

머신러닝 기반 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준 관련 경험 구조 탐색

김혜인* · 강태훈** 세종대학교

초록 본 연구는 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준을 구분하는 입시 준비 행동 및 경험 요인을 탐색하는 데 목적이 있다. 이를 위해 2025년 9월부터 12월까지 수집된 수도권 A 대학교 실용무용 실기전형 지원자 807명($N=807$)의 설문 데이터를 2차 자료로 활용하였다. 무용 경력 수준은 선행연구의 경력 구분 기준을 참고하여 3년 이하 집단(Short), 4년 이상 집단(Long)으로 구분하였다. 분석 전 불필요 변수 제거, 가족 구성 변수 재구성, 파생변수 생성, 원-핫 인코딩, 수치형 변수 표준화 등의 전처리를 실시하였으며, 데이터 누수를 방지하기 위해 집단 구분 기준으로 사용된 무용 경력 기간 변수는 분석 변수에서 제외하였다. 또한 학습용 자료와 검증용 자료를 7:3으로 층화 분할하고, 클래스 불균형 보정을 위해 학습 데이터에 한하여 SMOTE를 적용하였다. 분석에는 머신러닝 기반 지도 학습 방법을 활용하였으며, 랜덤포레스트(Random Forest) 변수 중요도 분석과 Mann-Whitney U 검정, 카이제곱 검정을 함께 적용하였다. 연구 결과, 공연 횟수, 시험장 이동 시간, 모의면접 경험, 입시 준비 기간, 대회 수상 경험 등이 입시 시점 무용 경력 수준 집단을 구분하는 주요 변수로 나타났다. 또한 파생변수를 포함한 분석에서는 Mobility_Challenge와 Preparation_Performance가 주요 변수로 확인되어, 단일 경험보다 공연 참여, 이동 특성, 입시 준비 경험이 결합된 구조가 집단 구분과 관련될 가능성을 보여주었다. 본 연구는 특정 대학 실기전형 지원자 자료를 활용한 탐색적 연구이므로 전체 실용무용 전공자에게 일반화하는 데에는 주의가 필요하며, 분석 결과는 인과관계가 아닌 연관성 수준에서 해석되어야 한다. 그럼에도 본 연구는 실용무용 전공 진입 이전 단계의 입시 준비 경험 구조를 머신러닝 기반으로 탐색하였다는 점에서 의의가 있다.

주요어 : 실용무용, 무용 경력 수준, 머신러닝(Machine Learning), 랜덤포레스트(Random Forest), 탐색적 자료분석(EDA)

I. 서론

1. 연구의 필요성 및 목적

국내 힙합 문화의 유입 이후 스트릿댄스는 대중문화 영역에서 큰 성장을 이루었으며, 다양한 콘텐츠 제작을 통해 국내외 대중의 관심이 확대되고 있다. 특히 댄서들의 움직임, 스타일, 패션 등이 대중문화의 유행 요소로 확산되면서 스트릿댄스 기반 실용무용에 대한 사회적 관심도 함께 증가하고 있다(김나현, 2023). 이러한 흐름은 10대 청소년들의 실용무용에 대한 관심을 높였으며, 댄스와 안무가를 진로로 인식하는 경향으로 이어지고 있다(박한솔, 2023). 이에 따라 실용무용 전공에 대한 수요가 증가하였고, 교육과정도 점차 제도적 기반을 갖추어 왔다(박영하, 2015; 정한결, 2020; 조준희, 2025).

이러한 변화 속에서 실용무용 전공 진학을 준비하는 청소년들의 입시 준비 과정도 점차 전문화되고 있다. 실용무용 실기전형 지원자들은 입시 시점에 이르기까지 전문 교육기관 수강, 장르 선택, 공연 참여, 대회 출전, 워크숍

* 세종대학교 무용학과 박사과정, waackingmarid029@naver.com, 제1저자

** 세종대학교 무용학과 박사과정, kangrumbles@naver.com, 교신저자

참여, 영상 촬영 등 다양한 학습 경험과 실기 활동을 축적한다. 이 과정에서 일부 지원자는 비교적 이른 시기부터 장기간 실기 경험을 누적하는 반면, 다른 지원자는 입시를 앞둔 시점에서 단기간에 집중적으로 준비하기도 한다. 따라서 실용무용 실기전형 준비는 단순한 시험 대비 과정이라기보다, 입시 시점까지 축적된 학습 경험과 준비 행동이 반영된 과정으로 이해할 필요가 있다.

선행연구를 살펴보면, 실용무용 전공자의 전공 유지 및 진로에 관한 연구는 주로 개인의 심리적·인식적 요인을 중심으로 이루어져 왔다. 전공 만족도, 학습 몰입, 진로 인식, 지도자 요인, 진로 결정 장애 등은 전공자의 전공 유지 의도나 진로 태도를 설명하는 주요 변인으로 꾸준히 다루어져 왔으며(김은영, 2025; 문다예, 2022; 박수현, 2021; 이보람, 권소영, 2016; 최용원, 2020), 이러한 요인들이 진로 인식과 준비 양상에 영향을 미친다는 점이 제시되었다. 또한 고정민, 안병순(2025)은 경력 기간에 따른 진로 결정 장애 요인과 진로 교육 유형에 주목하여, 실용무용 전공자가 보유한 경력의 양과 질이 진로 인식 및 준비행동 수준과 관련될 수 있음을 논의하였다.

그러나 이들 연구에서 다루어진 전공 만족도, 학습 몰입, 진로 인식 등의 요인은 전공 진입 이후에만 형성되는 것이 아니라, 전공 진입 이전 단계에서 축적된 학습 경험과 실기 활동의 영향을 받을 가능성이 있다. 특히 경력 기간에 따라 진로결정 장애 요인, 진로교육 유형, 취업 준비 여부에 차이가 나타날 수 있음을 제시한 점은, 전공 진입 이후의 진로 형성 과정이 입시 이전에 축적된 경험과도 관련될 수 있음을 시사한다. 따라서 전공 진입 이후의 전공 유지 및 진로 준비를 보다 입체적으로 이해하기 위해서는, 그 이전 단계인 입시 시점에서 지원자가 어떠한 학습 경험과 준비 행동을 축적해 왔는지를 함께 살펴볼 필요가 있다. 그럼에도 기존 연구는 대체로 대학 전공자 또는 전공 진입 이후의 학습자를 대상으로 하였으며, 실용무용 실기전형 지원자의 입시 시점 경험 구조에 대한 논의는 상대적으로 부족한 실정이다.

입시 시점의 무용 경력 수준을 변별하는 것은 단순히 경력 기간의 차이를 확인하는 데 그치지 않는다. 이는 전공 교육과정에 진입하기 이전에 지원자가 어떠한 경험적 기반을 형성해 왔는지를 파악하고, 실용무용 전공자의 교육 및 진로 형성 과정을 입시 이전 단계까지 확장하여 이해하는 데 의미가 있다. 특히 실용무용 입시 준비 과정에서는 공연 참여, 대회 출전, 장르 선택, 연습량 조절, 모의평가 참여, 진로 정보 탐색 등 다양한 경험과 행동이 함께 작용한다. 따라서 실용무용 실기전형 지원자의 입시 준비 행동을 이해하기 위해서는 개별 요인의 영향뿐 아니라, 다양한 실기 경험과 학습 환경이 결합되는 구조를 함께 살펴볼 필요가 있다.

