되어 이후에 진행되는 학습부터 업데이트된 정책신경망의 결과가 반영된다.

IV. 실험 및 결과 분석

본 논문에서 제안한 딥러닝 기반의 농산물 가격 예측 시스템은 기존 연구와의 차별화를 위해 기후 정보를 적용하기에 부적합한 시설농작물을 대상으로 하며 실험에는 시설농작물 중 가격 변동성이 높은 상추를 대상으로 지정하였다. 실험 데이터는 KAMIS 시스템에서 제공하는 정보를 기준으로 학습 데이터로 평년에 해당하는 가격변동 정보를 기반으로 학습하였고 학습을 완료한 정책신경망 모델에 2018년 1월에서 12월까지 1년간의 데이터를 적용하여 분석하였다.

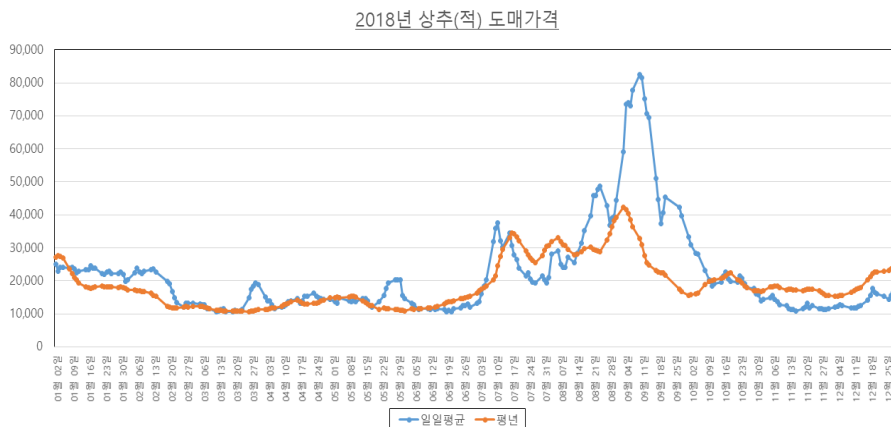


그림 4. 실험 데이터 가격변동 추이 그래프

Fig. 4. Graph of price fluctuation of experimental data

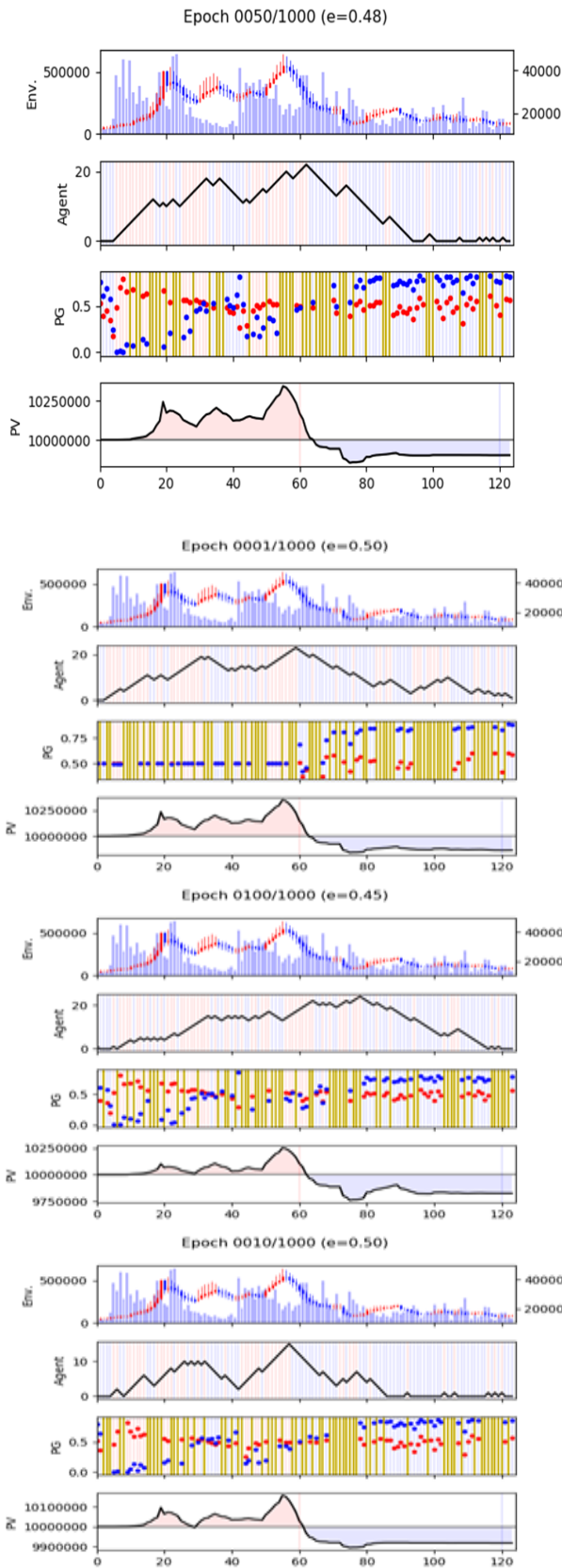


그림 5. 대상 농산물 강화학습 과정(Epoch 1, 10, 50, 100)
 Fig. 5. Target agricultural product reinforcement learning process(Epoch 1, 10, 50, 100)

그림 4와 같이 시설농작물 중 가격 변동성이 높은 상추의 경우 평년 평균가격에 비교하여 해당연도 일일 평균 가격변동 흐름이 전반적으로 비슷한 양상으로 변화되나 8월 말에서 10월 초까지 상당한 가격 차이가 나타나는 것을 확인할 수 있다. 또한, 본 실험에서는 시설농작물인 상추에 대하여 2017년 이전 5년간의 평균 가격변동 값을 기반으로 학습을 진행하고 이후 2018년 해당 농산물의 가격 변동 정보로서 검증하는 백테스트(back test)를 진행하였다. 