

연구논문

직업계고 졸업자의 진로이행과정 유형 및 주요 예측 변인 탐색: 시퀀스 분석과 설명가능한 인공지능을 활용하여*

백예은** · 정혜원***

본 연구에서는 직업계고 졸업자의 진로이행과정 유형을 탐색하고, 그 유형을 예측하는 변인을 탐색하였다. 이를 위해 한국교육고용패널 II의 3차~7차년도 자료를 활용하여 현재 학력(대학 재학 여부), 근무 형태(정규직·비정규직), 구직활동 경험 여부, 취업을 위한 준비 활동 경험 여부와 같은 진로이행상태의 변화에 대하여 시퀀스 분석을 수행하고, 이를 k-means 군집분석을 통해 유형화하였다. 이에 도출된 진로이행과정 유형을 예측하는 변인을 탐색하기 위하여 자동화 머신러닝을 통해 가장 우수한 성능을 보이는 머신러닝 기반의 예측 모형을 도출하고, SHAP 알고리즘과 로지스틱 회귀분석을 통하여 예측 모델에 투입한 다양한 개인 특성, 고등학교 생활 및 가정 생활 변인과 진로이행과정 유형 간의 관계를 살펴보았다. 주요 결과를 정리하면 다음과 같다. 첫째, 직업계고 졸업자의 진로이행과정의 유형은 '초기 정규직 정착형', '점진적 정규직 정착형', '대학 진학형', '장기 NEET형' 4가지로 구분되었다. 둘째, '초기 정규직 정착형'을 준거집단으로 설정하고, '점진적 정규직 정착형', '대학 진학형', '장기 NEET형'을 비교집단으로 각각 설정했을 때, 자격증 개수가 많고 취업·자격증 방과후학교, 취업상담에 참여한 경험이 있는 학생일수록 '초기 정규직 정착형'에 속할 확률이 높은 것으로 나타났다. 셋째, 진로 역량이 낮을수록 '장기 NEET형'에 속할 확률이 높았으며, 희망교육수준이 높을수록 대학 진학형'에 분류될 확률이 높은 것으로 나타나는 등 비교집단 설정에 따라 차별적으로 나타난 연구결과 또한 확인할 수 있었다. 그리고 분석결과를 통해 나타난 주요 예측 변인과 진로이행과정 간 관계를 바탕으로 진로이행과정 유형에 맞는 직업계고 졸업자의 진로 지원을 위한 시사점을 논의하였다.

주제어: 진로이행, 직업계고 졸업자, 시퀀스 분석, AutoML 알고리즘, SHAP 알고리즘

* 이 논문은 2024 KRIVET 패널 학술대회 발표내용을 수정·보완하여 작성하였음.

** 충남대학교 교육학과 박사과정 수료(byeunn7@gmail.com), 제1저자.

*** 충남대학교 교육학과 교수(chw7@cnu.ac.kr), 교신저자.

I. 서론

정부에서는 매년 「직업계고 졸업자 취업통계조사」를 실시하여 직업계고등학교 졸업자의 취업 및 대학 진학 여부 등의 졸업 후 진로 현황을 발표하고 있다. 직업계고등학교는 중등 직업교육단계에서 산업 수요에 맞춘 인력 양성을 목적으로 하는 고등학교로, 특성화고등학교, 마이스터고등학교, 전문교과를 운영하는 일반고를 포함한다(한국교육개발원·교육부 2024). 최근 조사결과에 따르면, 직업계고 졸업 후 취업이 아닌 대학 진학을 선택하거나 취업, 대학 진학을 모두 하지 않는 졸업자의 비율이 증가하는 등 오늘날 직업계고 졸업자의 진로이행 양상이 변화하고 있는 것으로 나타났다. 구체적으로 2022년부터 취업률이 2.1%, 0.4%씩 지속적으로 감소하는 추세를 보인 반면, 대학 진학률은 1.8%, 1.0%씩 증가하는 추세에 있으며 2024년의 경우 졸업자 중 진학자나 입대자를 제외한 21%가 군입대를 포함하여 취업 혹은 대학에 진학하지 않은 상태인 것으로 보고되었다(한국교육개발원·교육부 2023, 2024).

이에 따라 교육학 분야에서는 직업계고 졸업자의 진로이행 상태를 취업 여부로만 구분하는 것의 한계를 인식하여, 직업계고 졸업자의 경력 경로나 노동시장 지위 변화의 유형을 분석한 연구들이 수행되었다(김남식 2023; 김진모 외 2018; 류지은 외 2022 등). 그러나 최근 사회적 문제로 제기되는 청년 NEET(Not in Employment, Education, or Training)와 같이 단순히 취업을 하지 않을 뿐만 아니라 정규교육을 받거나 구직활동을 하지 않는 청년의 증가와 짧은 근속기간과 같은 현실적 문제를 반영하는 연구는 부족한 실정이다. 이에 본 연구에서는 직업계고 졸업자의 대학 재학 여부, 근무 형태(정규직·비정규직), 구직활동 및 취업을 위한 준비 활동 경험 여부(NEET 여부) 등 다양한 지표를 활용하여 직업계고 졸업자의 진로이행과정을 유형화하고자 한다.

이를 위해 본 연구에서는 한국직업능력연구원에서 노동시장 진입 과정에 대한 보다 체계적이고 다양한 자료를 구축하기 위해 2016년부터 수행하고 있는 한국교육고용패널조사 II(Korean Education & Employment Panel II, KEEP II)의 3차년도(2019년 조사)부터 7차년도(2023년 조사) 자료를 활용하여 시퀀스 분석(sequence

analysis)를 실시하고자 한다. 시퀀스란 여러 시점에 걸쳐 나타나는 연쇄적인 상태의 변화를 의미하며(Abbott 1983), 사회과학분야에서 시퀀스 분석은 주로 개인의 중단적인 상태 변화를 추출하고, 유사한 패턴을 지닌 시퀀스로 유형을 추출하는 단계로 확장되어 활용된다(김남식 2023; 노범래·문영민 2020; 오소옥 외 2023 등). 이에 본 연구에서는 졸업 이후 5년간 변화하는 취업, 대학 재학, 구직활동 등과 같은 진로이행상태의 변화를 시퀀스로 설정하여 직업계고 졸업자의 진로이행과정을 살펴 보고, 시퀀스 분석을 통해 도출된 개인의 시퀀스 간 거리 행렬을 기반으로 k-means 군집분석 알고리즘을 적용하여 직업계고 졸업자의 진로이행과정을 유형화함으로써 직업계고 졸업자의 중단적인 진로이행과정의 유형을 탐색하고자 한다. 더불어 본 연구에서는 분류된 유형을 머신러닝 기반의 예측 모형의 결과 변인으로 설정하여 진로이행과정에 영향을 미치는 다양한 개인 및 환경 변인을 탐색하고, 전통적인 통계모형이 갖는 제한적인 변인 활용과 모형 설정의 한계로부터 비교적 자유로운 머신러닝 모델을 활용하여 KEEP II에서 제공하는 다양한 개인 특성, 고등학교 생활 및 가정 생활 변인을 모형에 투입하고자 하였다. 이때, 머신러닝 기반의 예측 모형 구축에 있어서 다양한 머신러닝 모델에 대하여 하이퍼파라미터 탐색 및 교차 검증을 수행함으로써 모델 간 성능 비교 후 최적의 예측 모델을 탐색할 수 있도록 하는 자동화 머신러닝(Automated Machine Learning, AutoML) 알고리즘을 활용한다. 이를 통해 성능이 가장 우수한 최적의 직업계고 졸업자의 진로이행과정 예측 모델을 탐색하고, 최종 모델에 대하여 진로이행과정 예측에 높은 기여를 하는 것으로 나타난 변인을 도출한다.

한편 사회과학 분야에서는 연구결과를 해석할 때 종속 변인에 대한 예측 변인의 설명과 해석이 중요하다. 이에 본 연구에서는 최적의 머신러닝 예측 모델 탐색에서 더 나아가 Lundberge et al.(2020)이 제안한 설명 가능한 인공지능의 일종인 SHAP 알고리즘(Shapley Additive ExPlanation)을 통해 주요 예측 변인과 진로이행과정 유형 간 관계를 시각화하고, 전통적인 회귀분석 방법인 로지스틱 회귀분석을 통해 주요 예측 변인의 통계적 유의성을 검증함으로써 개인 특성, 고등학교 생활 및 가정 생활 변인과 진로이행과정 유형 간 관계를 설명한다. 이를 통해 도출되는 주요 분석결과를 기반으로 오늘날 다양화된 직업계고 졸업자의 진로이행과정에 맞는 진로 지원 정책 수립을 위한 시사점을 제공하는 것이 본 연구의 목적이다. 연구목적을 달성하기 위하여 설정한 연구문제는 다음과 같다.

- 연구문제 1. 직업계고 졸업자의 졸업 후 진로이행과정은 어떻게 유형화되는가?
연구문제 2. 직업계고 졸업자의 진로이행과정 유형을 예측하는 주요 변인은 무엇인가?
연구문제 3. 직업계고 졸업자의 진로이행과정 유형과 주요 예측 변인 간 어떠한 관계가 있는가?

II. 선행연구 고찰

1. 직업계고 졸업자의 진로이행

진로이행이란, 개인 본인이 진로를 선택·결정하여 실제로 이행하는 것에 강조점을 둔 용어로(김현주·김준영 2010; 양준모·배관표 2024) 고등학교를 졸업한 학생은 고등교육기관인 대학에 진학하거나, 취업을 통해 노동시장에 진입하는 등의 진로이행과정을 거치게 된다. 교육학계에서는 ‘중등 직업교육단계에서 특정 분야에 종사하기 위한 인력 양성’이라는 직업계고의 설립 목적에 따라, 주로 노동시장에서의 지위 변화, 경력이동과 같은 변화 양상을 탐색한 연구가 수행되어 왔다. 이를 살펴보면, 김진모 외(2018)의 연구에서는 한국교육고용패널(이하 KEEP) I의 3차년도(2006년)~12차년도(2015년) 자료를 활용하여 직업계고 졸업자의 경력유형을 분석하였는데, 직장 규모와 정규직·비정규직 여부에 따라 ‘중소기업 정규직형’, ‘불안정 고용형’, ‘4년제 대학 진학형’, ‘대기업 정규직형’으로 구분하였으며, ‘중소기업 정규직형’의 비율이 가장 높았고, ‘대기업 정규직형’의 비율이 가장 낮은 것으로 나타났다.

또한 류지은 외(2022)의 연구에서는 KEEP I, II 자료 중 임금근로자를 중심으로 경력경로 유형을 구분하였는데, 2018년 졸업자(KEEP II)의 경우 저임금 비율이 높고 다른 유형보다 직종 및 일자리 이동이 가장 잦은 것으로 나타난 ‘저임금 일자리 이동형’, 첫 직장에서 계속해서 근무하는 ‘첫 일자리 유지형’, ‘저임금 일자리 이동형’보다 일자리 이동 비율은 낮지만 직종 이동이 계속해서 이루어지는 ‘직무 이동형’ 세 가지 유형이 도출되었으며, ‘직무 이동형’이 전체의 60% 이상을 차지하는 것으로 나타났다.

이와 더불어 KEEP I 자료를 활용하여 직업계고 졸업자의 10년간의 경력 경로를 분석한 김남식(2023)의 연구에서는 경력 유형을 대학, 취업, 직업교육훈련 참여로 구분하여, 고교 시절 높은 직업교육훈련 참여율을 보였으나 성인기로 갈수록 직업교육훈련에 참여하지 않지만 정규직 취업 상태를 유지하는 ‘연계저조형’, 직업계고 졸업 후에도 상대적으로 높은 직업교육훈련 참여율을 보이거나 정규직, 비정규직, 미취업의 상태가 혼재되어 나타나는 ‘연계유지형’, 졸업 후 직업교육훈련 참여가 상승하고 늦은 시기 정규직에 취업하는 ‘성인기참여상승형’, 그리고 성인기에 높은 직업교육훈련 참여율을 보이거나 취업과 미취업 상태가 혼재되어 나타나는 ‘성인기참여지속형’으로 구분하였다.

앞서 살펴본 선행연구를 종합하면, 직업계고 졸업자는 서로 다른 진로이행과정을 경험하며 연구에서 설정한 진로이행상태 지표 변인에 따라(예: 대학, 취업, 직업교육훈련 참여, 정규직·비정규직 취업 여부 등) 다양하게 유형화되는 것을 알 수 있었다. 이를 고려하여 본 연구에서는 직업계고 졸업자가 취업을 통한 노동시장 진입뿐만 아니라, 대학에 진학하거나 대학 진학과 취업을 하지 않은 채 취업을 유예하고 있는 비율이 증가하는 추세에 있고, 특히 실업 상태에서 교육 및 구직 활동을 하지 않는 청년들의 NEET 상태가 사회적 문제로 꾸준히 제기된다는 점을 고려하여 보다 다양한 진로이행상태를 포괄하여 분석하고자 한다. 이에 따라 본 연구에서는 현재 학력(대학 재학 여부), 근무 형태(정규직·비정규직), 구직활동 경험 여부, 취업을 위한 준비 활동 경험 여부와 같은 다양한 진로이행상태를 나타내는 지표를 활용하여 직업계고 졸업자의 진로이행과정을 분석하고자 한다.

2. 직업계고 졸업자의 진로이행과 개인, 학교, 가정 특성 간의 관계

그간의 선행연구에서는 직업계고 졸업자의 진로이행의 유형을 분류하고 유형별로 개인, 학교 및 가정 관련 특성을 비교하거나 유형 분류에 영향을 미치는 변인을 탐색하였다. 선행연구 결과를 개인, 학교, 가정의 범주로 나누어 정리하면 다음과 같다.