실용무용 입시 준비 행동은 다양한 경험 요인이 함께 작용하는 과정이므로, 이를 분석하기 위해서는 여러 변수를 동시에 고려할 수 있는 탐색적 접근이 필요하다. 머신러닝은 데이터를 바탕으로 규칙과 패턴을 학습하여 예측과 분류를 수행하는 방법으로(Domingos, 2012), 변수 간 관계를 사전에 단일한 경로로 설정하지 않더라도 집단 구분에 기여하는 주요 변수를 탐색할 수 있다. 또한 머신러닝 기반 분석은 다양한 분야에서 예측모델 구축과 변수 중요도 탐색에 활용되어 왔으며(문재웅, 2022; 백세룡, 2022), Halat et al.(2023) 역시 학생 성과와 관련된 주요 변수를 탐색하는 데 머신러닝 기반 특징 중요도 분석을 활용한 바 있다. 따라서 본 연구에서는 머신러닝 기반 분류 모델을 활용하여 입시 시점 무용 경력 수준 집단을 구분하는 주요 입시 준비 행동 및 경험 요인을 탐색하고자 한다.

이에 본 연구는 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준을 결과 변수로 설정하고, 전공 진입 이전의 입시 준비 단계에서 형성된 학습 경험과 준비 행동 요인이 어떠한 구조로 나타나는지를 탐색하고자 한다. 이를 위해 김수산나와 안귀호(2023), 김정숙(2008, 2009), 이준모(2019)의 연구 기준을 참고하여 입시 시점 무용 경력 기간 3년을 기준으로 3년 이하 경력 집단과 4년 이상 경력 집단으로 구분하였다. 본 연구에서 사용하는 Short와 Long은 분석 편의를 위한 집단 명칭으로, Short는 입시 시점 무용 경력 기간 3년 이하 경력 집단을, Long은 입시 시점 상대적으로 긴 무용 경력 기간을 가진 4년 이상 경력 집단을 의미한다. 다만 이 기준은 실용무

용 학습자의 숙련 수준이나 장기적 활동 지속 여부를 판단하기 위한 절대적 기준이 아니라, 입시 시점에서 지원자가 보유한 무용 경력 수준을 상대적으로 구분하기 위해 설정한 탐색적 기준이다.

또한 본 연구는 인구통계학적 변수, 학습 경험 관련 변수, 입시 준비 행동 관련 변수, 파생변수를 단계적으로 포함하여 입시 시점 무용 경력 수준 집단을 구분하는 주요 변수와 변수 중요도 구조의 변화를 확인하고자 한다. 나아가 3년 이하 경력 집단 중 모델상 입시 시점에서 상대적으로 긴 무용 경력 기간을 가진 4년 이상 경력 집단과 유사한 변수 패턴을 보이는 사례의 입시 준비 행동 특성을 탐색함으로써, 실용무용 실기전형 지원자의 무용 경력 수준이 단일 요인만으로 설명되기보다 다양한 경험과 조건의 조합 속에서 나타나는지를 살펴보고자 한다. 이를 통해 실용무용 전공 진입 이전 단계의 입시 준비 행동을 이해하고, 실용무용 교육 및 입시 지도에 필요한 기초 자료를 마련하는 데 목적이 있다.

2. 연구문제

본 연구의 목적을 달성하기 위하여 설정된 연구문제는 다음과 같다.

첫째, 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준(3년 기준)에 따라 입시 준비 행동 및 경험 특성은 어떠한 차이를 보이는가?

둘째, 기본 변수와 파생변수를 포함함에 따라, 무용 경력 수준 집단을 구분하는 변수 중요도 구조는 어떻게 변화하는가?

셋째, 3년 이하 경력 집단에서 4년 이상 경력 집단으로 분류될 가능성과 관련된 주요 입시 준비 행동 및 경험 요인은 무엇인가?

II. 연구방법

1. 자료수집

본 연구는 2025년 9월부터 12월까지 수집된 수도권 A 대학교 실용무용 실기전형 전형 지원자를 대상으로 수집된 설문 데이터를 행정 부서의 협조를 통해 제공받아 2차 데이터로 활용하였다. 최종적으로 사용된 표본은 총 807명($N=807$)이며, 해당 자료는 실기전형 지원자의 무용 활동 특성 및 입시 준비 경험을 파악하기 위해 구축된 설문 데이터이다. 본 연구에서는 수험번호와 연령 등 개인 식별 가능성이 있는 정보가 제외된 비식별 분석자료를 제공받았으며, 입시 평가와 연구 분석은 분리하여 진행하였다.

설문 문항은 무용 경력 기간, 공연 경험, 대회 성취 경험, 입시 준비 기간, 소속 여부, 수업 형태, 경제, 가족 환경 등 수험생의 무용 활동 및 준비 특성 관련 항목이 포함되어 있다. 본 연구는 입시 시점 기준 무용 경력 수준에 따라 나타나는 경험 특성 및 준비 행동의 차이를 탐색하는 데 목적이 있다. 연구대상자의 일반적 특성은 <표 1>과 같다. 본 연구에서는 제공받은 분석자료의 범위 내에서 성별, 지원조건 유형, 고등학교 유형, 거주 지역, 무용 경력 기간 및 무용 경력 수준을 제시하였다.

표 1. 연구대상자의 일반적 특성

구분	항목	빈도(N)	비율(%)
성별	여성	732	90.7
	남성	75	9.3
지원조건 유형	검정고시	49	6.1
	고등학교 졸업(편입 포함)	758	93.9
고등학교 유형 I	예술계열	74	9.2
	인문·상공계열	727	90.1
	무응답/기타	6	0.7
고등학교 유형 II	여고/남고	194	24.0
	남녀공학	608	75.3
	무응답/기타	5	0.6
거주 지역	서울	169	20.9
	경기·인천	353	43.7
	강원	32	4.0
	충청·대전	54	6.7
	경상	98	12.1
	전라	54	6.7
	제주	3	0.4
	그 외	44	5.5
무용 경력 기간	1년 이하	193	23.9
	1~3년	465	57.6
	4~6년	113	14.0
	7~9년	25	3.1
	10년 이상	11	1.4
무용 경력 수준	3년 이하 경력 집단	658	81.5
	4년 이상 경력 집단	149	18.5
전체		807	100.0

본 연구는 세종대학교 생명윤리심의위원회의 심의면제를 받았으며(IRB No. SUIRB-HR-E-2026-006), 개인 식별정보가 제거된 비식별화 자료를 제공받아 분석하였다. 연구자는 개인을 식별할 수 있는 정보에는 접근하지 않았다.

다만 본 연구는 단일 대학 실기전형 지원자를 대상으로 수집된 자료를 활용하였으므로, 본 연구 결과를 모든 실용무용 입시 지원자에게 일반화하는 데에는 한계가 있다.

