

<https://doi.org/10.7236/JIIBC.2023.23.4.9>  
JIIBC 2023-4-2

## 심층 강화학습 기반의 대학 전공과목 추천 시스템

### Recommendation System of University Major Subject based on Deep Reinforcement Learning

임덕선, 민연아, 임동균\*

Ducsun Lim, Youn-A Min, Dongkyun Lim\*

**요약** 기존의 단순 통계 기반 추천 시스템은 학생들의 수강 이력 데이터만을 활용하기 때문에 선호하는 수업을 찾는 것에 많은 어려움을 겪고 있다. 이를 해결하기 위해, 본 연구에서는 심층 강화학습 기반의 개인화된 전공과목 추천 시스템을 제안한다. 이 시스템은 학생의 학과, 학년, 수강 이력 등의 정형 데이터를 기반으로 학생들 간의 유사도를 측정하며, 이를 통해 각 전공과목에 대한 정보와 학생들의 강의 평가를 종합적으로 고려하여 가장 적합한 전공과목을 추천한다. 본 논문에서는 이 DRL 기반의 추천 시스템을 통해 대학생들이 전공과목을 선택하는 데에 유용한 정보를 제공하며, 이를 통계 기반 추천 시스템과 비교하였을 때 더 우수한 성능을 보여주는 것을 확인하였다. 시뮬레이션 결과, 심층 강화학습 기반의 추천 시스템은 통계 기반 추천 시스템에 비해 수강 과목 예측률에서 약 20%의 성능 향상을 보였다. 이러한 결과를 바탕으로, 학생들의 강의 평가를 반영하여 개인화된 과목 추천을 제공하는 새로운 시스템을 제안한다. 이 시스템은 학생들이 자신의 선호와 목표에 맞는 전공과목을 찾는 데에 큰 도움이 될 것으로 기대한다.

**Abstract** Existing simple statistics-based recommendation systems rely solely on students' course enrollment history data, making it difficult to identify classes that match students' preferences. To address this issue, this study proposes a personalized major subject recommendation system based on deep reinforcement learning (DRL). This system gauges the similarity between students based on structured data, such as the student's department, grade level, and course history. Based on this information, it recommends the most suitable major subjects by comprehensively considering information about each available major subject and evaluations of the student's courses. We confirmed that this DRL-based recommendation system provides useful insights for university students while selecting their major subjects, and our simulation results indicate that it outperforms conventional statistics-based recommendation systems by approximately 20%. In light of these results, we propose a new system that offers personalized subject recommendations by incorporating students' course evaluations. This system is expected to assist students significantly in finding major subjects that align with their preferences and academic goals.

**Key Words** : Deep Reinforcement Learning, Deep Q-Network, Dueling Deep Q-Network, Markov Decision Process, Recommendation System

\*정회원, 한양사이버대학교 컴퓨터공학과  
접수일자 2023년 7월 4일, 수정완료 2023년 7월 31일  
게재확정일자 2023년 8월 4일

Received: 4 July, 2023 / Revised: 31 July, 2023 /

Accepted: 4 August, 2023

\*Corresponding Author: eiger07@hycu.ac.kr

Dept. of Computer Science Engineering, Hanyang Cyber University, Korea

## I. 서 론

인공지능은 과거의 하드웨어 한계를 극복하며 지속해 발전해 왔다. 강화학습은 인공지능의 주요 분야 중 하나로, 이는 기계학습의 한 방법이며, 에이전트가 환경과의 상호작용을 통해 최적의 행동을 선택하고, 이에 따른 보상을 통해 학습을 진행하는 방식이다<sup>[1]</sup>. 강화학습의 목표는 환경과의 상호작용을 통해 받는 보상을 최대화하는 것이다.