우선, 직업계고 졸업자의 진로이행 유형에 따른 개인 관련 특성을 비교한 선행연구를 살펴보면, 류지은 외(2022)의 연구에서는 직업계고 졸업자의 경력경로에 따라 ‘저임금 일자리 이동형’, ‘첫 일자리 유지형’, ‘직무 이동형’으로 분류하였는데 이때 각 유형별 성별 분포의 차이를 검증한 결과, ‘저임금 일자리 이동형’의 경우 여성이

가장 높은 비율(80%)을 보인 반면, ‘첫 일자리 유지형’, ‘직무 이동형’은 남성이 더 높은 비율인 것으로 나타나 남성이 여성보다 첫 일자리를 유지하는 경향이 있는 것으로 보고하였다. 한편, 직업계고 졸업자의 대학 진학 영향 요인을 탐색한 변수용·김경근(2012)의 연구에 따르면 개인의 희망교육수준이 높을수록 대학 진학 가능성이 상승하는 것으로 보고되었으며 박소영·정혜원(2022)의 연구에서는 진로개발역량이 높은 직업계고 졸업생의 취업률이 더 높은 것으로 보고하였다.

이어서, 학교와 관련된 선행연구 결과를 정리하면 다음과 같다. 김진모 외(2018)의 연구에서는 직업계고 졸업자의 ‘대기업 정규직형’, ‘중소기업 정규직형’, ‘불안정 고용형’, ‘4년제 대학 진학형’ 네 가지 경력유형에 따라 고등학교 생활과 관련된 변인(지각 빈도, 학업성취수준, 교우관계)의 차이를 검증하였다. 그 결과, ‘불안정 고용형’, ‘중소기업 정규직형’의 지각 빈도가 ‘대기업 정규직형’, ‘4년제 대학 진학형’보다 높게 나타났으며, 고등학교 시기 학업성취수준의 경우 ‘4년제 대학 진학형’의 학업성취수준이 그 외의 유형보다 유의하게 높은 것으로 나타났다. 이러한 결과와 관련하여 변수용·김경근(2012)의 연구에서도 학업성취도가 높을수록 직업계고 졸업자의 대학 진학 가능성도 향상하는 것으로 나타나, 직업계고 졸업자의 대학진학이 높은 학업성취도와 관련이 있음을 알 수 있었다. 한편, 긍정적인 교우관계의 경우에도 ‘4년제 대학 진학형’이 가장 높게 나타났으며, ‘불안정 고용형’이 가장 낮은 것으로 보고되어 ‘불안정 고용형’의 경우 고등학교 재학 중 성실성의 지표로 볼 수 있는 지각의 빈도가 가장 높고, 학업성취뿐만 아니라 긍정적인 교우관계의 수준 또한 가장 낮은 것을 확인할 수 있었다.

이와 더불어 류지은 외(2022)의 연구에 따르면, 직업계고 유형인 특성화고와 마이스터고 중, 특성화고 졸업자의 경우 ‘저임금 일자리 이동형’과 ‘직무 이동형’의 비율이 높게 나타난 반면, 마이스터고 졸업자의 경우 ‘첫 일자리 유지형’의 비율이 높게 나타나 고등학교 유형 간 차이가 존재함을 보고한 바 있다. 또한 김지효(2018)의 연구에서는 특성화고 졸업생의 진로이행 중 취업 여부와 관련하여 취업 결정 요인을 탐색하였는데, 그 결과 고등학교 시기 경험한 진로 체험에 해당하는 진로·진학 프로그램 유형 중 한 번이라도 경험한 바가 있다고 응답한 학생일수록, 졸업 후 취업할 확률이 높은 것을 확인할 수 있었다.

마지막으로, 가정과 관련된 선행연구 결과를 살펴보면, 직업계고 졸업자의 대학 진학 영향 요인을 탐색한 변수용·김경근(2012)의 연구에서는 부모의 교육수준이 4년제 대졸 이상이고, 월평균 사교육비가 높은 가정의 자녀일수록 직업계고 졸업 후

대학에 진학할 확률이 높은 것으로 나타났다. 한편 김지효(2018)의 연구에서 가정 관련 변인이 특성화고 졸업생의 취업결정에 미치는 영향력을 분석하였는데, 가정배경 변인은 유의한 영향을 미치지 않는 것으로 나타났으나 부모의 진학정보인식 수준이 높을수록 졸업 후 취업할 확률이 유의하게 높은 것으로 보고된다.

이상 살펴본 선행연구에서는 연구자의 관심 변인과 이론적 배경에 따라 연구 모형을 설정하여 직업계고 졸업자의 진로이행과정 유형, 취업 여부, 대학 진학 여부와 관련된 개인(성별, 희망교육수준, 진로역량), 고등학교(학교 유형, 학업성취도, 교우관계, 진로·진학 프로그램 경험), 가정(부모의 교육수준, 월평균 사교육비, 부모의 진학정보인식) 변인을 중심으로 분석을 수행하였다. 이에 본 연구에서는 선행연구에서 탐색된 변인뿐만 아니라 기존 연구에서 충분히 활용되지 않았던 패널자료에서 제공하는 다양한 변인과 진로이행 간의 관계를 탐색하고자, 전통적 모수 기반의 통계기법이 갖는 모형 설정의 한계에서 비교적 자유로운 비모수적 통계기법인 머신러닝 기법을 적용하고자 한다(김진희·김준엽 2021). 데이터를 기반으로 모형을 설정하는 머신러닝 기법을 통해 예측 모델에 투입한 수많은 변인 중, 예측 기여도가 높은 주요 예측 변인을 도출함과 더불어 더 나아가 로지스틱 회귀분석을 통한 통계적 유의성 검증 또한 실시하여 머신러닝을 통해 도출된 주요 예측 변인이 진로이행과정 유형에 미치는 영향력을 검증하고자 한다.

III. 연구방법

1. 연구대상

본 연구에서는 한국직업능력연구원에서 조사, 수집한 한국교육고용패널 II (Korean Education & Employment Panel II, KEEP II)의 3차년도(2019년 조사)부터 7차년도(2023년 조사) 자료를 활용하고자 한다. 한국교육고용패널 II는 2016년을 기점으로 고등학교 2학년 학생을 대상으로 조사되었다. 본 연구에서는 3차년도 조사시점을 기점으로 3, 4, 5, 6, 7차 5개 시점의 자료를 활용하여 직업계고 졸업 이후 진로이행상태 유형을 분류하였으며, 2차년도 조사시점의 개인 특성, 고등학교 생활 및 가정 생활 변인을 예측 변인으로 활용하였다. 최종적으로, 본 연구에서는

직업계고 졸업자 중 진로이행상태를 분류하기 위해 활용한 설문 문항(현재 학력, 근무 형태, 아르바이트 여부, 구직활동 경험 여부, 취업을 위한 준비 활동 경험 여부)에 모두 응답한 1,919명을 최종 분석대상으로 활용하였다.

2. 측정도구

1) 진로이행상태

본 연구에서는 직업계고 졸업자의 진로이행과정을 분석하기 위하여 3차년도부터 7차년도 5개 시점의 설문 문항을 활용해 각 시점의 진로이행상태를 코딩하였다. 이때 활용한 설문문항을 정리하면 <표 1>과 같다. 우선 현재 일자리 유무와 일자리 형태를 구분하기 위하여 ‘이 일자리에서 귀하의 근무 형태는 다음 중 무엇에 해당 됩니까?’ 문항을 통해 정규직과 비정규직 상태로 구분하였다. 이때, 아르바이트(수입을 목적으로 임시로 하는 일)의 경우 정규직·비정규직 집단에서 제외하였다. 정규직과 비정규직 상태를 제외한 학생들에 대하여 ‘2022년 12월 31일 기준으로 귀하는 학교를 어디까지 다녔습니까?’ 문항을 통해 대학에 재학 중인 집단과 미재학 집단으로 구분하였다.

이후 미재학 집단으로 분류된 학생들에 대하여 ‘2022년에 구직 활동(아르바이트 제외)을 해본 적이 있습니까?’와 같은 구직 활동 경험 여부, 취업을 위한 준비 활동 경험 여부(13개 문항)를 활용하여 구직 및 취업 준비 여부를 구분하고 구직 및 취업 준비를 하지 않은 집단에 대하여 ‘구직 활동을 하지 않은 주된 이유는 무엇입니까?’ 문항을 통해 비자발적 NEET 집단과 자발적 NEET 집단을 구분하였다. 이와 관련하여 진로이행상태 분류 방식 알고리즘에 대한 보다 구체적인 내용은 연구결과의 <그림 1>과 함께 서술하였다.

2) 예측변인

본 연구에서는 직업계고 졸업자의 진로이행과정의 유형을 예측하는 변인을 탐색하기 위하여 머신러닝 예측 모델을 구축하였다. 이때 예측 변인 투입에 있어서 본 연구에서 설정한 예측 모델의 종속 변인이 졸업 이후 시점인 3차년도부터 7차년도의 종단적인 진로이행과정 유형임을 고려하여, 3차년도 이전 시점인 2차년도에 조사된 변인을 활용하였으며, 2차년도의 경우 2016년 9월부터 2017년 12월까지의 고등학교 재학 기간 동안의 경험이 조사되었다.

<표 1> 진로이행상태 코딩 방식

구분	설문 문항	척도 및 코딩 방식
현재 일자리 유무	일자리 형태 (정규직 여부)	이 일자리(직장)에서 귀하의 근무 형태는 다음 중 무엇에 해당됩니까? ①정규직(파트타임), ②정규직(전일제) → 1로 코딩 ③비정규직(파트타임), ④비정규직(전일제) → 0으로 코딩
	아르바이트 여부	이 일자리(직장)에서 하시는 일은 아르바이트 (수입을 목적으로 임시로 하는 일)입니까? ①예, ②아니오 → ‘①예’라고 응답한 학생의 경우 정규직, 비정규직 집단에서 제외
대학 재학	현재 학력	2022년 12월 31일 기준으로 귀하는 학교를 어디까지 다녔습니까? ④2-4년제 전문대학 재학, ⑧4-6년제 대학교 재학, ⑬대학원(석사) 재학, ⑱대학원(박사) 재학 → 1로 코딩, 그 외 응답 0으로 코딩
구직 및 취업 준비활동 여부	구직활동 경험 여부	2022년에 구직 활동(아르바이트 제외)을 해본 적이 있습니까? ①예 → 1으로 코딩, ②아니오 → 0으로 코딩
	취업을 위한 준비 활동 경험 여부	‘1) 직업훈련(사설학원 포함) 참여 경험 여부’ 등 13개 문항 ①예, ②아니오 → 13개 중 ‘①예’ 응답이 하나 이상이면 1로 코딩, 그 외 0으로 코딩
	구직활동 안 한 이유	구직활동을 하지 않은 주된 이유는 무엇입니까? ①재학 중이어서, ⑦취업이 확정되어 발령 대기 상태여서, ⑧군입대 예정이어서, ⑨결혼·육아나 가사로 인하여 ⑫건강상의 이유로, ⑭군복무 중이어서 → 1로 코딩, 그 외 응답 0으로 코딩

본 연구에서는, 결측률이 20% 이상으로 나타난 ‘취업/자격증 관련 월평균 사교육 비용’, ‘여가시간 활동빈도(기타)’, ‘흡연량’ 등의 47개 변인을 제외하였으며 그 외 변인의 결측치의 경우 R(ver. 4.0.2)의 mice(ver 3.12.0) 패키지를 통해 대체하여 최종적으로 예측 모델에 투입한 115개 변인을 <표 2>에 정리하였다. 이를 살펴보면 개인의 여가 생활, 진로계획 및 직업의식, 개인 배경 및 일반적 특성이 조사된 개인 변인을 비롯하여 학교 생활, 사교육 및 방과후 활동, 현장체험 및 현장실습, 재학 중 근로 경험, 자격증과 관련된 변인 등을 포함하는 고등학교 생활 변인과 가정 생활 전반과 관련된 변인과 보호자와 관련된 변인이 조사된 가정 생활 변인을 활용하였다.

한편 <표 2>에서 이텔릭체로 표시된 문항의 경우 개별 문항들의 평균을 계산하여 분석에 활용하였다(예: 독서 선호도, 자신에 대한 평가, 진로 역량 등). 반면 개별 문항을 활용한 문항의 경우, 관련 문항의 개수를 함께 표기하였는데 구체적으로 ‘여가시간 활동 빈도(12)’, ‘진로체험 경험 여부(6)’와 같이 개별 문항이 서로 측정하고 있는 여가 활동 및 프로그램 유형이 구분되거나 ‘전문교과(전공) 수업태도(6)’와 같이 ‘전문교과 수업 내용을 충분히 이해할 수 있었다’, ‘전문교과의 비중이 지금보다 확대될 필요가 있었다’와 같이 개별 문항의 의미가 명확히 구분될 필요가 있는 경우가 이에 해당된다. 더불어 ‘미래직업 미결정 이유(더미)’ 등의 변인은 더미 처리하여 분석에 활용하였다.