2. 자료분석

1) 데이터 전처리 및 변수 구성

본 연구에서는 수집된 설문 데이터를 Google Colab 환경에서 전처리 및 분석하였다. 총 50개의 설문 변수를 활용하였으며, 변수는 인구통계학적 특성(성별, 신체 특성 등), 생활 특성(MBTI, 독서량, 스마트폰 사용 시간 등), 입시 준비 행동(입시 준비 기간, 모의면접 경험, 중복지원 여부 등), 무용 활동 경험(공연 경험, 대회 수상 경험, 무용 경력 기간 등), 시험 당일 행동 특성(수면 시간, 식사 여부, 카페인 섭취 여부 등), 환경 및 이동 특성(거주 지역, 시험장 이동 시간 등)으로 구성되었다.

분석 전 불필요한 변수를 제거하고 변수 구조를 변환하였다. 분석에 직접 활용되지 않는 Enrolled(등록 여부) 변수는 제거하였으며, Admitted(합격 여부)는 입시 결과 이후의 정보로 해석될 가능성이 있어 제외하였다. DancingPeriod(무용 경력 기간)는 선행연구에서 활용된 경력 구분 기준을 참고하여 3년 이하와 4년 이상 집단을 구분하는 종속변수(Target) 생성에만 활용하였으며, 데이터 누수를 방지하기 위해 예측변수에서는 제외하였다.

FamilyMembers(가족 구성)는 단일 변수로 해석하기 어려운 점을 고려하여 GuardianStatus_1(보호자 동반 여부)와 SiblingStatus_1(형제자매 여부)로 재구성한 후 기존 변수는 제거하였다.

이후 범주형 변수에는 원-핫 인코딩(One-Hot Encoding)을 적용하여 수치형 변수로 변환하였으며, 수치형 변수는 StandardScaler를 적용하여 표준화하였다. 결측치 확인 결과 전체 807명 자료에서 총 10건의 결측치가 확인되었으며, TypeofHighSchool1 5건, TypeofHighSchool2 4건, PropsUsage 1건으로 확인되었다. 해당 변수들은 범주형 변수이므로 최빈값 대체 방식으로 처리하였다. 또한 단일 요인으로는 설명하기 어려운 입시 준비 경험의 결합 특성을 반영하기 위해 파생변수를 생성하였다. 각 파생변수는 서로 다른 경험 요인의 결합 구조를 반영하도록 설계되었으며 <표 2>와 같다.

표 2. 파생변수 정의 및 산출 방식

파생변수명	구성변수	산출 방식
Performance_Award_Sum	공연 횟수, 수상 경험	$0.5 \times \text{공연횟수} + 0.5 \times \text{수상경험}$
Performance_Award_Product	공연 횟수, 수상 경험	$\text{공연횟수} \times \text{수상경험}$
Mobility_Challenge	공연 횟수, 이동 시간	$\text{공연횟수} \times \log(\text{이동시간}+1)$
Preparation_Performance	준비 기간, 공연 횟수	$\text{준비기간} \times (\text{공연횟수} + 1)$

공연 경험과 수상 경험을 동일한 수준의 경험 축적으로 보기 위해 동일 가중치(0.5:0.5)를 부여하였으며 (Performance_Award_Sum), 공연 경험과 수상경험이 동시에 존재할 때 효과를 보기 위하여 곱셈 방식으로 생성하였다(Performance_Award_Product). 수상 경험은 이분형 변수(없음=0, 있음=1)로 구성되어 있어, 수상 경험이 있는 경우에만 공연 경험이 반영되도록 구성하였다. 따라서 이 변수는 단순한 공연 경험이 아니라 수상 경험을 동반한 공연 경험의 특성을 탐색적으로 반영한다. 또한 본 연구에서는 거주지역에 따라 공연 활동 및 참여 기회에 차이가 존재할 수 있다는 점을 고려하여 Mobility_Challenge 변수를 생성하였다. 시험장까지의 이동시간이 길다는 것은 공연 활동을 위해 상대적으로 더 많은 이동을 감수했을 가능성을 보고, 공연 경험과 이동시간을 결합하여 이동 부담 속에서도 축적한 공연 경험의 특성을 탐색적으로 반영하고자 하였다. 이동시간은 값의 차이가 지나치게 커지는 것을 방지하기 위해 로그변환을 적용하였다. Preparation_Performance는 입시작품 준비기간과 공연 경험의 결합 구조를 반영하기 위해 생성하였다. 공연경험 변수에 1을 더한 것은 준비기간과 공연 경험을 곱하는 과정에서 공연 경험 수준이 낮은 집단의 정보가 지나치게 축소되는 것을 방지하기 위한 보정이다. 본 연구의 파생변수의 산출 방식은 특정 공식을 적용한 것이 아닌, 입시 준비 경험 간의 결합 구조를 탐색하기 위해 연구 목적에 맞게 구성하였다. 이러한 과정을 통해 기본 설문 변수를 활용한 기존 변수(feature set 1)와 파생변수를 추가한 feature set 2로 설정하였다.

전처리 과정에서는 불필요 변수 제거, 가족 구성 변수 재구성, 파생변수 생성 과정을 수행하여, 최종적으로 인코딩 및 스케일링 후 125개의 변수로 확장하여 분석에 활용하였다. 전처리 단계에 따른 변수 수 변화는 <표 3>에 제시하였다.

표 3. 전처리 단계에 따른 변수 수 변화

단계	행 수	변수 수
기존 데이터	807	50
불필요 변수 제거 후	807	47
가족 구성 변수 재구성 후	807	48
파생변수 추가 후	807	52
최종 분석용 X(인코딩 전)	807	52
인코딩 및 스케일링 후	807	125

2) 모델 구성 및 학습 방법

본 연구의 타깃 변수는 입시 시점 무용 경력 수준을 기준으로 구분한 이진 범주형 변수이므로, 분류(Classification) 문제로 설정하였다. 이에 따라 분류 알고리즘 중에서 랜덤포레스트(Random Forest), Gradient Boosting(GB), Light Gradient Boosting Machine(LGBM), eXtreme Gradient Boosting(XGBoost) 등 트리 기반 앙상블 모델을 포함한 다양한 분류 기법을 적용하였다(임승준, 2023). 각 feature set에 대해 데이터를 학습용과 검증용으로 7:3 비율로 분할하고, 층화 추출(Stratified Sampling)을 적용하여 Long과 Short 비율을 유지하였다. SMOTE(Synthetic Minority Over-sampling Technique) 적용 전 클래스 분포를 확인한 결과, Short 집단은 658(81.5%), Long 집단은 149명(18.5%)으로 나타나 클래스 불균형이 존재하였다(그림 1).