이러한 강화학습은 딥러닝과 결합하여 심층 강화학습이라는 새로운 분야로 발전하였으며, 다양한 분야에서 응용되고 있다<sup>[2-5]</sup>. 예를 들어, 사용자의 취향과 관심을 파악하여 개인에게 맞는 제품이나 콘텐츠를 추천하는 시스템에서 더욱 효과적인 추천을 위해 활용될 수 있다. 심층 강화학습 기반의 추천 시스템은 사용자와 시스템 사이의 상호작용을 통해 최적의 추천을 찾아낸다. 다양한 아이템을 시스템이 제시하고 사용자의 피드백을 수집하며, 이를 통해 시스템은 보상을 최대화하는 추천 전략을 학습하고 점진적으로 사용자에게 맞춤형 추천을 제공한다. 이러한 심층 강화학습 기반의 시스템은 최적화된 제품이나 콘텐츠를 추천하는 데 큰 역할을 할 것이다.

많은 대학에서 대학 혁신사업을 통해 다양한 콘텐츠와 서비스를 제공하고자 노력한다. 학내의 조직들은 기본적으로 학생들을 올바르게 지도하고 지원하는 것이 목표지만, 학생들의 전공과목 선택과 학습을 더욱 효율적으로 돕기 위해 노력하고 있다. 전공과목의 다양성과 수요를 파악하기 어려워서 전공 선택에 어려움을 겪는 학생들도 많다. 또한, 학생들은 전공과목을 졸업 요건이나 학점을 충족하기 위한 강의로만 생각하는 경우도 적지 않다. 이에 본 논문에서는 대학 내의 전공과목 수강 이력 데이터를 활용하여 각 학생의 전공 수강 패턴을 분석하고, 이를 심층 강화학습 기반으로 전공과목을 추천하는 시스템을 제안한다. 이를 통해 학생 개개인에게 맞는 전공과목을 추천하여 전공 선택을 돕고, 학생들의 학습 효율과 전공에 대한 흥미를 높일 수 있다.

이러한 방법을 대학교의 전공과목 추천 시스템에 강화학습을 적용하면 많은 학생에게 개인화된 지도와 학습 환경을 제공할 수 있을 것이다. 학생들의 수강 이력 데이터를 분석하여 각 학생의 특성과 성향을 파악하고, 이를 기반으로 전공과목 추천이나 학습 지원을 제공하는 시스템의 개발에 심층 강화학습 기반의 추천 시스템은 사용자의 취향과 요구를 더욱 정확하게 파악하고, 맞춤형 추천을 제공함에 있어 도움을 줄 수 있다. 본 연구는 심층

강화학습을 기반으로 대학의 전공과목 추천 스케줄링 시스템을 제안하고, 이를 통해 학생들의 학습 경험을 개선하고 최적의 맞춤형 추천을 제공하는 방법에 대해 기여하고자 한다.

본 논문은 다음과 같이 구성되어 있다. II장에서는 연구 데이터 및 관련 연구에 대해 설명하고 III장에서는 심층 강화학습 기반의 제안 방법에 대해 기술하였다. IV장에서는 시뮬레이션 결과를 확인하고 V장에서는 결론을 통해 마무리 짓는다.

## II. 연구 데이터 및 관련 연구

### 1. 연구 데이터

본 연구에서는 심층 강화학습 기반의 전공과목 추천 시스템을 활용하기 위해 2016년부터 2022년까지의 H 대학 재학생과 졸업생의 수강 이력 데이터를 사용하였다. 이 데이터는 학생의 학번, 수강 년도, 수강 학기, 학수 번호, 이수 구분, 과목명, 성적 등 16개의 변수로 구성되어 있다. 또한, 강의 평가 데이터를 활용하여 각 강의의 특성을 나타내는 전공과목의 기본정보 데이터를 분석에 활용하였다.

연구에 사용된 데이터는 H 대학의 정보처에서 추출되었으며, 데이터의 안전성과 보안을 위해 적절한 처리가 이루어졌다. 수강 이력 데이터와 강의 평가 데이터는 강화학습 기반의 전공과목 추천 시스템의 테스트를 위해 필요한 속성과 변수들을 포함하고 있다. 표 1은 본 연구에 사용된 데이터의 속성을 기재하였다.

표 1. 데이터 속성 및 설명.

Table 1. Data attributes and descriptions.