<표 2> 진로이행과정 유형 예측 모델에 투입된 변인

대분류	중분류	변인
	여가생활	여가시간 평균, 여가시간 활동 빈도(12), 독서 경험 여부, 독서량, 독서 선호도, 동아리 활동 여부, 일주일 평균 동아리 활동시간, 해외여행 경험
개인	진로계획 및 직업의식	희망교육수준, 미래직업 결정 여부, 미래직업 미결정 이유(더미), 직업 보유 이유(더미), 직업선택 요소 1순위(더미), 직업생활의 성공조건(더미), 다문화수용성, 자신에 대한 평가
	배경 및 일반적 특성	성별, 건강 상태, 일주일 평균 운동시간, 수면 시간, 아침 식사, 고민거리, 흡연 여부, 흡연량, 음주 여부, 음주빈도, 행복도

대분류	중분류	변인
고등학교 생활	학교생활	고등학교 생활 만족도, <i>학교 시설 만족도</i> , 교사만족도(4), 고3 담임교사 만족도, <i>국어 교과교사 만족도</i> , <i>영어 교과교사 만족도</i> , <i>수학 교과교사 만족도</i> , 수업분위기(부정적), <i>내신등급(국어, 영어, 수학)</i> , 전문교과 내신등급, 전문교과(전공) 수업태도(6), <i>전문교과 교사 만족도</i> , 전공 동아리 경험여부, 기능반 경험여부, 진로와 직업수업 경험 여부, 창의적 체험활동 경험 여부, 진로심리검사 경험 여부, 진로상담 경험 여부, 진로동아리 경험 여부, 진로체험 경험 여부(6), 진로와 직업수업 만족도, 창의적 체험활동 만족도, 진로심리검사 만족도, 진로상담 만족도, 진로동아리 만족도, 진로체험 만족도(6), 취업상담 경험 여부, 채용기업정보제공 경험 여부, 구직활동지도 경험 여부, 취업박람회·채용설명회 경험 여부, 취업캠프 경험 여부, 졸업자멘토링 경험 여부, 취업처연계 경험 여부, 취업상담 만족도, 채용기업정보제공 만족도, 구직활동지도 만족도, 취업박람회·채용설명회 만족도, 취업캠프 만족도, 졸업생멘토링 만족도, 취업처연계 만족도, <i>진로역량</i> 방과후자율학습 운영 여부, 학교 내외 수상 여부
	사교육 및 방과후학교	취업 자격증 방과후학교 참여 여부, 취업 자격증 사교육 경험 여부, 취업 자격증 사교육 종류 개수, 사교육 이수 기간, <i>주중 평균 공부시간</i>
	현장체험 및 현장실습	현장체험 경험 여부, 현장체험 경험 횟수, 도제학교 프로그램 참여 여부
	재학 중 근로경험	유급 경험 여부, 유급노동 일한 총 기간, 유급노동 주당 근무 일수, 하루 근무 시간
	자격증	자격증 유무, 자격증 개수
	가정 생활 전반	현재 거주 지역(더미), 하숙 자취 기숙사 월평균 생활비용, 하숙 자취 기숙사 생활 비용 중 본인충당비율, 주택마련금액의 지원 정도, 가정생활 전반 만족도, <i>부모님과의 여가활동 빈도</i> , 부모님과의 대화 빈도(5)
	보호자	남성보호자의 최종학력, 남성보호자의 경제활동 상태, 남성보호자 월평균 소득, 남성보호자 직장 종사자 수, 여성보호자의 최종학력, 여성보호자의 경제활동 상태, 여성보호자 월평균 소득, 여성보호자 직장 종사자 수, 여성 남성 보호자 부동산 금융소득

3. 분석방법

본 연구에서는 직업계고 학생의 고등학교 졸업 이후 진로이행과정을 탐색하여 유형화하고, 분류된 진로이행과정 유형을 예측하는 학생 개인, 고등학교 생활 및 가정 변인을 탐색하기 위하여 <표 3>에 정리된 분석절차에 따르고자 한다.

<표 3> 분석 절차

	진로이행과정 유형 탐색	최적의 진로이행과정 유형 예측 모형 탐색	도출된 주요 예측 변인의 영향력 탐색 및 검증
분석 절차	시퀀스 분석 (Sequence analysis), k-means 군집 분석	자동화 머신러닝(AutoML) 알고리즘	SHAP (SHapley Additive exPlanation) 알고리즘, 로지스틱 회귀분석
분석 프로 그램	R 프로그램 (ver. 4.4.0) · TraMineR 패키지 (ver 2.2-10) · stats 패키지 (ver. 4.5.0)	Python 프로그램 (ver. 3.11) · PyCret 라이브러리 (ver. 2.3.10)	Python 프로그램 (ver. 3.11) · SHAP 라이브러리 (ver. 2.3.10) R 프로그램 (ver. 4.4.0) · stats 패키지 (ver. 4.5.0)

분석 절차를 구체적으로 설명하면, 먼저 시퀀스 분석을 수행하여 진로이행과정의 유형을 분석하였다. 사회과학 분야에서의 시퀀스 분석은 개인의 연속적인 사건을 유형화할 때 활용되는 분석기법으로(Abbott 1983) 시각화, 최적화 매칭(optimal matching) 등의 알고리즘을 통하여 각 개인의 상태가 전환되는 양상에 따라 유형화할 수 있다. 이에 본 연구에서는 직업계고 졸업자의 시점별 진로이행상태의 변화, 즉 5년간 이루어진 중단적인 진로이행과정을 하나의 시퀀스로 가정하고 최적화 매칭 알고리즘을 적용하여 진로이행상태의 전환 확률 행렬과 각 개인의 시퀀스 간 거리 행렬을 도출하였다. 이에 진로이행과정의 유형을 분류하기 위하여 개인의 시퀀스 간 거리 행렬을 기반으로 k-means 군집분석 알고리즘을 적용하였다. k-means

군집분석은 비계층적 군집분석 방법 중 하나로, k 개의 군집(cluster)과 개인 간 거리를 기반으로 각 군집에 개인을 할당시키며(MacQueen 1967; Everitt 1993) 연구자가 군집의 수(k)를 설정하기 때문에 군집 개수에 따른 WSS(Within cluster Sum of Square) 지수와 같은 통계적 검토와 더불어 군집 분류의 결과가 이론적 배경이나 실제 현장과 적합한 양상을 보이는지 타당성을 검토하는 과정을 거쳐 진로이행과정 유형을 분류하였다.

이후, 진로이행과정 유형을 결과 변인으로 설정하고 KEEP II에서 제공하는 다양한 학생 개인, 고등학교 생활, 가정 생활 변인을 예측 변인으로 설정하여 진로이행과정 유형 분류를 예측하는 머신러닝 모형을 구축하였다. 이때 최적의 예측 모델을 탐색하기 위하여 다양한 머신러닝 모델을 생성 후 성능을 비교하는 자동화 머신러닝(Automated machine learning, AutoML) 알고리즘을 적용하였다. 구체적으로 ① Decision Tree(Quinlan 1986), ② Extra Tree Classifier(Geurts et al. 2006), ③ Random Forest(Breiman 2001), ④ Light Gradient Boosting Machine(Ke et al. 2017), ⑤ Extreme Gradient Boosting(Chen & Guestrin 2016), ⑥ Gradient Boosting(Friedman 2001), ⑦ CatBoost(Prokhorenkova et al. 2018) 7개 모델을 비교·분석하였으며 각각의 모델 구축에 있어서 10-fold 교차 검증(cross-validation) 및 하이퍼파라미터(hyper parameter) 튜닝 과정을 거쳐 성능이 가장 우수한 것으로 나타난 모델을 최적의 예측 모델로 결정하였다.

최종적으로 결정된 예측 모델을 통해 모형에 투입된 학생 개인, 고등학교 생활 및 가정생활 변인 중 진로이행과정 유형 분류 예측에 높은 기여를 하는 주요 예측 변인을 도출하기 위하여 Tree SHAP 알고리즘을(Lundberg et al. 2020) 적용하여 SHAP 값(SHAP value)의 절대값 평균을 예측에 미치는 기여도를 의미하는 중요도 지수로 활용하였다. 이를 통해 머신러닝 기법을 활용한 연구에서 투입한 예측 변인 개수의 5%~20% 내외에 해당하는 변인을 주요 예측 변인으로 제시하고 있음을 참고하여(이선영 외 2020) 중요도 지수 상위 10개 변인을 주요 예측 변인으로 도출하였다. 이에 예측 변인에 대한 SHAP 지수의 분포를 나타내는 시각화 자료인 SHAP 의존성 도표(SHAP dependence plot)를 확인하여 주요 예측 변인과 진로이행과정 유형 간의 관계를 살펴보았다.

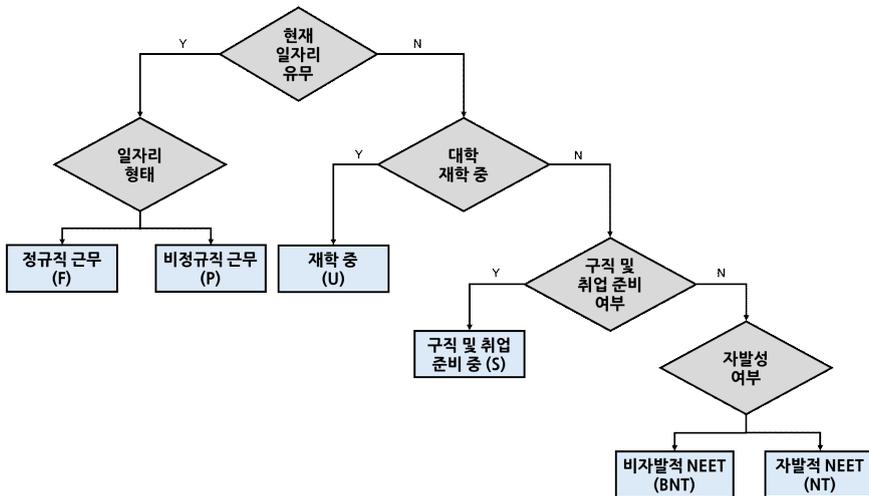
또한 머신러닝 모델을 통해 도출된 주요 예측 변인이 진로이행과정 유형에 미치는 영향력을 통계적으로 검증하기 위해 로지스틱 회귀분석을 추가로 수행하였다. 이는 머신러닝 모델 탐색만으로는 확인하기 어려운 종속 변인에 대한 예측 변인의

설명력과 해석의 한계를 보완하여 이를 통해 개인, 고등학교 생활, 가정생활과 관련된 주요 변인이 진로이행과정 유형에 미치는 영향에 대한 결과 해석의 신뢰성을 높이기 위해서이다.

IV. 연구결과

1. 직업계고 졸업자의 진로이행과정 유형 탐색

본 연구에서는 최적일치법을 통해 직업계고 졸업자의 진로이행과정의 유형을 탐색하였다. 이때 각 시점의 진로이행상태를 <그림 1>의 알고리즘에 따라 분류하였다.



<그림 1> 진로이행과정 분류 알고리즘 순서도

구체적으로, 졸업 후 현재 일자리(아르바이트 제외)를 가지고 있다고 응답한 학생에 대하여(‘현재 일자리 유무’ = ‘Y’) 정규직으로 근무 중인 경우(F)와 비정규직으로 근무 중인 경우(P)로 분류하였다. 현재 일자리가 없다고 응답한 학생들에 대하여(‘현재 일자리 유무’ = ‘N’) 대학에 재학 중인 경우(U)와 대학에 재학 중이지 않

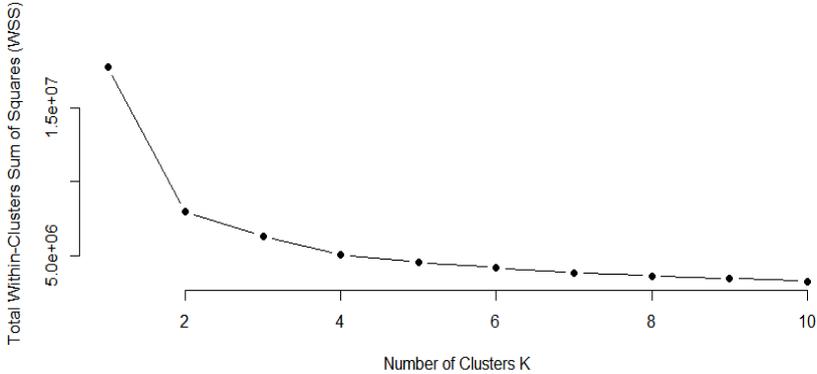
지만, 자격증 준비와 같은 취업 준비와 구직 중이라고 응답한 경우(S)로 분류하여 코딩하였다. 이때 구직 및 취업 준비 경험이 없는 학생들을 군입대 예정, 군복무 중, 건강상의 이유, 결혼·육아 가사 등의 이유에 해당되는 경우 비자발적 NEET 집단(BNT)으로 분류하고 그 외의 학생은 자발적 NEET 집단(NT)으로 분류하였다.

분류된 진로이행상태에 대한 각 시점별 빈도 분석 결과를 <표 4>에 제시하였다. 고등학교 졸업 직후 시점인 3차년도에 정규직에 근무 중이라고 응답한 학생들의 비율이 33.6%로 가장 높았으며, 대학에 재학 중이라고 응답한 비율이 28.4%로 두 번째로 높았다. 한편 가장 마지막 시점인 7차년도에 정규직에 근무 중이라고 응답한 비율이 52.9%로 나타났으며, 자발적 NEET 상태로 응답한 비율이 15.8%로 두 번째로 높게 나타난 것을 확인할 수 있었다. 또한 모든 시점에서 공통적으로 비정규직 근무(P) 비율이 가장 낮게 나타났다.