일반적으로 강화학습에서 사용되는 백테스트에는 몇 가지 환경 설정 변수를 설정하여야 하며 본 연구에서는 농산물 가격 예측을 농산물의 현재 가격을 기반으로 판매와 수급을 통해 최대 이익을 추구하는 것을 목표로 학습을 진행하였다. 이에 필요한 학습 파라미터로는 초기 자본금과 최소·최대 거래 단위, 총 에포크, 초기 탐험률, 지연 보상 임계치, 학습 속도를 적절히 설정하여 학습을 진행한다. 따라서 본 실험에서는 학습 속도를 0.0001로 설정한 확률적 경사 하강법(stochastic gradient descent : SGD)으로 학습하며 초기 투자금 3,000만 원과 에포크는 1,000번, 탐험률 0.5로 설정하였다.

그림 5는 평년 기준 상추 가격변동 정보를 기반으로 학습한 결과로 1,000번의 학습 과정 중 초기 학습 단계에 해당하는 에포크 1과 10, 50, 100에 해당하는 것이다. 학습 초반 50%의 탐험확률에 의해 무작위로 판매와 수급을 진행한 것으로 판매와 수급에 대한 다양한 결정들을 볼 수 있다.

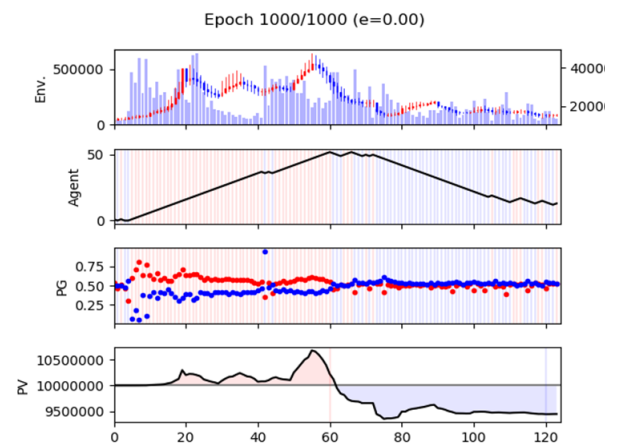


그림 6. 대상 농산물 강화학습 과정(Epoch 1,000)
 Fig. 6. Target agricultural product reinforcement learning process(Epoch 1,000)

그림 6은 학습의 최종 결과인 에포크 1000번째의 그래프로써 학습을 시작한 초기의 무작위 패턴에서 학습이 진행됨에 따라 전체 가격변동에 대해 적절한 행동을 수행하는 것을 알 수 있으며 판매와 수급에 대한 일관된 행동을 선택하고 있는 것을 알 수 있다.