속성명	설명	범위
교과과정명	교과과정의 이름	문자형
분반	과목의 학습 번호	문자형
교과목명	분반 정보	정수형
이수 학년	교과목의 이름	정수형
이수 구분	과목을 이수해야 하는 학년	문자형
기준학점	과목의 이수 구분 (전공필수, 전공 선택 등)	정수형
담당 교수	과목을 담당하는 교수	문자형
수강인원수	과목을 수강하는 학생 수	정수형
청강신청	청강신청 여부	정수형

## 2. 관련 연구

추천 시스템 분야에서의 연구들은 다양하며, 일부는 대학 수업 데이터를 활용해 최적의 수업 선택 전략을 도출하는 데 중점을 두었다. 기존 추천 시스템의 문제점은 모델의 결과를 실제 환경에서 확인하고 검증하기 어렵다는 점이다. 이는 모델의 예측이 불확실하며 미래에 발생할 수도 있고 아닐 수도 있다는 것 때문이다. 따라서, 추천 시스템은 실제 사용자에게 제공된 후 사용자의 반응 데이터를 수집하여 지속해 개선하는 방식이 필요하다.

이러한 문제를 해결하기 위해, 본 연구에서는 심층 강화학습을 대학 전공과목 선택에 적용하였다. 먼저, 학생들의 과거 수강 이력 데이터를 기반으로 학습 환경을 설정하였다. 그 후에는 어떤 상황에서 어떤 수업을 선택하면 가장 이상적인 결과를 얻을 수 있는지에 대한 정책을 학습하였다.

본 연구는 H 대학의 학생들을 대상으로 서비스를 제공하기 위해, 이들의 사용 패턴과 피드백을 분석하여 서비스의 문제점을 파악하였다. 이를 통해, 학생들이 실제로 추천받은 과목을 언제 선택하였는지, 그 선택에 만족하였는지 등의 정보를 수집하고, 이를 토대로 강화학습 모델의 성능을 평가하고 개선하는 과정을 거쳤다. 이렇게 함으로써, 심층 강화학습 기반의 수업 추천 시스템이 실제 환경에서 어떠한 작용을 하는지를 검증하였다.

## III. 제안 방법

### 1. 원-핫 인코딩 기반의 데이터 전처리

본 논문의 첫 단계는 대학교의 수업 및 학생 데이터를 수집하고 전처리하는 것이다. 이 단계는 추천 시스템이 학생들의 수강 과목을 적절히 추천할 수 있도록 필요한 정보를 확보하는 근간을 제공한다.

데이터 수집 과정에서는 학생들의 과거 수강 이력, 각 과목의 학점, 성적 분포, 과목 간 선후수 관계 등이 고려되어야 한다. 이렇게 수집된 데이터는 학생들이 어떤 과목을 언제 수강했는지, 그 과목을 수강한 결과 어떤 성적을 받았는지, 그리고 다른 과목들과 어떤 관계가 있는지에 대한 중요한 정보를 포함하고 있다.

다음으로, 수집된 데이터는 전처리 과정을 거쳐야 한다. 이 과정에서는 불필요한 정보를 제거하고, 데이터의 형식을 통일하며, 결측치나 이상치를 처리하는 등의 작업을 진행한다. 또한, 필요에 따라 특정 데이터를 추출하

거나 변환하여 모델이 학습할 수 있는 형태로 만드는 작업도 이루어진다.

본 연구에서는 원-핫 인코딩 기법을 사용하여 전처리를 진행하였다. 이 기법은 범주형 데이터를 컴퓨터가 이해할 수 있는 숫자 형태로 변환한다<sup>[6]</sup>. 예를 들어, 각 과목 데이터를 원-핫 인코딩을 적용하면, 각 과목이 별도의 특징으로 간주 되어 0 또는 1의 이진값을 가진다. 학생이 특정 과목을 선택하면 해당 과목에는 1이 할당되고, 모든 다른 과목에는 0이 할당된다.

표 2. 전공과목의 원-핫 인코딩 형식.

Table 2. One-hot encoding format for major subjects.

자료구조	알고리즘	컴퓨터 네트워크	운영체제
1	0	0	0
0	1	0	0
0	0	1	0
0	0	0	1

이렇게 전처리된 데이터는 심층 강화학습에서 효과적으로 활용할 수 있다. 각 학생의 상황에 따라 어떤 과목을 수강하면 최적의 결과에 대한 정보를 학습하여, 학생들에게 최적의 과목을 추천할 수 있게 된다.