<표 4> 시점별 진로이행상태 분포 (N=1,919)

진로이행상태	3차		4차		5차		6차		7차	
	빈도	비율	빈도	비율	빈도	비율	빈도	비율	빈도	비율
정규직 근무(F)	644	33.6	612	31.9	791	41.2	922	48.0	1016	52.9
비정규직 근무(P)	81	4.2	57	3.0	90	4.7	117	6.1	118	6.1
대학 재학 중(U)	545	28.4	353	18.4	229	11.9	279	14.5	227	11.8
구직 및 취업 준비 중(S)	209	10.9	150	7.8	125	6.5	143	7.5	145	7.6
비자발적 NEET(BNT)	166	8.7	545	28.4	440	22.9	174	9.1	110	5.7
자발적 NEET(NT)	274	14.3	202	10.5	244	12.7	284	14.8	303	15.8

앞서 서술한 알고리즘에 따라 분류된 진로이행과정의 유형을 탐색하기 위하여 최적화 매칭을 통해 도출된 진로이행상태의 전환 확률 행렬에 대해 k-means 군집 분석을 실시하여 유형(clusters)을 탐색하였다. 본 연구에서는 유형의 개수를 결정함에 있어서 통계적 검토와 유형 분류 결과의 타당성을 점검하였다. 우선 통계적 검토를 위하여 유형 개수에 따른 WSS 변화를 살펴보았다. WSS는 군집 즉, 유형의 개수가 증가할수록 감소하는 경향이 있으므로 감소율이 급격하게 감소하는 유형의 개수를 중심으로 살펴보면 4개일 때를 기점으로 그래프가 완만해지는 것을 확인할 수 있다.



<그림 2> 진로이행과정 유형 고정 군집 개수에 따른 WSS 변화

이에 따라 4개로 분류된 진로이행과정 유형의 특성을 검토한 결과, 각각의 유형이 진로이행과정의 뚜렷한 특성으로 구분되며 의미 있는 차이를 반영하는 것으로 판단되어 최종적인 유형의 개수를 4개로 설정하였다. 각 유형별 특성에 따라 각 유형을 명명한 결과를 <표 5>와 <그림 3>에 제시하였다.

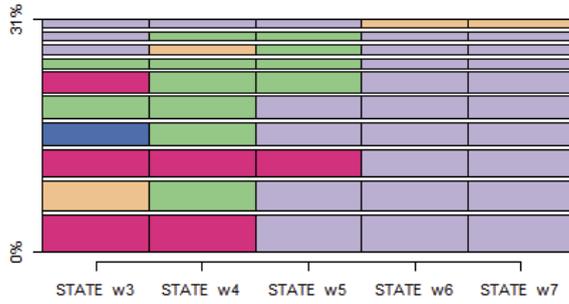
<표 5>를 살펴보면, 우선 첫 번째 집단의 경우 졸업 직후에는 일자리를 갖지 못하였으나 대학 재학 후 정규직을 일자리를 갖게 되거나 비자발적·자발적 NEET 상태, 혹은 구직 및 취업 준비 과정을 거쳐 2년 이상 정규직에 근무하면서 정착하는 패턴을 보였다. 이에 해당 집단을 ‘점진적 정규직 정착형’으로 명명하였으며, 해당 집단은 가장 많은 학생들이(30.0%) 할당된 것으로 나타났다. 두 번째로 많은 학생들이 할당된 집단은 졸업 직후부터 정규직에 정착하는 패턴을 보여 ‘초기 정규직 정착형’으로 명명하였다. 해당 집단에는 26.3%의 학생들이 할당되었다. 다음으로, 졸업 후 정규직 및 비정규직 일자리로 취업하지 않고 대학에 진학하여 재학 상태인 유형이 도출되어 해당 집단을 ‘대학 진학형’으로 명명하였으며 22.1%의 학생들이 포함되어 있었다. 해당 유형의 경우 대다수의 사례가 고등학교 졸업 직후부터 대학에 진학하여 대학에 재학 중인 상태이거나, 3년 이상 대학에 재학한 뒤 정규직에 막 취업하였거나, NEET 상태인 것을 알 수 있다. 마지막으로 졸업 이후에 일자리를 갖거나 대학에 진학하지 않고 장기간 자발적·비자발적 NEET 상태에 머무르는

패턴을 보인 유형인 ‘장기 NEET형’ 집단이 도출되었다. 해당 집단에는 21.7% 학생들이 할당되어, 가장 낮은 빈도를 보였다.

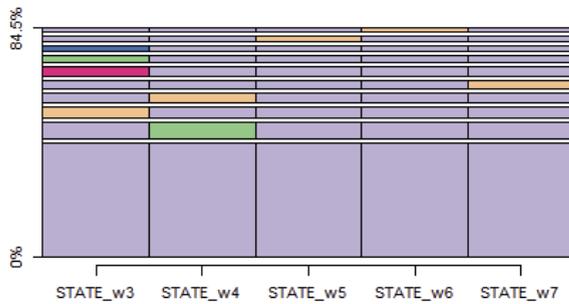
<표 5> 진로이행과정 유형별 특성 및 빈도

유형	특성	진로이행과정 예시 ¹⁾	빈도 (%)
점진적 정규직 정착형	졸업 후 대학 재학(U), 구직 및 취업 준비 과정(S), 자발적·비자발적 NEET 상태(NT, BNT) 등의 과정을 거쳐 정규직(F)에 2년 이상 정착하는 유형	U-U-F-F-F, NT-BNT-F-F-F, U-U-U-F-F S-BNT-F-F-F, BNT-BNT-F-F-F	575 (30.0%)
초기 정규직 정착형	졸업 직후 정규직(F)에 정착하는 유형	F-F-F-F-F, F-BNT-F-F-F, NT-F-F-F-F, F-NT-F-F-F, F-F-F-F-NT	504 (26.3%)
대학 진학형	졸업 후 대학교에 진학하여 재학 중이거나, 대학교에 3년 이상 재학한 뒤 졸업 후 정규직에 막 취업 또는 NEET상태인 유형	U-BNT-BNT-U-U, U-U-U-U-NT, U-U-U-U-U, U-BNT-BNT-U-F, U-U-U-U-F	424 (22.1%)
장기 NEET형	일자리를 갖거나, 대학에 진학하지 않고, 비자발적·자발적 NEET(BNT, NT) 상태에 4년 이상 머무르는 유형	BNT-BNT-BNT-NT-NT, BNT-BNT-NT-NT-NT, NT-BNT-BNT-NT-NT, NT-BNT-BNT-BNT-NT NT-NT-NT-NT-NT,	416 (21.7%)

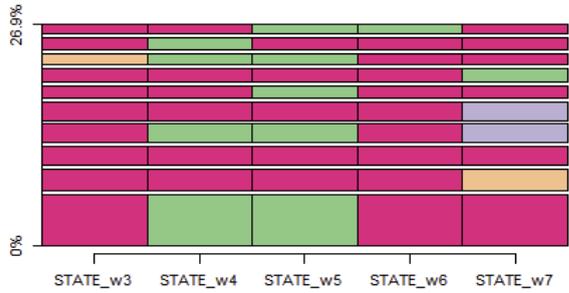
주: 1) 각 유형별로 가장 높은 빈도로 나타난 진로이행과정을 다음과 같이 약어로 표기하였음:
정규직 근무(F), 비정규직 근무(P), 대학 재학 중(U), 구직 및 취업 준비 중(S), 비자발적
NEET(BNT), 자발적 NEET(NT).



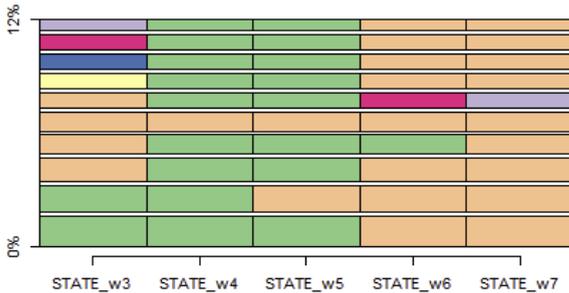
점진적 정규직 정착형(n=575)



초기 정규직 정착형(n=504)



대학 진학형(n=424)



장기 NEET형(n=416)



<그림 3> 유형별 진로이행과정 시각화 결과

2. 진로이행과정 유형을 예측하는 고등학교 및 가정 생활, 개인 변인 탐색

앞서 도출된 진로이행과정 패턴에 따른 네 가지 유형(점진적 정규직 정착형, 초기 정규직 정착형, 대학 진학형, 장기 NEET형)을 예측하는 개인, 고등학교 생활, 가정 생활 변인을 탐색하였다. 이때 직업계고등학교의 경우 일반계고와 달리 중등 직업교육단계에서 산업 수요에 맞춘 인력 양성을 목적으로 한다는 점에 주목하여, 고등학교 졸업 직후 정규직에 취업한 후 정규직에 장기간 정착하는 유형인 ‘초기 정규직 정착형’을 참조집단(0)으로 설정하고 세 가지 유형을 각각 비교집단(1)으로 설정하였다. 이에 따라 세 가지 예측 모형(①초기 정규직 정착형-장기 NEET형, ②초기 정규직 정착형-대학 진학형, ③초기 정규직 정착형-점진적 정규직 정착형)에 대하여 자동화 머신러닝 알고리즘을 적용하여 예측 모델을 비교하여 가장 높은 성능을 보인 최종 예측 모델을 활용하여 주요 예측 변인을 도출하였다.

1) 진로이행과정 예측 모델 탐색

자동화 머신러닝 알고리즘을 통해 6개의 머신러닝 분류 모델의 성능 평가 결과를 아래의 <표 6>에 제시하였다. 우선, ‘초기 정규직 정착형’과 ‘장기 NEET형’ 간 분류에 대하여 모델의 성능을 비교한 결과를 살펴보면 Catboost의 정확도, 정밀도, 민감도가 0.661, 0.661, 0.662 가장 높게 나타나고 그 외의 평가 지표 또한 가장 우수한 것으로 나타났다. ‘초기 정규직 정착형’과 ‘대학 진학형’ 간의 분류를 수행한 예측 모델 또한 Catboost의 성능이 정확도, 정밀도, 민감도가 각각 0.724, 0.728, 0.724으로 가장 우수한 것을 확인할 수 있었다. 한편, ‘초기 정규직 정착형’과 ‘점진적 정규직 정착형’ 간의 분류를 수행한 예측 모델을 살펴보면, 정확도, 정밀도, 민감도가 각각 0.592, 0.618, 0.592로 나타난 ET(Extra Tree classifier) 모델의 성능이 그 외의 모델과 비교했을 때 상대적으로 가장 높은 것을 확인할 수 있었다.

이렇듯 비교집단 설정에 따라 최종적으로 서로 다른 머신러닝 모델이 선택되었으며 비교집단을 ‘점진적 정규직 정착형’으로 설정했을 때 모델의 성능이 전반적으로 낮게 나타나는 등 자동화 머신러닝 알고리즘 분석결과가 유형에 따라 상이하게 나타나는 것을 확인할 수 있었다. 이러한 분석결과는 본 연구에서 기존 선행연구에서 모형 설정에 따른 제한된 활용 변인의 한계를 극복하고자 설명 변인의 선별(variable selection) 절차를 거치지 않고 패널자료에서 제공하는 다양한 변인들을 모형에 모두 투입한 결과 일부 모형에서 예측력이 낮게 나타난 것으로 유추된다.

<표 6> 자동화 머신러닝 알고리즘을 통한 머신러닝 모델별 성능 비교

모델 ¹⁾		정확도 (Accuracy)	AUC	민감도 (Recall)	정밀도 (Precision)	F1 점수	Kappa	MCC
초기 정규직 정착형 vs 장기 NEET형	DT	0.574	0.571	0.574	0.575	0.572	0.141	0.142
	ET	0.633	0.709	0.633	0.633	0.625	0.245	0.252
	RF	0.644	0.699	0.644	0.644	0.636	0.267	0.274
	LightGBM	0.661	0.706	0.661	0.664	0.657	0.310	0.315
	XGBoost	0.638	0.692	0.638	0.641	0.634	0.264	0.270
	GBC	0.646	0.678	0.646	0.647	0.642	0.280	0.284
	Catboost	0.661	0.717	0.661	0.662	0.659	0.312	0.314
초기 정규직 정착형 vs 대학 진학형	DT	0.587	0.585	0.587	0.590	0.585	0.168	0.171
	ET	0.709	0.770	0.709	0.714	0.703	0.403	0.413
	RF	0.700	0.767	0.700	0.703	0.693	0.385	0.393
	LightGBM	0.681	0.739	0.681	0.682	0.678	0.351	0.355
	XGBoost	0.682	0.752	0.682	0.685	0.678	0.353	0.359
	GBC	0.670	0.729	0.670	0.674	0.667	0.332	0.337
	Catboost	0.724	0.775	0.724	0.728	0.721	0.439	0.445
초기 정규직 정착형 vs 점진적 정규직 정착형	DT	0.513	0.512	0.513	0.514	0.512	0.024	0.024
	ET	0.592	0.618	0.592	0.591	0.588	0.175	0.177
	RF	0.582	0.608	0.582	0.581	0.577	0.153	0.156
	LightGBM	0.542	0.571	0.542	0.541	0.538	0.075	0.077
	XGBoost	0.554	0.587	0.554	0.552	0.551	0.100	0.100
	GBC	0.567	0.575	0.567	0.566	0.564	0.125	0.127
	Catboost	0.574	0.596	0.574	0.572	0.570	0.136	0.138

주: DT: Decision Tree Classifier,
ET: Extra Trees Classifier,
RF: Random Forest Classifier,
LightGBM: Light Gradient Boosting Machine,
XGBoost: Extreme Gradient Boosting,
GBC: Gradient Boosting Classifier,
Catboost: CatBoost Classifier.