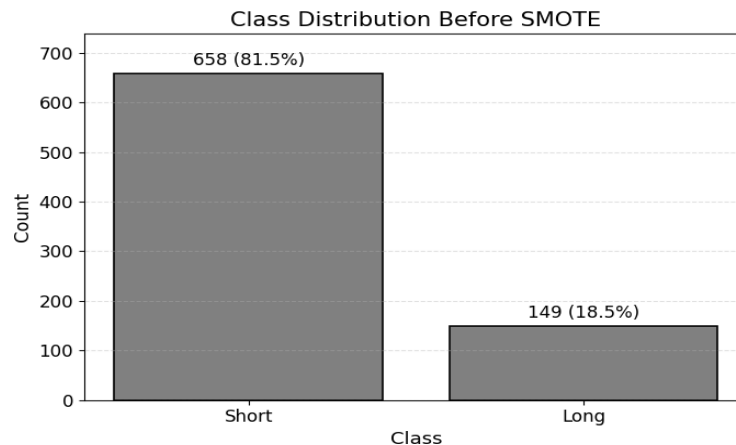


그림 1. SMOTE 적용 전 무용 경력 집단 클래스 분포

이후 학습 데이터의 불균형으로 인해 모델이 특정 집단에 편향되는 문제를 보정하기 위해 SMOTE 기법을 적용하여 모델 학습의 안정성을 확보하였으며, SMOTE는 검증 데이터에 영향을 주지 않도록 학습 데이터에만 적용하였다. 다만, SMOTE를 통해 생성된 데이터가 변수 중요도 해석에 영향을 미칠 가능성이 있으며, 특히 원-핫 인코딩된 범주형 변수에서는 실제 존재하지 않는 조합이 생성될 수 있다는 한계가 있다. 따라서 본 연구에서는 결과 해석 시 이러한 점을 고려하였다. 또한 모든 실험은 동일한 랜덤 시드(Random state)를 사용하여 재현 가능하도록 설정하였다.

3) 평가 방법 및 분석 전략

모델의 분류 성능은 ROC-AUC를 중심으로 평가하고 Accuracy를 보조 지표로 활용하였으며, 소수 집단(Long)에 대한 분류 성능 확인을 위해 Precision과 Recall을 함께 산출하였다. 그러나 본 연구의 핵심 목적은 최고 성능의 예측 모형을 제시하는 것이 아닌, 머신러닝 기반 분류 모델을 활용하여 입시 시점 무용 경력 수준 집단과 관련된 주요 변수 패턴과 구조를 탐색하는 데 있다. 이에 본 연구는 변수 중요도 해석과 분류 조건 탐색에 중점을 두었으며, 변수 중요도 산출과 해석 보조 가능성을 고려하여 본 연구는 랜덤포레스트 모델을 중심으로 분석을 수행하였다. 랜덤포레스트는 다수의 의사결정나무를 결합하여 분류 및 회귀를 수행하는 앙상블 기법으로, 변수 간 복잡한 관계와 비선형적 패턴을 탐색하는 데 활용될 수 있다(Breiman, 2001). 또한 랜덤포레스트는 변수 중요도 지표를 제공하므로, 예측 결과에 상대적으로 기여한 변수를 탐색하는 데 유용하다(Breiman, 2001; Liaw & Wiener, 2002). 이에 입시 준비 행동 및 경험 요인이 복합적으로 작용하는 본 연구의 탐색적 분석 목적에 적합한 모델로 판단하였다. 다만 랜덤포레스트의 변수 중요도 결과는 변수 중요도 산출 방식의 특성을 고려하여 이를 보완하기 위해 Mann-Whitney U 검정과 카이제곱 검정을 함께 실시하여 해석의 타당성을 확인하였다. 변수 중요도 결과를 위해 검정을 보조적으로 활용하였으며, 다중비교 보정은 적용하지 않았다. 결과 해석 시 제1종 오류 가능성을 함께 고려하였다.

또한 랜덤포레스트 기반 예측 모형을 이용하여 Short 집단이 Long 집단으로 분류될 예측 확률을 산출하고, 주요 경험 변수의 값이 변화에 따라 어떻게 달라지는지를 분석하였다. 이를 통해 입시 시점 무용 경력 수준 집단 구분과 관련된 핵심 요인을 도출 및 해석하고자 하였다.

본 연구는 관찰 데이터를 기반으로 하므로, 변수 간 관계는 연관성 수준에서 해석되며 인과관계로 보기는 어렵다. 따라서 변수 간의 관계는 관련성 수준에서만 해석하였으며, 특정 변수의 영향이나 인과관계를 의미하는 것은 아니다.

3. 자료의 신뢰도 및 타당성

분석 결과 해석의 타당성을 높이기 위해 무용학 박사 3인이 변수 중요도 결과와 해석 내용을 독립적으로 검토하였다. 이후 전문가 검토 내용을 종합하여 핵심 변수의 의미와 연구 결과에 대한 최종 해석을 도출하였으며, 이를 통해 머신러닝 분석 결과 해석에서 발생할 수 있는 주관적 판단을 최소화하고 연구 결과의 신뢰성을 보완하였다.

III. 연구 결과

1. 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준(3년 기준)에 따른 입시 준비 행동 및 경험 특성 차이

실용무용 실기전형 지원자의 입시 시점 무용 경력 수준 집단을 구분하는 주요 변수 특성을 확인하기 위해 분석을 수행하였다. 무용 경력 수준 집단(Long/Short)을 예측하는 주요 변수를 파악하기 위해 본 연구에서는 랜덤포레스트 모델을 적용하여 변수 중요도를 분석하였다. 랜덤포레스트 모델은 다수의 결정트리를 결합하여 예측을 수행하는 앙상블 기법으로, 각 변수의 중요도를 산출할 수 있는 특성을 지닌다(김세준, 2024). 이러한 특성은 공연 경험, 준비 과정, 환경 요인이 복합적으로 작용하는 본 연구 자료에 적합한 방법으로 판단되어, 랜덤포레스트 모델을 기반으로 집단 구분 변수의 중요도를 산출하였다. 랜덤포레스트 분석 결과 상위 10개 변수는 <표 4>와 같다.

표 4. 랜덤포레스트 기반 무용 경력 수준 집단 구분 변수의 중요도 상위 요인 10개

순위	변수	변수 설명	중요도
1	NumberOfPerformances	공연 횟수	0.0769
2	TimeTakenFromHomeToTheExamHall	시험장 이동 시간	0.0513
3	MockInterview	모의 면접 경험	0.0483
4	Weight	몸무게	0.0463
5	SleepDurationtheDayBefore	전날 수면 시간	0.0394
6	Height	키	0.0365
7	MBTI	성격 유형	0.0332
8	Part-timeJob	아르바이트 여부	0.0324
9	GivenMoney	용돈	0.0323
10	EntranceExamPractice(months)	입시 준비 기간	0.0300

랜덤포레스트 변수 중요도는 무용 경력 수준 집단을 분류하는 과정에서 각 예측변수가 기여한 상대적 중요도를 의미한다. 범주형 변수는 원-핫 인코딩(One-Hot Encoding)을 통해 분석하였으며, 결과는 기존 변수 기준으로 정리하여 비교하였다.