### 2. 환경 설정

본 연구에서는 대학의 수업 시스템을 환경으로 설정하여 심층 강화학습을 적용하였다. 환경 설정은 학생(에이전트)이 선택할 수 있는 과목의 선택지를 제공하며, 선택된 과목에 따른 보상을 결정하는 역할을 한다. 보상은 해당 과목을 이수했을 때 얻을 수 있는 학점을 기반으로 한다. 또한, 과목 간의 선후수 관계는 에이전트가 과목을 선택하는 데 있어 제약 조건으로 작용한다. 예를 들어, 어떤 과목을 수강하기 위해선 특정 과목을 먼저 이수해야 하는 경우, 에이전트는 그 선행 과목을 먼저 수강하는 행동을 선택해야 한다.

본 연구에서는 심층 강화학습의 에이전트를 학생으로 설정하였다. 에이전트인 학생은 매 학기 어떤 과목을 수강할지를 결정하게 된다. 이러한 문제 상황은 시간이 지나면서 상태가 전이되고, 에이전트의 결정에 따라 보상이 결정되는 마코프 결정 프로세스(Markov Decision Process, MDP)로 모델링 될 수 있다<sup>[7]</sup>. 그림 1은 환경과 에이전트로 구성된 강화학습의 구조를 보여준다.

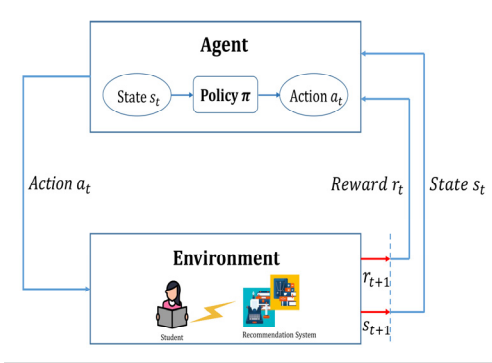


그림 1. 강화학습의 구조.  
Fig. 1. Structure of Reinforcement Learning.

에이전트는 시행착오를 통해 학습하며, 어떤 행동이 가장 높은 보상을 주는지 알아낸다. 이는 마치 학생이 어떤 공부 방법이 시험에서 가장 좋은 성적을 내는지 배우는 과정과 유사하다. 이러한 과정을 통해 에이전트는 정책을 찾게 되며, 이 정책은 주어진 상태에서 어떤 행동을 취할 것인지를 결정한다. 이를 구성하는 주요 구성요소들은 다음과 같다.

- 상태(State): 상태는 학생의 현재 상황을 반영한다. 현재까지 수강한 과목들, 그 과목들의 성적, 남은 학점, 현재까지 획득한 총 학점 등으로 구성된다.
- 행동(Action): 행동은 학생이 다음에 수강할 과목을 선택하는 것입니다. 수강 가능한 과목의 집합이 학생의 행동 공간을 형성한다.
- 보상(Reward): 보상 함수는 학생이 선택한 과목을 이수하였을 때 받게 되는 학점을 기반으로 합니다. 더 나은 성적을 받은 과목은 더 높은 보상을 제공하게 된다.

이러한 설정을 통해, 에이전트인 학생은 강화학습 알고리즘을 통해 자신의 수강 과목을 선택하는 정책을 학습하게 된다. 이는 본 연구의 목표인 학생 개개인에게 최적화된 수업 추천을 실현하는 핵심 방법이라 할 수 있다.