따라서 본 연구에서는 주요 예측 변인과 진로이행과정 유형 간의 관계를 해석함에 있어서, 최종 머신러닝 모델에서 SHAP 알고리즘 분석을 통해 예측 기여도가 높은 것으로 나타난 상위 10개의 변인을 도출하고 시각화 도표를 확인함과 더불어 로지스틱 회귀분석을 실시함으로써 주요 예측 변인과 진로이행과정 유형 간 관계를 통계적으로 검증하는 절차를 거쳤다. 이어지는 분석결과를 서술하면 다음과 같다.

2) 진로이행과정 유형을 예측하는 주요 변인 탐색

앞서 탐색한 최적의 예측 모델을 통해 진로이행과정 유형을 예측하는 주요 변인을 탐색하기 위해 SHAP 알고리즘을 통해 중요도 지수를 산출하여 중요도 지수가 높게 나타난 상위 10위 변인을 주요 예측 변인으로 도출하였다. 이에 세 가지 예측 모형(① 초기 정규직 정착형-장기 NEET형, ② 초기 정규직 정착형-대학 진학형, ③ 초기 정규직 정착형-점진적 정규직 정착형)에 대하여 각각 주요 예측 변인과 예측 변인 각각의 중요도 지수 및 중요도 지수 순위를 <표 7>에 제시하였다.

그 결과를 살펴보면, 우선 ‘초기 정규직 정착형’과 ‘장기 NEET형’ 집단 분류에 높은 기여를 하는 변인은 중요도 지수가 높은 순으로 성별, 자격증 개수, 부정적인 수업분위기, 전문교과 내신등급, 음주 여부, 취업 자격증 방과후학교 참여 여부, 행복도, 국어·영어·수학 내신등급, 진로 역량, 건강 상태인 것으로 나타났다. ‘정규직 정착형’과 ‘대학 진학형’ 집단 분류에 높은 기여를 하는 변인은 희망교육수준, 부정적인 수업 분위기, 자격증 개수, 취업상담 경험 여부, 취업 자격증 방과후학교 참여 여부, 취업처연계 경험 여부, 여가시간 활동 빈도(독서), 국어·영어·수학 내신등급, 여성 보호자의 경제활동 상태, 전문교과 수업태도(수업 내용 이해) 순으로 중요도 지수가 높은 것으로 나타났다. 마지막으로 ‘초기 정규직 정착형’과 ‘점진적 정규직 정착형’의 분류에 높은 기여를 하는 것으로 나타난 변인을 살펴보면 취업 자격증 방과후 학교 참여 여부, 채용기업정보제공 경험 여부, 취업처연계 경험 여부, 학교 내·외 수상 여부, 방과후자율학습 참여 여부, 직업선택 요소(0: 경제적 안정, 1: 삶의 질), 취업상담 경험 여부, 성별, 진로체험학과체험 경험 여부, 미래직업 결정 여부 순으로 중요도 지수가 높게 나타났다.

<표 7> 주요 예측 변인별 중요도 지수(순위)

영역	변인	초기 정규직 정착형(참조집단)			
		① 장기 NEET형	② 대학 진학형	③ 점진적 정규직 정착형	
개인	여가생활	독서 활동 빈도	0.0836(7)		
	진로계획 및 직업의식	진로 역량	0.1550(9)		
		직업선택 요소 (0: 경제적 안정, 1: 삶의 질)	0.0103(6)		
		미래직업 결정 여부	0.0079(10)		
		희망교육수준	0.6409(1)		
	배경 및 일반적 특성	성별	0.6980(1)	0.0101(8)	
		음주 여부	0.2277(5)		
		행복도	0.2009(7)		
		건강상태	0.1491(10)		
	고 등 학 교	학교생활	수업 분위기(부정적)	0.3996(3)	0.2129(2)
전문교과 수업태도 (수업 내용 이해)			0.0754(10)		
국어·영어·수학 내신등급			0.1786(8)	0.0779(8)	
전문교과 내신등급			0.2514(4)		
학교 내·외 수상 여부			0.0113(4)		
취업상담 경험 여부			0.1704(4)		0.0102(7)
채용기업정보제공 경험 여부			0.0152(2)		
취업처연계 경험 여부			0.1412(6)		0.0115(3)
진로체험학과체험 경험 여부			0.0089(9)		
사교육 및 방과후학교			방과후자율학습 참여 여부	0.0104(5)	
	취업자격증 방과후학교 참여 여부	0.2080(6)	0.1503(5)	0.0184(1)	
자격증	자격증 개수	0.4574(2)	0.2097(3)		
가 정	보호자	여성 보호자의 경제활동 여부	0.0773(9)		

3. 주요 예측 변인과 진로이행과정 유형 간 관계

1) 초기 정규직 정착형과 장기 NEET형

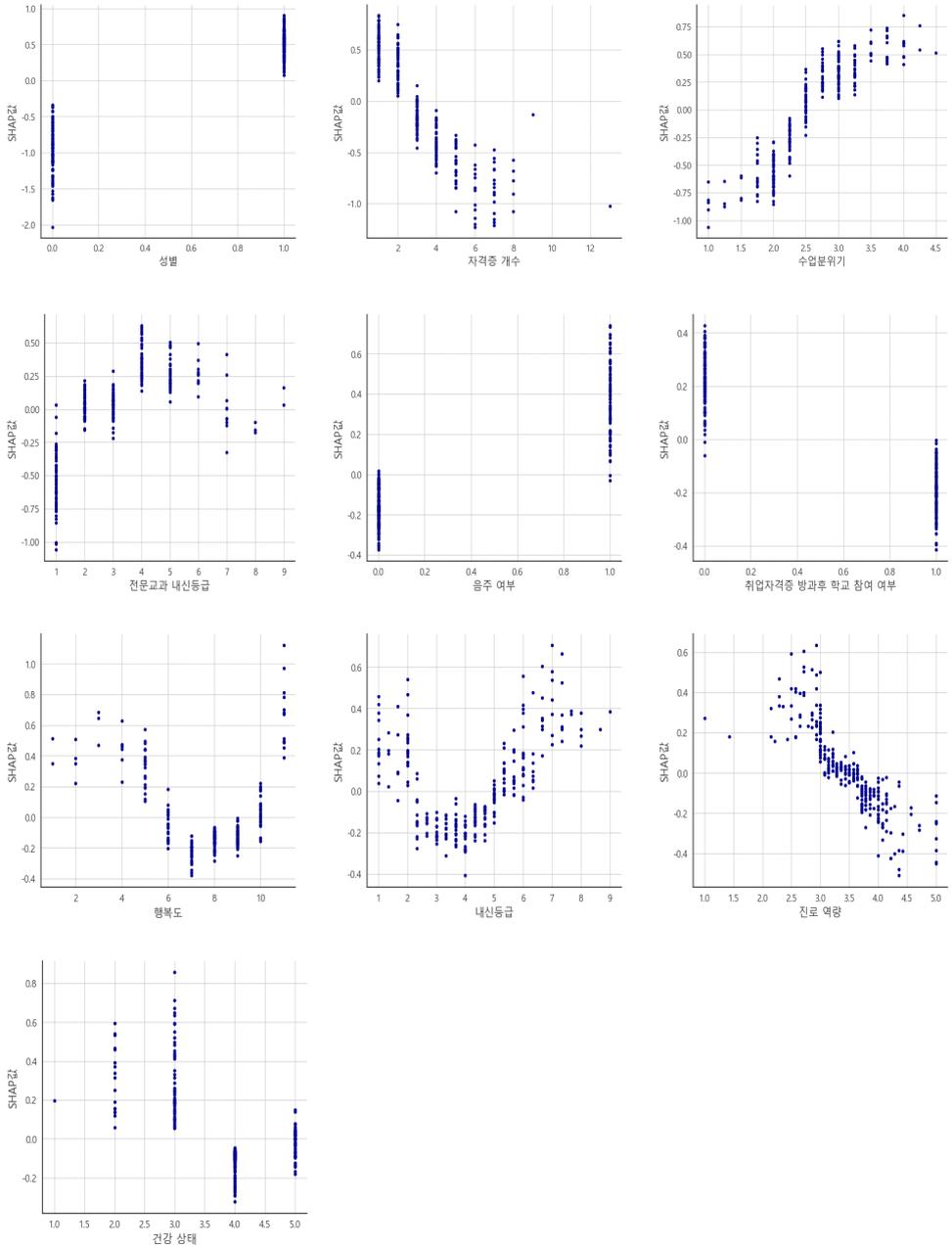
본 연구에서는 도출된 주요 예측 변인과 진로이행과정 유형 간의 관계를 설명하고 주요 예측 변인이 진로이행과정 유형 분류에 미치는 영향력을 검증하기 위해 SHAP 의존성 도표와 로지스틱 회귀분석을 실시하였다.

SHAP 의존성 도표의 경우 예측 변인의 응답 수준에 따른 SHAP 지수의 분포를 시각화한 도표로, Y축의 SHAP 지수가 0보다 클 때(+), 예측 변인의 응답 수준(X축)이 높을수록 특정 집단(1)에 속할 가능성이 높은 것을 의미한다. 이에 초기 정규직 정착형(0)과 장기 NEET형(1)의 주요 예측 변인의 기술통계 결과를 <표 8>에 정리하였으며 주요 예측 변인의 분포를 확인한 결과, 모든 변인의 왜도와 첨도가 Kline(2023)의 기준(왜도 |3| 미만, 첨도 |10| 미만)을 충족하여 정규성 가정을 위배하지 않는 것을 확인하였다.

<표 8> 초기 정규직 정착형(0)과 장기 NEET형(1) 주요 예측 변인 기술통계 (N=920)

변인	빈도 또는 평균 (표준편차)		변인	빈도 또는 평균 (표준편차)	
성별	여자(0)	330	취업자격증 방과후학교 참여 여부	없다(0)	422
	남자(1)	590		있다(1)	498
자격증 개수	2.97(1.95)		행복도	8.02(1.92)	
부정적인 수업 분위기	2.60(0.69)		국어·영어·수학 내신등급	4.30(1.75)	
전문교과 내신등급	3.22(1.79)		진로 역량	3.51(0.57)	
음주 여부	없다(0)	615	건강 상태	3.95(0.80)	
	있다(1)	305			

<표 8>의 주요 예측 변인에 대한 SHAP 의존성 도표를 중요도 지수가 높은 순서로 <그림 4>에, 로지스틱 회귀분석 결과를 <표 9>에 제시하였다.



<그림 4> 초기 정규직 정착형(0)-장기 NEET(1) 예측 모형에 대한 SHAP 의존성 도표

<표 9> 초기 정규직 정착형(0)과 장기 NEET형(1)에 대한 로지스틱 회귀분석 결과 (N=920)

변인	B	S.E.	exp(B)	p-value
성별	0.952***	0.162	2.590	0.000
자격증 개수	-0.148***	0.044	0.863	0.001
부정적인 수업 분위기	0.478***	0.112	1.613	0.000
전문교과 내신등급	0.031	0.055	1.031	0.572
음주 여부	0.212	0.157	1.236	0.176
취업자격증 방과후학교 참여 여부	-0.434**	0.148	0.648	0.003
행복도	-0.070	0.041	0.932	0.087
국어·영어·수학 내신등급	0.099	0.055	1.104	0.073
진로 역량	-0.436**	0.143	0.647	0.002
건강상태	-0.051	0.097	0.950	0.601

주: * $p < .05$, ** $p < .01$, *** $p < .001$.

분석결과를 살펴보면 다음과 같다. 성별, 자격증 개수, 부정적인 수업 분위기, 취업자격증 방과후학교 참여 여부, 진로역량이 초기 정규직 정착형(0)과 장기 NEET형(1) 분류에 유의한 영향을 미치는 것으로 나타났으며, SHAP 의존성 도표의 결과와 회귀계수의 방향성이 일치하는 것을 확인할 수 있었다. 여성보다 남성이 장기 NEET형(1)으로 분류될 확률이 높았으며, 자격증 개수가 적고, 부정적인 수업 분위기를 경험하고, 취업 자격증 방과후학교 참여 경험이 없고, 진로역량이 낮을수록 장기 NEET형(1)에 속할 확률이 높은 것으로 나타났다.

반면 전문교과 내신등급, 음주 여부, 행복도와 국어·영어·수학 내신등급, 건강상태의 경우 회귀분석 결과 통계적으로 유의하지 않았으나, SHAP 의존성 도표를 통해 유형과의 관계를 살펴볼 수 있었다. 우선 전문교과 내신등급의 경우 1등급에 해당될수록, 음주경험이 없다(0)고 응답한 학생일수록 SHAP 지수가 음수로 나타나 초기 정규직 정착형(0)에 속할 확률이 높은 것으로 나타났다. 반면 행복도, 국어·영어·수학 내신등급, 건강상태의 경우 예측 변인의 응답 수준에 따라 비선형적인 양상을 보였으며 세 변인 모두 예측 변인의 수준이 매우 높거나 낮을 경우 초기 정규직 정착형(0)이 아닌 장기 NEET형(1)에 속할 확률이 높은 것을 확인할 수 있었다.