또한 변수 중요도만으로는 변수의 통계적 유의성을 판단하기 어렵기 때문에, Mann-Whitney U 검정과 Chi-square 검정을 추가적으로 실시하였다. 랜덤포레스트 분석 결과와 통계검정 결과를 종합적으로 확인하며 최종 변수를 선정하였다. 최종 변수는 예측 기여도와 통계적 유의성과 연구 목적에 따른 해석 가능성을 함께 고려하여 선정하였다.

그 결과 공연 횟수, 수상 경험, 입시 준비 기간, 모의 면접 경험, 거주 지역, 시험장 이동 시간이 총 6개의 핵심 구분 변수로 선별되었다(표 5)). 순서형 변수인 공연 횟수, 시험장 이동 시간 및 입시 준비 기간은 중앙값과 사분위범위(IQR)로 제시하였으며, 집단 간 차이는 Mann-Whitney U 검정을 통해 확인하였다. 또한 모의 면접과 수상 경험은 각각 1=경험 있음, 2=경험 없음으로, 거주 지역은 1=서울, 2=경기(인천), 3=강원, 4=충청(대전), 5=경상, 6=전라, 7=제주, 8=그 외로 코딩하였다. <그림 2>에는 최종적으로 선정된 6개 주요 변수의 랜덤포레스트 중요도를 시각화하여 제시하였다.

표 5. 핵심 구분 변수의 집단별 비교 및 변수 중요도

변수	Short	Long	<i>p</i>	중요도
NumberOfPerformances	3 (2-5)	5 (3-5)	<0.001	0.0769
TimeTakenFromHomeToTheExamHall	2 (2-3)	2 (2-4)	<0.001	0.0513
MockInterview	1 (1-2)	1 (1-2)	0.022	0.0483
EntranceExamPractice(months)	3 (2-3)	3 (2-3)	0.013	0.0300
CompetitionAwardsExperience	1 (1-2)	1 (1-2)	<0.001	0.0233
CurrentRegion	2 (2-4)	2 (2-5)	0.002	0.0112

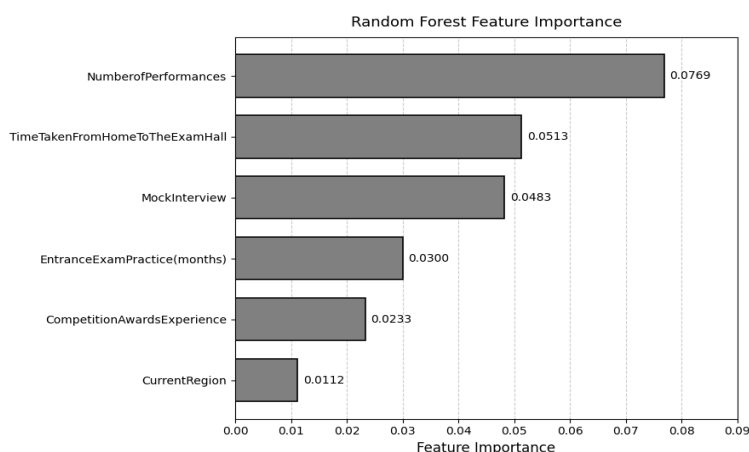


그림 2. 랜덤포레스트 기반 핵심 구분 변수 중요도

공연 횟수는 변수 중요도(0.0769)가 가장 높게 나타났으며, Long 집단의 중앙값이 Short 집단보다 높아 두 집단 간 유의한 차이가 확인되었다($p < .001$). 시험장 이동 시간 역시 Long 집단에서 상대적으로 높은 분포를 보여 두 집단 간 유의한 차이가 나타났다($p < .001$). 또한 모의면접 경험, 입시 준비 기간, 대회 수상 경험 및 거주 지역에서도 집단 간 유의한 차이가 확인되었다($p < .05$).

EntranceExamPractice(입시 준비 기간)는 입시 작품을 준비한 기간을 의미하며, 춤을 시작한 이후 입시 시점까지의 전체 무용 활동 기간을 의미하는 DancingPeriod(무용 경력 기간)와 구분된다. 따라서 입시 준비 기간은 무용 경력 기간과 별개의 입시 준비 행동 변수로 활용하였다. CurrentRegion(거주 지역)과 TimeTakenFromHomeToTheExamHall(시험장 이동 시간)은 입시 준비 과정에서 나타나는 환경적 특성과 관련된 변수로 볼 수 있으며, 무용 경력 수준 집단 구분 과정에서 일정한 관련성을 보였다.

공연 횟수, 대회 수상 경험 여부, 준비 기간과 같은 경험 관련 변수들이 높은 중요도를 나타낸 점에서, 무용 경력 수준은 환경적 특성뿐 아니라 반복적인 공연 참여와 준비 과정, 성취 경험 등 다양한 경험 요인이 함께 반영될 가능성이 있는 것으로 해석할 수 있다.

2. 파생변수 추가에 따른 변수 중요도 구조의 변화

파생변수를 포함할 경우 변수 중요도 구조의 변화를 확인하기 위해 본 연구에서는 변수 구성을 단계적으로 확장하여 <표 6>와 같이 구성하였다.

표 6. 단계별 feature set 구성

feature set	구성
feature set 1	기본 변수
feature set 2	파생변수 추가

각 feature set에 대해 10개 분류 알고리즘을 적용하여 분류 성능을 비교하였다. feature set 1에서는 Naive Bayes(ROC-AUC=0.704), Ada Boost(ROC-AUC=0.691), Gradient Boosting(ROC-AUC=0.679) 순으로 나타났으며, 파생변수를 포함한 feature set 2에는 Naive Bayes(ROC-AUC=0.710), Logistic(ROC-AUC=0.685),

Ada Boost(ROC-AUC=0.684)가 상대적으로 높은 성능을 보였다.

각 feature set ROC-AUC를 확인한 결과, 기본 변수 기반 모델의 ROC-AUC는 0.704, 파생변수를 포함한 모델은 0.710으로 유사한 수준의 구분력을 유지하였다. 이는 파생변수를 포함하더라도 전체 분류 성능은 비슷한 수준에서 유지되는 것을 보여준다. 또한 ROC-AUC는 전반적인 분류 성능을 평가하나, 본 연구 같은 Long 집단의 비율이 낮은 클래스 불균형 자료에서는 소수 집단에 대한 예측 성능을 충분하게 설명하기 어렵다. 이에 Long 집단을 positive class(1)로 설정하였으며, Precision, Recall 및 F1-score를 함께 산출하였고 그 결과는 <표 7>과 <표 8>에 제시하였다.