### 3. Dueling Deep Q-Network

본 논문에서는 Dueling DQN을 알고리즘을 활용하였다. Dueling DQN은 기존의 DQN 알고리즘에 추가로 가치함수를 분리하여 계산하는 구조를 가진다<sup>[8-9]</sup>. 이는 행동의 가치를 직접 계산하는 대신, 상태의 가치와 행동의 이점을 별도로 계산하여 학습의 효율성을 향상하는

기법이다. 그림 2는 제안하는 Dueling DQN 기반의 추천 알고리즘 의사 코드이다.

```

Initialize replay memory D
Initialize action-value function Q with random weights
Initialize target action-value function Q-bar with weights theta-bar = theta
for episode = 1, M do
  Initialize state s
  for t = 1, T do
    Choose action a from state s using policy derived from Q
    Take action a, observe reward r and next state s'
    Store transition (s, a, r, s') in memory D
    Sample random mini-batch of transitions (s_j, a_j, r_j, s'_j) from D
    Set y_j = r_j for terminal s'_j
    y_j = r_j + gamma * max_{a'} Q(s_j, a'; theta) for non-terminal s'_j
    Perform a gradient step on (y_j - Q(s_j, a_j; theta))^2 with respect to theta
    Every C steps, reset Q-bar = Q
    s = s'
  end for
end for
    
```

그림 2. Dueling DQN 기반의 전공과목 추천 코드.  
Fig. 2. Dueling DQN-based major subject recommendation code.

추천 알고리즘은 학습 데이터를 기반으로 Dueling DQN이 수강 과목을 학습하는 방법을 보여준다. 각 에피소드에서 에이전트는 정책을 통해 행동을 선택하고, 보상과 다음 상태를 관찰한다. 이후에 경험 리플레이를 통해 데이터를 메모리에 저장하고, 그 데이터를 사용하여 행동 가치함수 Q를 업데이트한다. 이 과정을 반복하면서 에이전트는 수강 과목을 추천하는 정책을 학습하게 된다.

Dueling DQN 알고리즘을 적용하면 에이전트는 각 상태에서 수강할 과목을 선택하는 행동의 가치를 추정하게 된다<sup>[9]</sup>. 즉, 학생은 상태와 행동의 가치를 고려하여 어떤 과목을 수강할 것인지 결정하게 된다. 이렇게 계산된 Q-가치는 각 상태에서 어떤 과목을 수강해야 하는지를 결정하는 데 사용된다. 즉, 상태에 따른 가치와 행동에 따른 가치를 독립적으로 평가하고 추정하므로, 더욱 상세하고 정확한 가치 추정이 가능하다는 것이다<sup>[10]</sup>. 이렇게 분리된 Value와 Advantage는 Dueling DQN에서 수식(1)에 따라 계산되며, 이 과정을 통해 더욱 정확한 학습이 가능하다<sup>[11]</sup>.

$$Q(s_t, a_t) = V(s_t) + A(s_t, a_t) \quad (1)$$

여기서  $V(s_t)$ 는 특정 상태  $s_t$ 에서 얻을 수 있는 보상의 기댓값(Expected Value)을 나타내며,  $A(a_t)$ 는 특정 행동의 상대적인 중요성을 나타내는 Advantage 기댓값이다<sup>[12]</sup>. 기존의 DQN과 달리 Dueling DQN은 Value 함수와 Advantage 함수를 별도로 학습하여 상태의 가치와

각 행동의 중요성을 추정한다. 이렇게 분리된 추정값으로 에이전트는 가치가 큰 행동을 선택하여 학습을 진행할 수 있다. 이를 통해 에이전트는 보상을 얻기에 좋지 않은 상태에서도 효율적으로 학습할 수 있으며, 최적의 정책을 더 빠르게 학습할 수 있다. Dueling DQN에서 사용되는 Q-value를 계산하는 수식 (2)와 같이 정의된다<sup>[13]</sup>.

$$Q(s_t, a_t; \theta) = V(s_t; \theta) + A(s_t, a_t; \theta) - \frac{1}{|A|} \sum_{a_{t+1}} A(s_t, a_{t+1}; \theta) \quad (2)$$

$s_t$ 는 현재 상태이고,  $a_t$ 는 현재 상태에서 선택할 수 있는 행동이다.  $V(s)$ 는 상태  $s_t$ 에서의 기댓값으로, 어떤 행동을 취하던 상태  $s_t$ 에서 기대할 수 있는 보상의 평균값이다.  $A(s_t, a_t)$ 는 행동  $a_t$ 가 선택되었을 때의 상태  $s_t$ 에서 얻을 수 있는 Advantage 값을 의미하며,  $A(s_t, a_t) = Q(s_t, a_t) - V(s_t)$ 로 계산된다.  $A(s_t, a_{t+1})$ 는 다음 상태  $s_{t+1}$ 에서의 얻을 수 있는 Advantage 값으로, Target Network를 이용하여 추정된다.  $|A|$ 는 취할 수 있는 행동의 수를 나타내며, 전체 행동의 평균값을 구하기 위한 정규화 상수이다<sup>[14]</sup>.