2) 초기 정규직 정착형과 대학 진학형

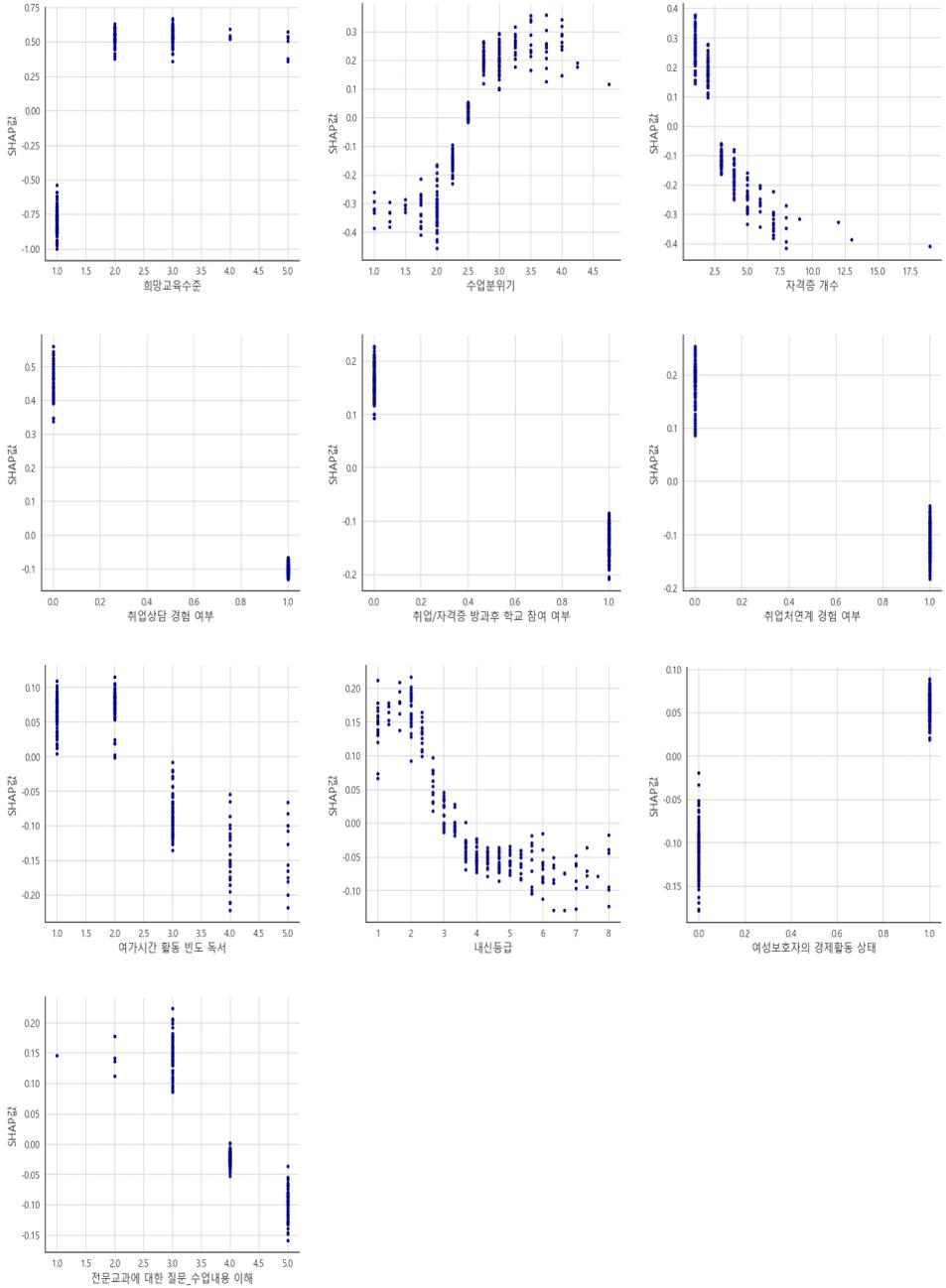
초기 정규직 정착형(0)과 대학 진학형(1)의 주요 예측 변인의 기술통계를 <표 10>에 정리하였으며, 모든 변인의 왜도와 첨도가 Kline(2023)의 기준을 충족하였다. 해당 변인에 대한 SHAP 의존성 도표를 <그림 5>에, 로지스틱 회귀분석 결과를 <표 11>에 제시하였다.

<표 10> 초기 정규직 정착형(0)과 대학 진학형(1) 주요 예측 변인 기술통계 (N=928)

변인	빈도 또는 평균(표준편차)		변인	빈도 또는 평균(표준편차)	
희망교육수준	2.00(0.96)		취업처연계 경험 여부	없다(0)	368
				있다(1)	560
부정적인 수업 분위기	2.57(0.69)		독서활동빈도	2.26(1.11)	
자격증 개수	3.05(2.04)		국어·영어·수학 내신등급	3.93(1.68)	
취업상담 경험여부	없다(0)	173	여성 보호자의 경제활동 여부	없다(0)	337
	있다(1)	755		있다(1)	591
취업자격증 방과후학교 참여 여부	없다(0)	445	전문교과 수업 내용 이해도	3.96(0.76)	
	있다(1)	483			

그 결과를 살펴보면 희망교육수준, 부정적인 수업 분위기, 자격증 개수, 취업상담 경험 여부, 취업자격증 방과후학교 참여 여부, 취업처연계 경험 여부, 여성 보호자의 경제활동 여부, 전문교과 수업 내용 이해도가 집단 분류에 유의한 영향을 미치는 것으로 나타났다. 해당 변인 또한 SHAP 의존성 도표의 결과와 회귀계수의 해석이 일치하는 것을 확인할 수 있었다. 구체적으로, 희망교육수준이 높고, 부정적인 수업 분위기를 경험하며, 자격증 개수가 적고 취업상담과 취업자격증 방과후학교, 취업처연계 경험이 없을수록 대학 진학형(1)에 속할 확률이 높은 것으로 나타났다. 또한 여성 보호자가 경제활동에 참여하고 있으며 전문교과 수업이해도가 낮을수록 대학 진학형(1)에 속할 확률이 높은 것으로 나타났다.

반면 독서 활동 빈도와 국어·영어·수학 내신등급의 경우 회귀분석 결과 통계적으로 유의하지 않았으나 SHAP 의존성 도표를 통해 유형과의 관계를 살펴볼 수 있는데, 독서 활동 빈도가 낮고, 국어·영어·수학 내신등급이 우수할수록(1등급) 대학 진학형(1)에 속할 확률이 높은 것으로 나타났다.



<그림 5> 초기 정규직 정착형(0)-대학 진학형(1) 예측 모형에 대한 SHAP 의존성 도표

<표 11> 초기 정규직 정착형(0)과 대학 진학형(1)에 대한 로지스틱 회귀분석 결과
(N=928)

변인	B	S.E.	exp(B)	p-value
희망교육수준	0.510***	0.082	1.666	0.000
부정적인 수업 분위기	0.320**	0.111	1.377	0.004
자격증 개수	-0.151***	0.041	0.860	0.000
취업상담 경험 여부	-0.901***	0.215	0.406	0.000
취업자격증 방과후학교 참여 여부	-0.455**	0.153	0.635	0.003
취업처연계 경험 여부	-0.662***	0.161	0.516	0.000
독서 활동 빈도	-0.130	0.069	0.878	0.059
국어·영어·수학 내신등급	-0.087	0.048	0.917	0.068
여성 보호자의 경제활동 여부	0.403*	0.158	1.497	0.011
전문교과 수업 내용 이해도	-0.259*	0.106	0.772	0.015

주: * $p < .05$, ** $p < .01$, *** $p < .001$.

3) 초기 정규직 정착형과 점진적 정규직 정착형

초기 정규직 정착형(0)과 점진적 정규직 정착형(1)의 주요 예측 변인의 기술통계 결과를 <표 12>에 정리하였으며 해당 변인에 대한 SHAP 의존성 도표를 중요도 지수가 높은 순서로 <그림 6>에, 로지스틱 회귀분석 결과를 <표 13>에 제시하였다.

결과를 살펴보면, 직업선택요소, 취업상담 경험 여부, 성별이 유형 분류에 유의한 영향을 미치는 것으로 나타났으며 SHAP 의존성 도표의 결과와 회귀계수의 방향성이 일치하였다. 이를 살펴보면, 직업선택 시 소득 및 고용보장보다 여가시간 확보와 쾌적한 환경이 더 중요하다고 응답하고, 취업상담 경험이 없고, 남성일수록 점진적 정규직 정착형(1)으로 분류될 확률이 더 높은 것으로 나타났다.

반면 회귀분석 결과, 통계적으로 유의하지 않게 나타난 변인의 SHAP 의존성 도표를 살펴보면, 학교에서 제공하는 취업 지도 경험에 해당되는 취업 자격증 방과후 학교, 채용 기업정보 제공, 취업처연계, 취업상담 경험이 없을수록 초기 정규직 정착형보다 점진적 정규직 정착형(1)에 속할 확률이 높은 것으로 나타났다. 이외에도 학교 내·외 수상 경험을 비롯하여 방과후자율학습, 진로체험학과 경험이 없을수록 점진적 정규직 정착형(1)에 속할 확률이 높았다.

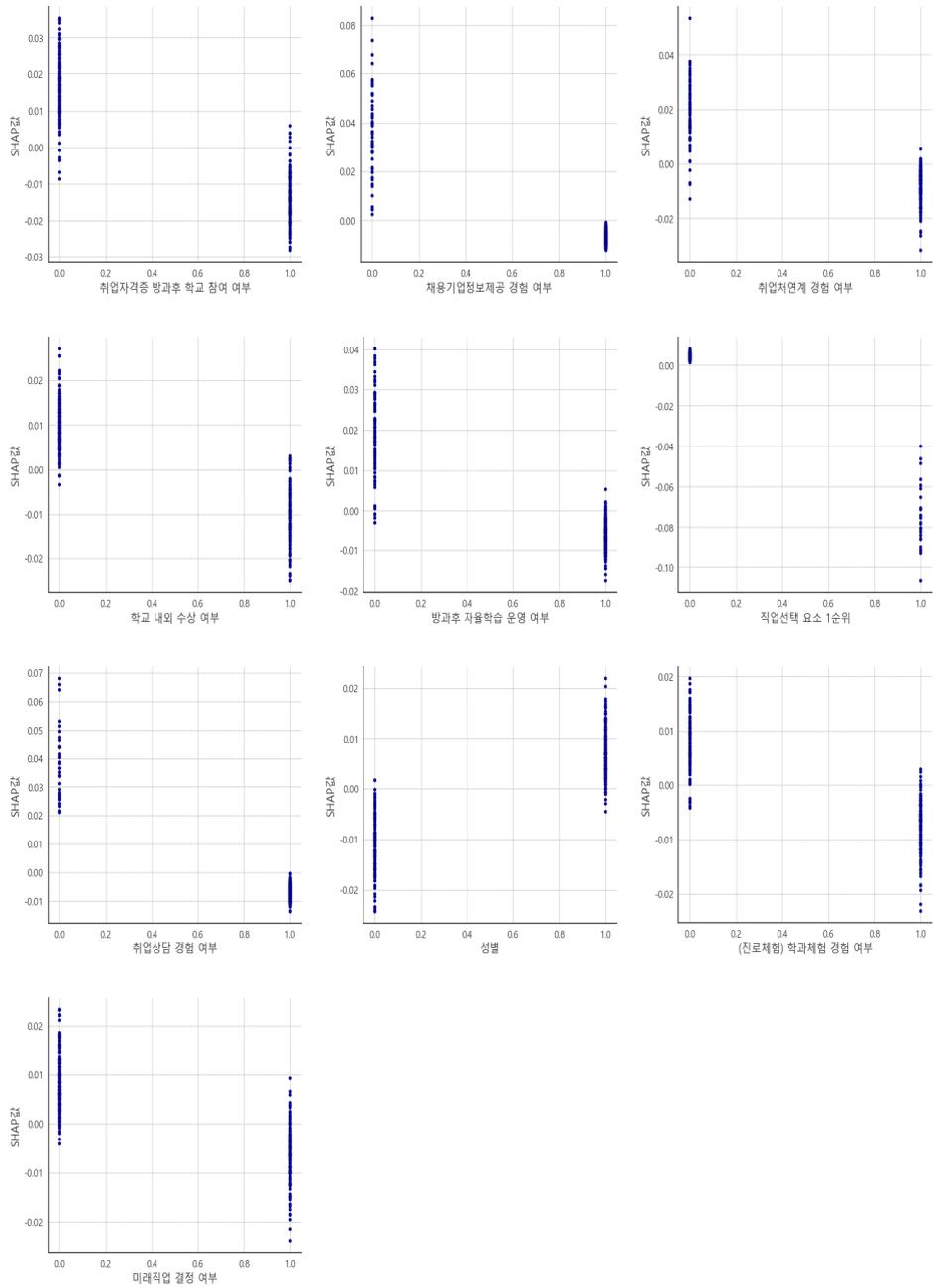
<표 12> 초기 정규직 정착형(0)과 점진적 정규직 정착형(1) 주요 예측 변인 기술통계 (N=1,079)

변인	빈도		변인	빈도	
취업자격증 방과후학교 참여 여부	없다(0)	472	직업선택요소	소득 및 고용보장(0)	1012
	있다(1)	607		여가시간확보, 쾌적한 환경(1)	67
채용기업정보제공 경험 여부	없다(0)	161	취업상담 경험 여부	없다(0)	141
	있다(1)	918		있다(1)	938
취업처연계 경험 여부	없다(0)	353	성별	여자(0)	426
	있다(1)	726		남자(1)	653
학교 내·외수상 여부	없다(0)	567	진로체험학과 경험 여부	없다(0)	520
	있다(1)	512		있다(1)	559
방과후자율학습 운영 여부	없다(0)	313	미래직업 결정 여부	없다(0)	516
	있다(1)	766		있다(1)	563

<표 13> 초기 정규직 정착형(0)과 점진적 정규직 정착형(1)에 대한 로지스틱 회귀분석 결과 (N=1,079)

변인	B	S.E.	exp(B)	p-value
취업자격증 방과후학교 참여 여부	-0.179	0.134	0.836	0.180
채용기업정보제공 경험 여부	-0.169	0.212	0.845	0.426
취업처연계 경험 여부	-0.170	0.151	0.844	0.261
학교내외수상 여부	-0.130	0.132	0.878	0.325
방과후자율학습 운영 여부	-0.225	0.144	0.798	0.119
직업선택요소 (0: 소득 및 고용보장, 1: 여가시간 확보와 쾌적한 환경)	-0.716**	0.270	0.489	0.008
취업상담 경험 여부	-0.448*	0.216	0.639	0.038
성별	0.416**	0.129	1.515	0.001
진로체험학과 경험 여부	-0.105	0.131	0.900	0.421
미래직업 결정 여부	-0.116	0.129	0.890	0.369

주: * $p < .05$, ** $p < .01$, *** $p < .001$.