표 7. Feature set 1의 ROC-AUC 및 분류 성능 지표

Model	ROC-AUC	Precision_Long	Recall_Long	F1_Long
Logistic	0.664	0.272	0.622	0.378
KNN	0.590	0.254	0.600	0.310
DecisionTree	0.594	0.224	0.244	0.234
RandomForest	0.602	0.243	0.200	0.220
ExtraTrees	0.569	0.257	0.200	0.225
GradientBoosting	0.680	0.323	0.222	0.263
AdaBoost	0.691	0.328	0.489	0.393
XGBoost	0.626	0.282	0.244	0.262
SVM	0.671	0.283	0.578	0.380
NaiveBayes	0.705	0.284	0.644	0.395

표 8. Feature set 2의 ROC-AUC 및 분류 성능 지표

Model	ROC-AUC	Precision_Long	Recall_Long	F1_Long
Logistic	0.685	0.309	0.644	0.417
KNN	0.598	0.250	0.356	0.294
DecisionTree	0.611	0.240	0.267	0.253
RandomForest	0.609	0.225	0.200	0.212
ExtraTrees	0.572	0.229	0.178	0.200
GradientBoosting	0.671	0.333	0.267	0.296
AdaBoost	0.685	0.375	0.600	0.462
XGBoost	0.629	0.237	0.200	0.217
SVM	0.640	0.280	0.578	0.377
NaiveBayes	0.711	0.260	0.600	0.362

Feature set 1에서는 Precision이 0.224~0.328, Recall은 0.200~0.644 수준으로 나타났으며, Feature set 2에서는 Precision이 0.225~0.375, Recall은 0.178~0.644 수준으로 확인되었다. 두 feature set 모두 Long 집단에 대한 재현율은 일부 모델에서 상대적으로 높게 나타났으나 Precision은 전반적으로 제한적인 수준이다. 이는 소수 집단을 분석하는 과정에서 일부 Short 집단이 Long으로 분류되는 경우가 존재함을 의미하며, ROC-AUC 결과와 마찬가지로 전체 분류 성능 역시 제한적인 수준으로 해석할 수 있다.

다양한 분류모형을 비교한 결과 Naive Bayes가 가장 높은 ROC-AUC를 보였다. 다만, 본 연구에서는 다양한

모형의 예측 성능을 비교·확인하기 위한 목적으로 ROC-AUC를 산출하였으며, 이를 최종 분석 모형의 선정 기준으로 사용하지는 않았다. 본 연구의 목적은 최고 성능 모델을 선정하는 것이 아닌 변수 중요도와 경험 구조를 해석하는데 있기에 변수 중요도 산출과 변수 간 관계에 강점을 지닌 랜덤포레스트를 중심 분석 모형으로 사용하였다.

각 feature set의 랜덤포레스트 변수 중요도를 비교한 결과, 기본 변수 모델에서는 공연 횟수가 주요 변수로 나타난 반면, 파생변수를 포함한 모델에서는 Mobility_Challenge(이동 특성과 공연 경험 결합 변수)와 Preparation_Performance(입시 준비 및 공연 경험 결합 변수)가 높은 중요도로 나타났다(그림 3).

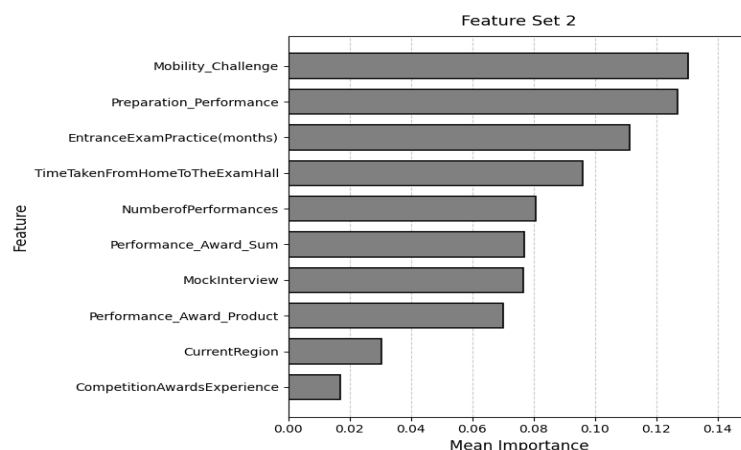


그림 3. 랜덤포레스트 기반 파생변수 포함 중요도 분석 결과

이는 파생변수가 포함된 모델에서는 변수 중요도가 변화하였다는 것을 보여준다. 특히 Mobility_Challenge와 Preparation_Performance가 높은 중요도를 보인 것은 무용 경력 수준 집단의 구분 과정에서 단순한 공연 참여 경험 횟수뿐만이 아닌 이동 특성과 공연 경험, 입시 준비 기간 등과 같은 경험 결합 구조가 함께 반영될 가능성을 보여준다. 특히 파생변수는 단일 경험 변수만으로 해석하기 어려운 입시 준비 경험의 복합적 특성을 반영한다는 점에서 의미가 있다.

이러한 결과는 변수 중요도가 단일 경험 변수 중심에서 경험 결합 구조 중심으로 확장되는 경향이 있음을 나타낸다.

3. Short 집단의 Long 집단 분류 가능성과 관련된 주요 경험 조건 탐색

Short 집단이 모델상 Long 집단으로 분류될 가능성과 관련된 경험 조건을 확인하기 위해 랜덤포레스트 기반 변수 중요도 분석 결과를 중심으로 분석하였다. 모든 변수에서 뚜렷한 차이가 나타난 것은 아니며, Mobility_Challenge, Preparation_Performance, NumberofPerformances, Performance_Award_Sum과 같은 일부 경험 변수에서만 상대적으로 차이가 확인되었다. 반면 단순 수상 경험과 모의면접 경험은 랜덤포레스트 중요도 및 통계검정에서는 무용 경력 수준 집단 구분과 관련된 변수로 나타났으나, Short 집단 내부에서 Long 집단으로 분류될 확률 변화 분석에서는 뚜렷한 관련이 확인되지 않았다. 이는 모의면접 경험이 전체 표본의 집단 구분 과정에서는 일정한 관련성을 보일 수 있으나, Short 집단 내부의 Long 집단 분류 확률 변화를 단독으로 관련된 변수로 해석하기에는 제한이 있음을 의미한다. 또한 공연 횟수와 수상 경험 곱으로 구성된 Performance_Award_Product와 입시 준비 기간도 Long 집단 분류 확률과 관련이 있으나, 그 관련성은 비교적 제한적인 것으로 나타났다(그림 4).

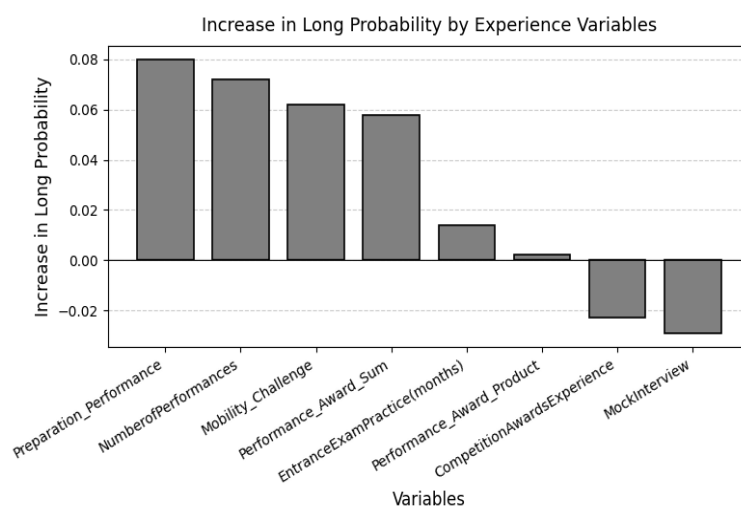


그림 4. Short 집단의 Long 집단 분류 확률 변화

상대적으로 높은 중요도를 보인 경험 기반 변수 중 실제 경험 특성과 관련된 변수들을 중심으로 구성하였다. 이에 따라 Preparation_Performance, NumberofPerformances, Mobility_Challenge, Performance_Award_Sum 등 주요 경험 변수를 중심으로 분석하였다(〈표 9〉).