이를 통해, Dueling DQN 기반의 강화학습 알고리즘이 학생들에게 개인화된 수강 과목 추천을 제공하게 된다. 이는 본 연구의 핵심 목표를 충족시키는 중요한 과정이다. 에이전트는 Q-value가 가장 높은 행동을 선택하며, Dueling DQN을 사용한 추천 시스템은 학생에게 가장 이상적인 수강 과목을 추천할 수 있게 된다.

#### IV. 실험 및 결과

제안 방법의 검증을 위해, 다양한 실험을 수행하였다. 실험 데이터는 대학교에서 수집된 실제 학생들의 수강 이력 데이터를 사용하였다. 이 데이터는 학생들의 개인 정보를 제외하고, 과목의 선택, 성적, 학점 등의 정보만을 포함하였다. 제안 방법의 성능을 평가하기 위해, 우리는 학생들의 총 학점을 최대화하는 것을 목표로 하였다. 즉, 우리의 모델이 추천한 과목을 수강한 학생들의 총 학점이, 무작위로 과목을 선택한 학생들의 총 점수보다 높아야 한다.

우선 추천 시스템의 성능을 평가하기 위해 강화학습 알고리즘이 수렴하는 과정을 관찰하는 것이 중요하다.

이 과정을 통해 우리는 학습이 제대로 이루어지고 있는지, 그리고 언제 알고리즘이 최적의 수강 과목 선택 정책을 찾아냈는지를 알 수 있다.

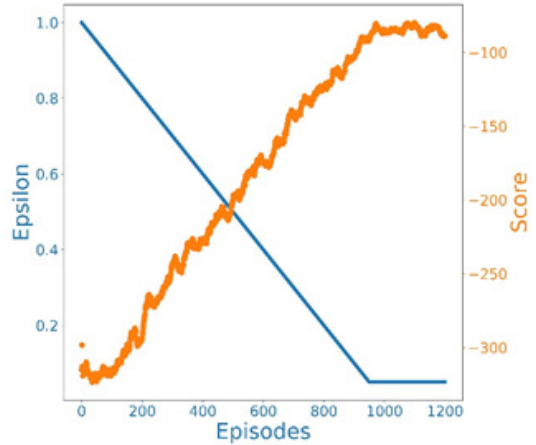


그림 3. Dueling DQN의 수렴 분석.  
 Fig. 3. Convergence Analysis of Dueling DQN

그림 3은 Dueling DQN의 학습 관련으로 수렴이 제대로 되는지를 보여준다. 초기에는 에이전트가 임의의 정책을 따르기 때문에 총 보상이 낮을 수 있다. 그러나 학습이 진행됨에 따라, 에이전트는 점점 더 높은 보상을 얻을 수 있는 수강 과목 선택 정책을 찾아낸다. 그 결과, 총 보상은 점차 증가하고 일정한 수준으로 수렴하게 된다. 이러한 수렴 과정을 관찰함으로써, 우리는 추천 시스템이 실제로 학생의 수강 과목 선택을 최적화하는 방향으로 학습하고 있음을 확인할 수 있다.

표 3. 추천 시스템의 실험 결과 비교 테이블.

Table 3. Comparison table of experimental results of recommender system.

시스템	정확도	예측 성능	추천 다양성
통계 기반 추천 시스템	0.72	0.78	0.60
Dueling DQN 기반 추천 시스템	0.84	0.87	0.75

통계 기반 추천 시스템과 Dueling DQN 기반 추천 시스템의 성능을 비교한 결과에서 Dueling DQN 기반 추천 시스템이 통계 기반 추천 시스템에 비해 더 높은 정확도를 보여준다. Dueling DQN 기반 추천 시스템은 정확한 전공 과목을 예측하는 능력이 향상되었음을 보여준

다. 이는 Dueling DQN이 학생의 학과, 학년, 수강 이력 등의 정보를 활용하여 맞춤형 추천을 수행하는 데에 효과적이었음을 볼 수 있다.