<그림 6> 초기 정규직 정착형(0)-점진적 정규직 정착형(1) 예측 모형에 대한 SHAP 의존성 도표

V. 논의 및 결론

본 연구에서는 직업계고 졸업자의 진로이행과정 유형을 탐색하고 이를 예측하는 머신러닝 기반의 예측 모델에 대해 설명 가능한 인공지능의 일종인 SHAP 알고리즘과 로지스틱 회귀분석을 적용하여 직업계고 졸업자의 진로이행과정 유형과 예측 변인 간 관계를 살펴보았다. 주요 결과를 기반으로 한 시사점을 논의하면 다음과 같다.

본 연구는 직업계고 졸업자의 진로이행상태를 정규직·비정규직 근무, 대학 재학, 구직 및 취업 준비 중, 비자발적·자발적 NEET 상태를 포괄하여 구분하고 졸업 후 5년간 진로이행과정 시퀀스에 따른 유형을 도출함으로써 기존의 연구에서는 도출되지 않았던 세부적인 진로이행경로를 유형화하였다는 데 그 의의가 있다. 그 결과를 살펴보면 졸업 직후에 정규직에 취업하여 3~4년 이상 정규직에 정착하는 ‘초기 정규직 정착형’, 대학 진학이나 구직 및 취업 준비 과정, 자발적·비자발적 NEET 상태를 거쳐 정규직에 취업하여 2년 이상 정규직에 정착하는 ‘점진적 정규직 정착형’ 유형이 도출되었는데 ‘점진적 정규직 정착형’ 유형에 해당되는 비율이 30.0%로 나타나 가장 많은 졸업자들이 포함된 것을 확인할 수 있었다. 더불어 앞선 유형과 달리 직업계고 졸업 직후 대학에 진학하여 3년 이상 재학 중이거나, 대학 재학을 마치고 정규직에 막 취업한 ‘대학 진학형’, 그리고 대학 재학, 구직 활동, 취업 준비 중 어느 것도 하지 않는 자발적·비자발적 NEET 상태에 머무르는 ‘장기 NEET형’ 유형이 도출되었다. 이러한 결과는 최근 직업계고 졸업자의 대학 진학을 증가와 함께, 진학도, 취업도 하지 않는 졸업자의 비율이 증가하는 현실(한국교육개발원·교육부 2024)과 맥락을 같이하며 기존의 선행연구에서는 도출되지 않았던 본 연구의 차별적인 결과에 해당된다. 이러한 분석결과는 고졸 취업 수요의 감소와 더불어, 한국 사회 전반에 걸친 대학 진학을 증가로 인해 초래된 학력 인플레이션 등 노동시장의 구조적 문제(나상현 2024; 박고은 2023)에 대한 시사점을 제공한다. 이와 더불어 4년 이상 장기 NEET 상태를 유지하는 ‘장기 NEET형’ 유형에 주목할 필요가 있는데, NEET는 경제활동과 구직 활동에 참여하지 않아 사회에 경제적 부담과 잠재 성장력을 저하시키는 부정적인 영향을 미치는 사회 문제이다(OECD 2019). 더불어

장기간 NEET 상태에 머물렀을 때, 개인 삶의 질에 있어서도 부정적인 영향을 미치는데, 빈곤 위험과 더불어 사회적 고립감이나 우울증 등의 심리·정서적 문제를 유발하는 것으로 보고된다(안홍순 2016; 진미성 외 2019). 직업계고 졸업자의 경우, 대부분의 학생이 대학교에 진학하는 일반계고 졸업자와 달리 고등학교라는 공교육 제도권에서 벗어난 이후에는 이들을 위한 직접적인 정책 개입이 어려워지기 때문에 장기 NEET 상태에 머무르는 유형에 직업계고 졸업자에 대해서 보다 정책적 관심을 기울일 필요성이 제기된다. 이에 본 연구에서는 진로이행과정 유형 예측 모델 분석을 통해 도출된 주요 결과를 기반으로 진로이행과정 특성을 반영한 정책적 시사점을 도출하고자 한다.

우리는 앞서 도출된 진로이행과정 유형을 예측하는 개인, 고등학교 생활, 가정 생활 변인을 탐색하기 위하여 초기 정규직 정착형을 참조집단으로 설정하고 장기 NEET형, 대학 진학형, 점진적 정규직 정착형 각각을 비교집단으로 설정한 세 개의 예측 모델 분석을 수행함으로써 주요 예측 변인을 도출하고, 로지스틱 회귀분석을 통해 주요 예측 변인과 진로이행과정 유형 간의 관계를 통계적으로 검증하였다. 먼저, 유형 분류에 있어서 공통적으로 나타난 결과에 대해 통계적으로 유의하게 나타난 변인을 중심으로 주요 결과를 논의하면 아래와 같다.

첫째, 자격증 개수, 취업·자격증 방과후학교 참여 여부, 취업상담 참여 여부 변인은 ‘장기 NEET형’, ‘대학 진학형’, ‘점진적 정규직 정착형’ 각각을 비교집단으로 설정한 세 가지 모든 경우에서 주요 예측 변인으로 도출되었으며 자격증 개수가 많고 취업·자격증 방과후학교에 참여한 경험이 있을수록 ‘초기 정규직 정착형’ 유형으로 분류될 확률이 유의하게 높은 것으로 나타났다. 이는 직업계고 졸업자가 졸업 후 정규직 정착에 있어 자격증 취득이 중요한 요소임을 시사한다. 이찬 외(2023)의 연구에 따르면 직업계고 학생들이 가장 필요하다고 꼽은 정책이 ‘자격증 비용 지원 제도’라고 나타났으며, 교육부에서도 2022년 직업계고 취업역량 강화 사업을 통해 자격증 취득 비용을 지원하는 사업을 추진한 바 있다(교육부 2022). 이에 따라 정부에서는 현장의 요구 사항을 반영하여 자격증 취득 지원과 같이 실질적으로 도움이 될 수 있는 사업을 적극적으로 추진할 필요가 있다. 더불어 다양한 학교의 진로 교육 활동 유형 중 ‘취업상담’ 경험 여부가 주요 예측 변인으로 도출되었는데, 해당 경험이 있는 학생일수록 ‘초기 정규직 정착형’에 분류될 확률이 유의하게 높은 것으로 나타났다. 관련 선행연구에 따르면, 단순히 비상호작용적인 매체를 통해 정보를 수동적으로 전달하는 방식의 진로상담보다는 학교 선배나 전문종사자와의 상호

작용을 통해 이루어지는 진로상담과 같은 취업 지도 방식이 직업계고 학생의 적극적인 진로 탐색을 유도하는 데 효과적인 것으로 보고된 바 있다(김소영·홍세희 2013). 이를 종합했을 때 취업상담이 초기의 정규직 정착을 지원하는 효과적인 진로교육 활동임을 시사하며, 따라서 직업계고에서는 학생 개개인의 진로 계획을 지도하는 담임, 진로상담교사뿐만 아니라 실제 본인의 전공에 대한 진로이행 경험이 있는 선배나 전문종사자와의 취업상담이 정기적으로 이루어지도록 할 필요가 있다.

둘째, 개인 변인 중 성별 변인 또한 공통적으로 도출된 주요 예측 변인이었으며 여성보다 남성이 ‘초기 정규직 정착형’보다 ‘장기 NEET형’과 ‘점진적 정규직 정착형’으로 분류될 확률이 높은 것으로 나타났다. 본 연구에서는 선행연구를 참고하여(정재현 외 2021; OECD 2014) 군복무와 군복무 대기 또한 비자발적 NEET 유형으로 구분하였는데, 분류된 ‘장기 NEET형’ 진로이행과정을 살펴보면, 고등학교 졸업 직후 2~3년 이내에 비자발적 NEET 상태에 머무르는 빈도가 높은 것을 확인할 수 있었다. 이는 직업계고를 졸업한 남성이 병역 의무를 수행함에 따라 발생한 비자발적 NEET 상태가 반영된 결과로 해석되며, 앞서 살펴본 바와 같이 남성이 ‘장기 NEET’형에 분류될 확률이 유의하게 높게 나타난 결과와도 맥락을 같이하는 결과이다. 즉, 장기 NEET 유형은 군 입대와 같은 남성의 특수성과 관련이 깊은 유형으로, 이를 고려한 정책적 접근이 필요함을 시사한다. 우리나라에서는 성인 남성의 대다수가 초기 성인기에 병역 의무를 수행하며, 기업에서 군 복무를 마친 구직자 채용을 선호함에 따라 발생하는 남학생의 고용 지위 불안정에 대한 정책적 관심을 기울일 필요가 있다(관계부처합동 2023). 따라서 병역 의무로 인한 취업의 어려움을 완화하기 위해 군 복무 후 취업 연계를 강화하는 등의 직업계고 졸업자의 성별에 따른 고용 지위 차이를 해소하는 맞춤형 지원 정책이 활성화되어야 한다

한편, 비교집단 설정에 따라 차별적으로 나타난 연구 결과 또한 확인할 수 있었다. 이를 살펴보면 첫째, 진로 역량은 ‘초기 정규직 정착형’과 ‘장기 NEET형’을 분류하는 주요 예측 변인으로 도출되었으며, 진로 역량 수준이 낮을수록 ‘장기 NEET형’으로 분류될 확률이 높은 것으로 나타났다. 이는 직업계고 학생의 진로 역량이 졸업 후 취업 성과에 긍정적인 영향을 미치는 것으로 보고한 선행연구(박소영·정혜원 2022)를 지지하는 결과이다. 따라서 진로 역량이 부족한 학생을 대상으로 진로 지도 프로그램을 제공하거나 고교학점제를 활용하여 자신의 수준과 진로 계획에 맞는 교과목을 선택하여 수강할 수 있도록 지도해야 한다. 특히 진로 역량과 관련하여 직업계고 졸업자의 진로 역량과 노동시장 진입 간의 관계를 검증한 이정민·유

명환(2021)의 연구 또한 직업계고 학생이 자기 이해를 바탕으로 성숙한 진로 의식을 가졌을 때 보다 안정적인 노동시장 진입을 예측할 수 있는 것으로 보고한 바 있다. 이에 따라 직업계고에서는 학생이 졸업 후 노동 시장에 안정적으로 정착할 수 있도록 자기 이해, 적성 탐색의 기회를 제공하고 학생의 진로 결정 수준 및 진로 계획을 파악하여 개개인에게 맞춤형 진로 교육 프로그램을 제공할 필요가 있다.

둘째, 희망교육수준이 ‘초기 정규직 정착형’과 ‘대학 진학형’을 분류하는 주요 예측 변인으로 도출되었으며, 희망교육수준이 높을수록 대학 진학형으로 분류될 확률이 높게 나타났다. 이러한 결과는 교사는 직업계고 학생의 진로 상담 및 지도 시에 학생의 희망교육수준, 내신 성적을 고려하여 장기적인 진로 계획을 수립할 수 있도록 지도하여야 함을 시사한다. 또한 전문교과 수업 이해도가 낮고 학교에서의 취업 처연계 경험이 없을수록 ‘대학 진학형’으로 분류될 확률이 높은 것으로 나타났다. 이러한 결과는 직업계고 졸업자가 취업지도 지원의 부족이나 자신의 전공에 대한 어려움 등의 요인으로 인하여 취업 유예나 차선책으로서의 대학 진학을 결정한 것은 아닌지, 직업계고 졸업 후 대학 진학의 목적과 의도에 대한 보다 심도 있는 연구가 수행될 필요가 있음을 시사한다.

이와 더불어 여성 보호자의 경제활동 여부의 경우 어머니와 같은 가정 내 여성 보호자가 경제활동에 참여한 학생일수록 ‘초기 정규직 정착형’보다 ‘대학 진학형’으로 분류될 확률이 유의하게 높은 것으로 나타났다. 이러한 연구결과는 개인의 진로 이행과정에 영향을 미치는 사회문화적 요인 중 하나로서 부모의 특성이 중요한 요소임을 함의한다. 기존의 연구에서는 어머니의 소득 기여가 있는 맞벌이 가구의 경우 외벌이 가구에 비해 자녀 교육비에 더 많은 투자가 이루어지는 것으로 보고된 바 있으며(이재림 2022) 이에 따라 여성 보호자가 경제활동에 참여하는 가정은 자녀가 대학에 진학할 수 있도록 경제적·심리적 기반을 제공할 수 있는 것으로 해석된다. 한편 김영혜·안현의(2014)의 연구에서는 아버지와 어머니의 학력, 직업 정보를 통합한 변인과 초기 성인기 자녀의 진로결정수준 간의 관계를 분석한 바 있다. 그 결과, 부모의 사회경제적 지위가 진로결정수준에 유의한 직접적인 영향을 미치지 않는으나 부모의 성취압력을 매개로 간접적인 영향을 미치는 것으로 보고하였다. 이를 고려했을 때, 본 연구에서는 SHAP 알고리즘 분석을 통해 도출된 주요 변인인 ‘여성 보호자의 경제활동 여부’ 변인만을 활용하였으나 후속 연구에서는 부모의 사회경제적 지위를 나타내는 다양한 지표를 활용하고, 부모의 사회경제적 지위와 자녀의 진로이행과정 간의 복잡한 구조적 관계를 보다 심층적으로 검증할 필요

가 있다.