표 9. 주요 경험 변수에 따른 Short 집단의 Long 집단 분류 확률(랜덤포레스트 기반 예측)

변수	평균 확률	임계값 적용 후 분류 확률	증가량
Preparation_Performance	0.107	0.187	+0.080
NumberofPerformances	0.107	0.179	+0.072
Mobility_Challenge	0.107	0.169	+0.062
Performance_Award_Sum	0.107	0.165	+0.058

분석 결과 Preparation_Performance와 NumberofPerformances가 Short 집단이 Long 집단으로 분류될 가능성과 상대적으로 높은 관련 경향을 보였다. Long 집단으로 분류될 예측 확률은 평균 약 0.107 수준이었으나 Preparation_Performance가 일정 수준 이상일 경우 약 0.187까지, NumberofPerformances가 높은 경우에도 약 0.179까지 높게 나타났다. 이는 공연 경험과 입시 준비 경험이 함께 축적된 경우 Long 집단과 유사한 변수 패턴을 보일 가능성이 높아질 수 있음을 시사한다.

Mobility_Challenge가 높은 경우 Long 집단으로 분류될 예측 확률은 약 0.169, Performance_Award_Sum이 높은 경우에는 약 0.165로 나타났다. 이러한 결과는 이동 특성과 공연 경험의 결합, 공연 참여와 수상 경험의 결합이 Short 집단 내부의 Long 집단 분류 확률과 일정한 관련성을 보였음을 의미한다. 따라서 Short 집단의 Long 집단 분류 가능성은 단일 경험 요인보다는 참여, 준비 및 성취 경험이 함께 나타나는 복합적인 경험 구조와 관련된 것으로 해석할 수 있다.

IV. 논의

본 연구는 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준과 관련된 입시 준비 행동 및 경험 구조를 머신러닝 기반 분석을 통해 탐색하였다. 연구 결과를 바탕으로 다음과 같이 논의하고자 한다.

첫째, 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준 집단 구분과 관련된 주요 변수는 공연 경험으로 나타났다. 기본 변수 모델에서 NumberofPerformances(공연 횟수)는 가장 높은 변수 중요도를 보였으며, Short 집단과 Long 집단 간에도 유의한 차이가 확인되었다. 이는 공연 경험이 입시 시점 무용 경력 수준과 함께 나타나는 주요 경험적 특성 중 하나임을 보여준다. 조유연(2025)은 공연 활동 참여 경험이 실용무용 전공자의 수행 경험 및 무용 지속 의향과 관련될 수 있다고 보고하였으며, 공연을 실제 수행과 참여 과정이 포함된 교육적 경험으로 설명하였다. 이러한 논의는 본 연구에서 공연 횟수가 주요 변수로 나타난 결과와 일부 맥을 같이한다. 또한 최용원(2020)은 실용무용 전공자의 지속적인 대회 참여가 학업 지속 의향과 정적 관계를 가진다고 보고하였다. 이를 고려하면 반복적인 공연 및 대회 참여 경험이 입시 시점까지 축적된 무용 활동 경험의 한 부분으로 반영되었을 가능성이 있다.

다만 공연 횟수와 대회 수상 경험은 무용 경력 기간과 완전히 독립된 변수로 보기 어렵다. 무용 경력 기간이 긴 지원자는 공연 참여 및 대회 출전 기회를 상대적으로 더 많이 경험했을 가능성이 있기 때문이다. 따라서 본 연구에서 나타난 공연 경험과 수상 경험의 중요도는 무용 경력 수준을 독립적으로 형성하는 원인이라기보다, 입시 시점까지 축적된 무용 활동 경험의 차이를 반영한 결과로 해석할 필요가 있다. 즉, 특정 경험이 무용 경력 수준을 직접 결정한다는 의미가 아니라, 서로 다른 무용 경력 수준 집단에서 함께 나타나는 경험적 패턴을 탐색적으로 확인한 결과이다.

실용무용 입시 맥락에서 공연 경험은 단순히 무대에 오른 횟수만을 의미하지 않는다. 공연 과정에서 지원자는 음악, 공간, 관객과의 관계 속에서 자신의 움직임을 조절하고 수행 상황에 적응하는 경험을 축적하게 된다. 선행연구에서도 무용참여기간과 공연 횟수가 자기관리, 완벽성향 및 공연몰입과 관련되는 것으로 보고되었다(노정식, 홍혜전, 2011). 또한 스트릿댄스의 배틀문화는 음악, 심사위원, 관객, 참여자가 상호작용하는 개방적 수행 구조를 가지며(이주영, 박현옥, 2017), 배틀 참여 상황에서는 불안, 자신감, 심리기술 및 환경적 변수가 함께 작용한다(최정은 외, 2015). 따라서 본 연구에서 공연 횟수와 수상 경험이 주요 관련 변수로 나타난 것은 단순한 활동량의 차이라기보다, 실용무용 입시 준비 과정에서 수행 경험, 평가 경험, 자기 점검 경험이 누적되는 구조를 반영한 결과로 해석할 수 있다.

둘째, 파생변수를 포함한 결과, 중요도는 단순한 경험의 양보다 경험이 쌓이는 방식에 더 민감하게 나타났다. 기본 변수로 구성된 feature set 1에서는 공연 횟수가 핵심 변수로 나타났지만, 파생변수를 포함한 feature set 2에서는 Mobility_Challenge와 Preparation_Performance가 높은 중요도를 보였다. 이는 무용 경력 수준 집단 구분이 공연 횟수만으로 설명되기보다, 이동 거리와 같은 제약 속에서 공연 경험과 준비 경험이 어떻게 결합되었는지와 더 관련이 있음을 시사한다. 특히 파생변수는 단일 경험 변수만으로 설명하기 어려운 입시 준비 경험의 결합 구조를 반영한다는 점에서 의미가 있다.

이러한 결과는 실용무용 입시 준비 과정이 단순한 경험의 양보다 공연 참여, 이동 특성, 준비 과정 등이 결합된 방식과 관련될 가능성을 시사한다. 다만 시험장 이동 시간은 입시 준비 환경의 차이를 일부 반영할 가능성이 있으나, 이를 지역적 접근성이나 교육 기회 격차로 직접 해석하는 데에는 주의가 필요하다. 시험장 이동 시간은 거주 지역, 교통 조건, 입시 당일 이동 방식, 개인적 일정 등 다양한 요인의 영향을 받을 수 있으며, 본 연구 자료만으로는 이동 시간이 학원 접근성, 입시 정보 접근성, 공연 참여 기회와 직접적으로 연결되는지를 확인하기 어렵다.