표 4. 전공과목 추천 시스템 비교.

Table 4. Comparison of major subject recommendation system.

시스템	재현율	F1점수	Top-K 정확도
통계 기반 추천 시스템	0.72	0.78	0.60
Dueling DQN 기반 추천 시스템	0.84	0.87	0.75

또한, 예측 성능 지표인 F1 점수를 비교해보면, Dueling DQN 기반 추천 시스템이 더 높은 점수를 보여준다. 이는 Dueling DQN 기반 시스템이 정밀도와 재현율을 균형 있게 유지하는 데에 더 효과적이라는 것을 나타낸다. 그리고 추천 다양성 지표를 통해서 각 학생의 개별적인 특성을 고려하여 다양한 전공 과목을 추천하는 데에 성공했음을 확인할 수 있다. Top-K 정확도에서는 Dueling DQN 기반 추천 시스템이 0.88의 높은 정확도를 보였으며, 이는 Dueling DQN 기반 추천 시스템은 상위 순위에서 더 정확한 추천을 제공하는 것을 볼 수 있다.

결과적으로, 실험 결과는 Dueling DQN 기반 추천 시스템이 통계 기반 추천 시스템보다 더 우수한 성능을 보여준다는 것을 보여준다. Dueling DQN 기반 시스템은 정확도, 예측 성능, 추천 다양성 측면에서 모두 향상된 결과를 제공하며, 학생들에게 더욱 개인화된 전공과목 추천을 제공할 수 있을 것이다.

## V. 결 론

본 논문에서 제안하는 심층 강화학습 기반의 수강 과목 추천 시스템으로 기존의 추천 시스템보다 학생들의 과목 선택 과정을 지원할 방법을 제안하였다. 제안하는 시스템은 학생들의 과거 수강 이력, 각 과목의 학점, 성적 분포, 과목 간 선호 수 관계 등을 기반으로 과목 추천을 수행하였으며, 실험 결과에 따르면, Dueling DQN 기반 추천 시스템이 통계 기반 추천 시스템에 비해 더 우수한 성능을 보여주었다. 이는 강화학습 알고리즘인 Dueling DQN이 학생들의 학과, 학년, 수강 이력 등의 정보를 활용하여 개별적인 특성을 반영하고 맞춤형 전공

과목을 추천하는 데에 효과적이었음을 나타낸다.

따라서 심층 강화학습 기반의 추천 시스템을 통해 전공과목 추천을 제공하는 데에 더 적합하다는 결론을 도출할 수 있다. Dueling DQN은 강화학습 알고리즘을 통해 개별 학생의 특성과 선호를 학습하여 정확하고 다양한 전공과목을 추천하는 시스템을 구현할 수 있을 것이다. 이러한 결과는 학생들이 전공과목 선택에 있어 보다 높은 만족도와 성공적인 학업 경험을 제공할 수 있는 중요한 지표가 될 것이다.