셋째, 직업선택의 우선순위가 ‘점진적 정규직 정착형’과 ‘초기 정규직 정착형’을 분류하는 주요 예측 변인인 것으로 나타났는데 구체적으로, 여가시간과 쾌적한 환경이 보장되는 직업보다는 소득과 지속적인 고용이 보장되는 것이 중요하다고 응답한 학생이 초기 정규직 정착형에 속할 확률이 유의하게 높은 것으로 나타났다. 이러한 연구결과는 고등학교 시기 진로 역량과 진로 가치관, 진로 결정과 같은 진로 발달이 직업계고 졸업 후 진로이행에 영향을 미치는 요인임을 시사하며, 일과 삶의 균형보다는 소득과 지속적인 고용과 같은 경제적인 안정성을 추구하는 성향을 가진 학생들이 직업계고 졸업한 직후 정규직에 취업하여 이를 오랜 기간 유지하는 것으로 해석할 수 있다.

본 연구의 의의와 더불어 한계점을 기반으로 한 후속연구를 위한 논의는 다음과 같다. 본 연구에서는 시퀀스 분석을 통하여 직업계고 졸업자의 취업, 대학 진학, 구직 및 취업 준비, NEET 상태와 같이 다양한 진로이행상태를 포괄하여 종단적인 진로이행과정의 유형을 탐색하였다. 더불어 도출된 진로이행유형을 분류하는 예측 변인을 분석함에 있어서 장기 NEET형, 대학 진학형, 점진적 정규직 정착형을 각각 비교집단으로 설정하여 각 모형별로 공통적으로 나타난 주요 예측 변인과 비교집단 설정에 따라 특징적으로 나타난 주요 변인을 도출하고, 이를 기반으로 오늘날 정책적 관심의 대상인 ‘대학 진학형’, ‘장기 NEET형’과 같은 유형의 직업계고 졸업자의 진로이행과정별 맞춤형 지원 방향에 대해 논의하였다는 데 의의가 있다.

한편 본 연구에서는 행복도와 국어·영어·수학 내신등급, 건강 상태의 경우 진로이행과정 유형 분류에 미치는 영향력은 통계적으로 유의하지 않은 것으로 나타났으나 세 변인 모두 극단적으로 높거나 낮을 경우 ‘장기 NEET형’에 속할 확률이 증가하는 공통된 비선형적 양상을 나타냈다. 이는 스트레스, 낮은 자존감과 장애, 질병과 같은 부정적인 심리·건강상태에 있는 학생일수록 NEET 상태에 처할 위험이 높아지고(김기민·전지연 2023; 노현주·김운태, 2019; 전영창·이선화, 2024) 내신등급이 높더라도 기대 불일치로 인하여 진로이행에 어려움을 겪는 것으로 해석할 수 있다. 그러나 이러한 변인과 진로이행유형 간의 비선형적인 관계를 보다 명확하게 밝히기 위해서는 후속연구를 통해 다른 변인 간의 상호작용 관계를 파악하는 등 추가적인 분석이 수행될 필요가 있을 것으로 사료된다. 이외에도 머신러닝 분석을 통해 그간 선행연구에서 충분히 다루어지지 않은 독서 활동빈도가 주요 예측 변인으로 도출되었으며, 유형 분류에 미치는 영향력이 통계적으로 유의하지 않았으나 SHAP

의존성 지표 분석을 통해 독서 빈도가 높을수록 대학 진학형보다 초기 정규직 정착형에 속할 확률이 높은 것으로 나타났다. 이와 관련하여, 최진수·정혜원(2021)의 연구에서는 독서 선호도가 개인의 취업 선호도 유형에 영향을 미치는 주요 요인인 것으로 나타난 바 있는데, 독서 선호도가 높을수록 정규직 여부와 같은 취업 선호도에 대한 수준 즉, 취업준비가 높은 것으로 보고된 바 있어 개인이 여가시간을 보내는 경향과 진로이행 간의 관계를 검증하는 연구도 수행되길 기대한다.

참고문헌

- 교육부. 2022. “교육부는 직업계고 취업역량 강화 사업을 실효성 있게 추진해 나가겠습니다.” 교육부 설명자료(12월 6일).
- 관계부처합동. 2023. 《중등직업교육 발전방안》
- 김남식. 2023. “청소년-성인 이행기의 경력경로와 노동시장 성과분석: 직업교육훈련 연계를 중심으로.” 《직업능력개발연구》 26(3): 1-40.
- 김소영·홍세희. 2013. “전문계고 졸업자의 진로준비행동 잠재프로파일 분류와 고등학교에서 받은 진로지도의 관련성 검증” 《조사연구》 14(1): 93-116.
- 김기민·전지연. 2023. “청년(NEET)의 진로발달 유형 및 영향요인 분석.” 《진로교육연구》 36(3): 1-22.
- 김영혜·안현의. 2014. “대학생의 부모-자녀 유대, 부모의 사회경제적 지위와 진로결정관계에서 성취압력 및 진로결정자기효능감의 매개효과.” 《한국심리학회지: 상담 및 심리치료》 26(3): 657-682.
- 김지효. 2018. “특성화고 졸업생의 취업에 미치는 영향요인 분석: 연도별 비교를 중심으로.” 《취업진로연구》 8(3): 1-20.
- 김진모·전영욱·류지은·김영홍·정지용. 2018. “직업계고 졸업자의 경력유형별 학교생활과 노동시장성과 분석.” 《농업교육과 인적자원개발》 50(2): 105-126.
- 김진희·김준엽. 2021. “랜덤 포레스트 회귀를 이용한 중고등학교 1학년의 교우관계 영향요인 탐색.” 《교육방법연구》 33(3): 589-612.
- 김현주·김준영. 2010. “청소년의 진로 이행과 청소년, 부모의 인적 및 사회적 자본과의 관계.” 《청소년복지연구》 12(2): 23-52.
- 나상현. 2024. “취업난 직업계고 학생들 “일단 대학 진학합니다”... 고졸 취업자 수 큰폭 감소.” 중앙일보. 2024.05.28.
<https://www.joongang.co.kr/article/25252227>.

- 노법래·문영민. 2020. “장애인 노동시장 참여의 중단적 패턴은 어떻게 변화했는가?: 노동시장 참여 시퀀스의 코호트 비교를 중심으로.” 《한국사회복지학》 72(2): 349-371.
- 노현주·김윤태. 2019. “한국 청년 니트(NEET)의 특성과 영향 요인: 학력 수준별 비교 분석.” 《공공사회연구》 9(4): 73-105.
- 류지은·김성남·안재영·강서울. 2022. 《직업계고 졸업자의 경력경로 유형과 특성 (2022-11)》 세종: 한국직업능력연구원.
- 류지은·김성남·안재영·강서울. 2023. “청년층 직업계고 졸업자의 경력경로 형성 경험과 특성.” 《농업교육과 인적자원개발》 55(1):27-54.
- 박고은. 2023. “취업 힘들다보니... 대학 가능 직업계고 졸업생 또 늘었다.” 한겨레. 2023. 10.31.
<https://www.hani.co.kr/arti/society/schooling/1114316.html>.
- 박소영·정혜원. 2022. “직업계고 학생의 진로개발역량에 영향을 미치는 주요 변인 탐색 및 취업성과 예측.” 《교육학연구》 60(1): 157-182.
- 변수용·김경근. 2012. “전문계 고등학교 졸업자의 대학 진학 영향요인.” 《한국교육》 39(2): 79-107.
- 안홍순. 2016. “청년세대를 위한 사회적 형평성 제고 방안.” 《사회복지정책》 43(3): 59-83.
- 양준모·배관표. 2024. “특성화고등학교 진로상담의 영향에 관한 연구: 진로탐색수준의 조절효과를 중심으로.” 《청소년문화포럼》 79(-): 67-99,
- 오소욱·이다인·공민정. 2023. “청년층 초기 일자리의 시퀀스 특성이 경력정착기 노동시장 성과에 미치는 영향: OMA와 CatBoost 기법을 중심으로.” 《직업능력개발연구》 26(3): 135-171.
- 이선영·이호준·양희준·김현진·이승호·김나영·송경오·김훈호·김영식·박나실·조혜진. 2020. 《학생의 학교 참여 수준과 특징 분석 (RR 2020-08)》 진천: 한국교육개발원.
- 이정민·유명환. 2021. 《직업계고 학생의 개인적 자원과 노동시장 진입의 관계 (2021-30)》 세종: 한국직업능력연구원.
- 이재림. 2022. “맞벌이 가구의 구성과 특성.” 《한국의 사회동향 202》 대전: 통계개발원.
- 이찬·송낙현·최현정. 2023. “중등 직업교육 정책에 관한 직업계고 교원 및 학생의 인식 분석.” 《한국기술교육학회지》 23(1): 129-152.
- 전영창·이선화. 2024. “청년니트(NEET) 유형과 영향요인.” 《사회과학연구》 63(1): 329-350.
- 정재현·정연순·송수종·김태환·김기현·전예원·이혜나·양찬주. 2021. 《청년고용정책 사각지대 발굴 및 정책 제언 (기본연구 2021-15)》 충북: 한국고용정보원.

- 진미성·한준·노신애. 2019. “20-30대 청년세대의 결혼·출산 가치관의 잠재유형과 한국 사회 인식 및 개인적 미래 전망의 관련성.” 《가족과 문화》 31(1): 166-188.
- 최진수·정혜원. 2021. “K-means 군집분석 및 랜덤포레스트를 적용한 취업선호도 유형화 영향요인 분석” 《교육학연구》 59(8): 113-144.
- 한국교육개발원·교육부. 2023. 《2023년 직업계고 졸업자 취업통계연보》 진천: 한국교육개발원.
- 한국교육개발원·교육부. 2024. 《2024년 직업계고 졸업자 취업통계연보》 진천: 한국교육개발원.

- Abbott, A. 1983. “Sequences of Social Events: Concepts and Methods for the Analysis of Order in Social Processes.” *Historical Methods* 16(4): 129-147.
- Breiman, L. 2001. “Random Forests.” *Machine Learning* 45(1): 5-32.
- Chen, T. and C. Guestrin. 2016. “XGBoost: A Scalable Tree Boosting System.” *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: 785-794.
- Everitt, B. 1993. *Cluster Analysis* (3rd ed). New York, NY: Halsted Press.
- Friedman, J.H. 2001. “Greedy Function Approximation: A Gradient Boosting Machine.” *Annals of Statistics* 29(5): 1189-1232.
- Geurts, P., D. Ernst, and L. Wehenkel. 2006. “Extremely Randomized Trees.” *Machine Learning* 63(1): 3-42.
- Ke, G., Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.Y. Liu. 2017. “LightGBM: A Highly Efficient Gradient Boosting Decision Tree.” *Advances in Neural Information Processing Systems* 30: 3146-3154.
- Kline, R.B. 2023. *Principles and Practice of Structural Equation Modeling*. Guilford Publications.
- Lundberg, S.M., G. Erion, H. Chen, A. DeGrave, J.M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.I. Lee. 2020. “From Local Explanations to Global Understanding with Explainable AI for Trees.” *Nature Machine Intelligence* 2: 56-67.
- MacQueen, J. 1967. “Some Methods for Classification and Analysis of Multivariate Observations.” *Berkeley Symposium on Mathematical Statistics and Probability* 1(-): 281-29.
- OECD. 2014. *Education at a glance 2014*: OECD indicators. Paris: OECD Publishing.
- OECD. 2019. *Society at a Glance 2019*: OECD Indicators Paris: OECD Publishing.
- Prokhorenkova, L., G. Gusev, A. Vorobev, A.V. Dorogush, and A. Gulin. 2018. “CatBoost:

Unbiased Boosting with Categorical Features.” *Advances in Neural Information Processing Systems* 31: 6639–6649.

Quinlan, J.R. 1986. “Induction of Decision Trees.” *Machine Learning* 1(1): 81–106.

<접수 2025.03.21; 수정 2025.04.16; 게재확정 2025.05.18>

Exploring the Vocational High School Graduates' Type of Career Transition Process and Key Predictive Variables: Using Sequence Analysis and Explainable Artificial Intelligence

Yeheun Baek
(Chungnam university)
Hyewon Chung
(Chungnam university)

This study explored the types of career transition processes among vocational high school graduates and examined the key predictors of each type. Using data from Waves 3 (2019) to 7 (2023) of the Korea Education and Employment Panel II (KEEP II), sequence analysis was conducted to capture longitudinal patterns in career transition states, including college enrollment status, employment type (regular/non-regular), job-seeking experience, and participation in employment preparation activities. Based on the sequence results, k-means cluster analysis was used to categorize four distinct types of career transitions: 'Early Regular Employment', 'Gradual Regular Employment', 'College Enrollment', and 'Long-term NEET'. To identify predictors of these types, the best-performing prediction model was selected through an AutoML algorithm, and key variables were interpreted using the SHAP algorithm and logistic regression analysis. The findings revealed that students with a greater number of certifications and experiences in career-related after-school programs and career counseling were more likely to follow the 'Early Regular Employment'. In contrast, those with weaker career competencies were more likely to be classified as 'Long-term NEET', while students with higher desired education levels were more likely to follow the 'College Enrollment'. Based on these differentiated patterns, the study discusses implications for tailored career support strategies for vocational high school graduates, aligned with their specific career transition trajectories.

Key words: career transition, vocational high school graduates, sequence analysis, AutoML algorithm, SHAP algorithm