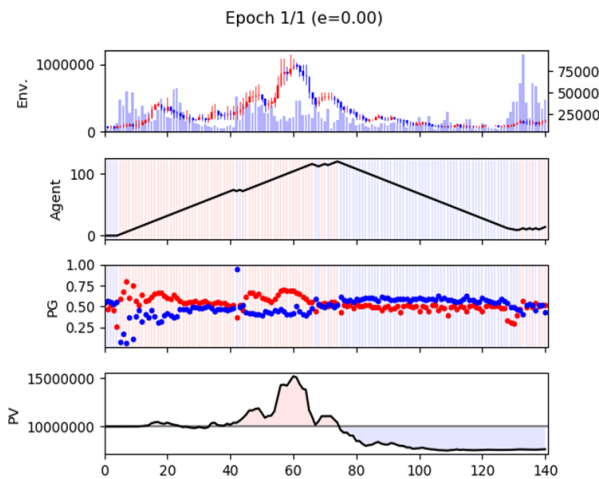


그림 7. 제안 시스템 검증
Fig. 7. Proposed system verification

그림 7은 평년 상추 가격을 기준으로 학습한 제안 시스템의 학습 결과에 대하여 2018년 상추 가격 변동 정보를 기반으로 검증을 실시한 내용이다. 2018년 상추 가격 변동은 평년과 비교하여 비슷한 흐름을 보이고 있으며 정책신경망의 결정이 보다 정밀한 결정을 동반하고 있으며 수익률도 비교적 추세에 따라 잘 형성되는 것을 알 수 있다.

V. 결론 및 향후 과제

농산물 가격 예측에 대한 여러 가지 연구와 실험, 통계적 자료들이 존재하고 있으나 가격 예측의 정밀도를 높이기 위해 농산물 가격 정보 외에 기후 정보와 농산물 생산량, 선호도 등의 많은 변수를 적용하는 사례가 대부분으로 이러한 변수들을 고려한 설계는 복잡하며 그에 따른 상관성 분석 또한 복잡하다. 이에 본 연구에서는 농산물 가격 정보를 이용한 강화학습 모델을 제안하였다. 제안 시스템을 이용하여 대상 농산물에 대하여 시뮬레이션한 결과 농산물 가격 변화와 비슷한 예측 결과를 보였다.

그러나 상추와 같은 시설 농산물 외에 생산량과 선호도 및 기후적 특성에 영향을 받는 농산물들에 대한 비교 실험이 향후 필요할 것이다.

References

- [1] B. S. Kim, "A Comparison on Forecasting Performance of the Application Models for Forecasting of Vegetable Prices", *The Korean Journal of Agricultural Economics*, Vol. 46, No. 4, pp. 89-113, Apr. 2005.
- [2] Y. S. Lee, B. O. Choi, and S. B. Shim, "A Time Series Analysis on Prices of Fruits and Vegetables", *Agri-Policy Focus R537*, Korea Rural Economic Institute, pp. 53-58, Nov. 2006
- [3] Haofei Z., Guoping X., Fangting Y., and Han Y., "A neural network model based on the multi-stage optimization approach for short-term food price forecasting in China", *Expert Systems with Applications*, Vol. 33, No. 2, pp. 347-356, Aug. 2007.
- [4] N. Hemageetha and G. M. Nasira, "Radial basis function model for vegetable price prediction", *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering*, IEEE, pp. 424-428, Apr. 2013.
- [5] Sungho Shin, Mikyoung Lee, and Sa-kwang Song, "A Prediction Model for Agricultural Products Price with LSTM Network", *The Korea Contents Society*, Vol. 18, No. 11, pp. 416-429, Nov. 2018.
- [6] R. S. Sutton and A. G. Barto, "Reinforcement Learning An Introduction", *Bradford Book*, pp. 2-3, Aug. 2018
- [7] G. Bolch, S. Greine, H. De Meer, and K. S. Trivedi, "Queueing networks and Markov chains: modeling and performance evaluation with computer science applications", *John Wiley & Sons.*, pp. 321-322, May 2006
- [8] Christopher J. C. H. Watkins and Peter Dayan, "Q-learning", *Machine learning*, Springer Link, Vol. 8, pp. 279-292, May 1992.

[9] Dong-Ha Shin, Kwang-Ho Choi, and Chang-Bok Kim, "Deep Learning Model for Prediction Rate Improvement of Stock Price Using RNN and LSTM", Korean Institute of Information Technology, Vol. 15, No. 10, pp. 9-16, Oct. 2017.

저자소개

김진 (Jin Kim)



2013년 2월 : 충북대학교
컴퓨터공학과(공학박사)
2013년 ~ 현재 : 중원대학교
컴퓨터공학과 조교수
관심분야 : 컴퓨터네트워크,
정보보호

이정일 (Jeong Il Lee)



2010년 8월 : 충북대학교
컴퓨터공학과(공학박사)
2013년 ~ 2018년 : 중원대학교
컴퓨터시스템공학과 겸임교수
2014년 ~ 2018년 : 주식회사
아인시스텍 대표이사
2019년 ~ 현재 : 한국폴리텍대학

안성캠퍼스 스마트소프트웨어과 초빙교수
관심분야 : 인공지능, 병렬처리, 영상처리