## References

- [1] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M. "Human-level control through deep reinforcement learning", Nature, 518(7540), 529-533, 2015.  
DOI: <https://doi.org/10.1038/nature14236>
- [2] Kim, Songhee, Sunhye Kim, and Byungun Yoon. "Deep Learning-Based Vehicle Anomaly Detection by Combining Vehicle Sensor Data." Journal of the Korea Academia-Industrial cooperation Society(JKAIS), Vol. 22(3), pp. 20-29, 2021.  
DOI: <https://doi.org/10.5762/KAIS.2021.22.3.20>
- [3] Eun-Gyu Ham, Chang-Bok Kim, "Model Implementation of Reinforcement Learning for Trading Prediction Using Deep Q Network," The Journal of KIIT, Vol. 17, No. 4, pp. 1-8, 2019.  
DOI: [10.14801/ikiit.2019.17.4.1](https://doi.org/10.14801/ikiit.2019.17.4.1)
- [4] Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." nature 529.7587, pp. 484-489, 2016.  
DOI: [10.1038/nature16961](https://doi.org/10.1038/nature16961)
- [5] Schulman, J., Levine, S., Abbeel, P., Jordan, M., Moritz, P. "Trust region policy optimization". In Proceedings of the 32nd International Conference on Machine Learning (ICML-15), pp. 1889-1897, 2015.  
DOI: [10.5555/3045118.3045280](https://doi.org/10.5555/3045118.3045280)
- [6] Hessel, M., Modayil, J., Van Hasselt, H., Schaul, T., Ostrovski, G., Dabney, W., Silver, D. "Rainbow: Combining improvements in deep reinforcement learning", In Thirty-Second AAAI Conference on Artificial Intelligence, 2018.  
DOI: [10.1609/aaai.v32i1.5503](https://doi.org/10.1609/aaai.v32i1.5503)
- [7] Castro, P. S., Precup, D. "Deep reinforcement learning for list-wise recommendations", In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 2077-2085, 2018.  
DOI: [10.1145/3269206.3269260](https://doi.org/10.1145/3269206.3269260)
- [8] Zeng, Y., Zhao, Q., Wang, X., "A deep reinforcement learning based recommendation framework with item-level adaptive sequence modeling", IEEE

Transactions on Knowledge and Data Engineering, 33(3), 1114-1128, 2020.  
DOI: [10.1109/TKDE.2019.2899727](https://doi.org/10.1109/TKDE.2019.2899727)

- [9] J. Park, "Self-Awareness and Coping Behavior of Smartphone Dependence among Undergraduate Students," Journal of the Korea Academia-Industrial cooperation Society (JKAIS), vol. 22, no. 2, pp. 336-344, 2021.  
DOI: <https://doi.org/10.5762/KAIS.2021.22.2.336>
- [10] Mabrouk, R. M., Kamal, E. M. "Reinforcement learning for dynamic personalized course recommendation systems", Journal of Ambient Intelligence and Humanized Computing, 1-13, 2021.  
DOI: 10.1007/s12652-021-03486-3
- [11] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning", arXiv preprint arXiv:1312.5602, 2013.  
DOI: <https://doi.org/10.48550/arXiv.1312.5602>
- [12] Hausknecht, Matthew, and Peter Stone. "Deep recurrent q-learning for partially observable mdps", 2015 aai fall symposium series. 2015.
- [13] Wang, Ziyu, et al. "Dueling network architectures for deep reinforcement learning." International conference on machine learning. PMLR, 2016.
- [14] Hessel, Matteo, et al. "Rainbow: Combining improvements in deep reinforcement learning" Proceedings of the AAAI conference on artificial intelligence. Vol. 32. No. 1, 2018.  
DOI: <https://doi.org/10.1609/aaai.v32i1.11796>

## 저 자 소 개

### 임 덕 선(정회원)



- 2004년 2월 : 한양사이버대학교 컴퓨터공학과 학사
- 2023년 8월 : 한양대학교 일반대학원 컴퓨터·소프트웨어학과 석·박사
- 2020년 : 한양사이버대학교 컴퓨터공학과 교수
- 관심분야 : 빅데이터, 강화학습, 옛지 컴퓨팅, Anomaly Detection

### 민 연 아(정회원)



- 2013년 2월 : 동국대학교 컴퓨터공학과 박사
- 2016년 ~ 2019년 : 가천대학교 소프트웨어 학과 교수
- 2020년 1월 ~ 현재 : 한양사이버대학교 응용소프트웨어학과 교수
- 관심분야 : 빅데이터, Blockchain, 원격교육, 심층 신경망

### 임 동 균(정회원)



- 1985년 2월 : 한양대학교 전자통신공학과 학사
- 1987년 2월 : 한양대학교 전자통신공학과 석사
- 2001년 2월 : 한양대학교 전자통신공학과 박사
- 2003년 3월 ~ 현재 : 한양사이버대학교 컴퓨터 소프트웨어 공학부 교수
- 관심분야 : 온라인 교육, 교육 콘텐츠, 컴퓨터 응용