따라서 본 연구에서 나타난 시험장 이동 시간의 중요도는 입시 시점 무용 경력 수준 집단 구분 과정에서 관찰된 환경적 관련 변수로 제한하여 해석할 필요가 있다.

Nam, Park, & Kim(2024)은 무용수의 최상 수행이 체력, 기술, 심리와 같은 개인적 요인뿐 아니라 무대 환경과 관객 반응 등 환경적 요인을 함께 고려하는 과정에서 이루어지며, 준비 단계와 무대 직전 단계에서 서로 다른 수행 전략이 작용한다고 설명하였다. 이러한 관점에서 볼 때, 이동 특성과 공연 참여, 준비 과정이 결합된 파생변수가 높은 중요도를 보인 본 연구의 결과는 입시 준비 경험이 단일 요인의 합산이 아니라 요인 간 관계 구조 속에서 집단을 변별할 수 있음을 시사한다.

셋째, Short 집단이 Long 집단으로 분류될 확률의 평균은 약 0.107이었으나, Preparation_Performance는 0.187, NumberofPerformances는 0.179, Mobility_Challenge는 0.169, Performance_Award_Sum은 0.165 수준의 분류 확률이 확인되었다. 반면 모의면접 경험은 전체 표본을 대상으로 한 랜덤포레스트 변수 중요도와 통계검정에서는 집단 구분과 관련된 변수로 나타났으나, 3년 이하 경력 집단 내부의 분류 확률 변화 분석에서는 뚜렷한 관련이 확인되지 않았다. 이는 분석 단계에 따라 모의면접 경험이 갖는 의미가 다르게 나타날 수 있음을 보여준다. 즉, 모의면접 경험은 입시 준비 행동의 한 요소로서 전체 집단 구분에는 일정 부분 반영될 수 있으나, 3년 이하 경력 집단 중 4년 이상 경력 집단과 유사한 변수 패턴을 보이는 사례를 설명하는 데에는 공연 경험, 입시 준비 기간, 이동 특성 등이 결합된 변수에 비해 상대적으로 제한적인 역할을 한 것으로 해석할 수 있다. 이러한 결과는 무용 경력 수준 집단 구분이 특정 단일 요인이나 준비 경험에 의해 설명되기보다, 준비, 참여, 성취가 함께 축적되는 과정과 관련된 경향을 보여준다. 이는 무용참여기간과 공연 횟수가 많을수록 자기관리와 공연물입이 높게 나타나고(노정식, 홍혜진, 2011), 공연 횟수 및 수상 경력에 따라 몰입 경험에 차이가 있다는(신숙경, 2008) 선행연구의 맥락과 유사하며, 무용에서 다양한 수행 경험이 경력 수준에 따른 차이를 설명하는 데 관련된 가능성을 시사한다.

본 연구는 기본 변수, 파생변수를 단계적으로 구성하고, 머신러닝 분류 모형과 랜덤포레스트 기반 변수 중요도 분석을 활용하여 실용무용 실기전형 지원자의 집단 구분 구조를 탐색하였다. 특히 다양한 분류 모형을 비교한 결과, 파생변수를 포함한 모델(ROC-AUC=0.710)이 기본 변수 모델(ROC-AUC=0.704)과 유사하거나 높은 구분력을 보였다. 다만 두 모델 성능 차이는 크지 않으며, 파생변수가 집단 구분과 관련된 정보를 일부 추가로 반영할 가능성을 보여준다고 해석할 수 있다. 이는 단일 변수보다 복합적인 변수 구성이 집단 구분 구조를 이해하는 데 유용할 수 있음을 시사한다.

또한 본 연구에서는 경험 변수들이 단일 변수보다 변수 간 결합 구조에서 중요도가 변화하는 경향이 나타났으며, 이는 변수 간 상호작용 구조를 함께 분석할 수 있는 머신러닝 기반 접근의 활용 가능성을 보여준다. 특히 본 연구는 해당 기법을 예측 정확도 자체보다 변수 중요도와 분류 가능성 해석을 위한 설명적 도구로 활용하였다는 점에서 차별성을 지닌다. 이는 실용무용 전공자의 경험 특성을 단일 변인 중심으로 분석해 온 선행 연구의 한계를 보완하고, 복합적 경험 요인의 상호작용을 기반으로 탐색할 수 있는 새로운 분석 틀을 제시한다는 점에서 중요한 학문적 의의를 지닌다.

다만 본 연구의 결과는 수도권 소재 A 대학교 실용무용 실기전형 지원자 자료에 기초한 것이므로, 전체 실용무용 입시생 또는 모든 실용무용 전공 희망자의 일반적 특성으로 확대하여 해석하는 데에는 제한이 있다. 특히 본 연구의 표본은 특정 대학, 특정 전형, 특정 시점의 지원자 특성을 반영하고 있으므로, 분석 결과는 해당 전형 지원자 집단 내에서 관찰된 입시 시점 무용 경력 수준 관련 경험 구조로 이해할 필요가 있다. 따라서 본 연구에서 나타난 공연 경험, 이동 특성, 입시 준비 경험 등의 관련성은 실용무용 입시생 전체에 대한 보편적 경향이라기보다, 본 연구 자료의 맥락 안에서 해석되어야 한다.

V. 결론 및 제언

본 연구는 실용무용 실기전형 지원자의 입시 시점 무용 경력 수준 집단 구분과 관련된 입시 준비 행동 및 경험 구조를 머신러닝 기반 분석을 통해 탐색하였다. 이를 위해 수도권 A 대학교 실용무용 실기전형 지원자 807명의 비식별 설문자료를 활용하였으며, 입시 시점 무용 경력 기간을 기준으로 3년 이하 경력 집단과 4년 이상 경력 집단을 구분하여 분석하였다.

입시 시점 무용 경력 수준 집단 구분 과정에서는 공연 횟수, 시험장 이동 시간, 모의면접 경험, 입시 준비 기간, 대회 수상 경험, 거주 지역 등이 주요 관련 변수로 나타났다. 특히 기본 변수 분석에서는 공연 횟수가 가장 높은 중요도를 보였으며, 파생변수를 포함한 분석에서는 Mobility_Challenge와 Preparation_Performance가 주요 변수로 확인되었다. 이는 실용무용 실기전형 지원자의 무용 경력 수준 집단 구분이 단일 요인에 의해 설명되기보다, 공연 참여, 입시 준비 기간, 이동 특성, 성취 경험 등이 결합된 구조 속에서 나타날 수 있음을 보여준다.

이러한 결과는 실용무용 입시 준비 과정을 단순한 시험 대비 과정으로만 보기보다, 전공 진입 이전 단계에서 학습자들이 실기 경험과 준비 행동을 어떠한 방식으로 축적해 왔는지를 함께 살펴볼 필요가 있음을 시사한다. 다만 공연 횟수, 수상 경험, 입시 준비 기간은 무용 경력 기간과 구조적으로 관련될 수 있으므로, 이를 무용 경력 수준을 독립적으로 설명하는 원인 변수로 해석하는 데에는 주의가 필요하다. 본 연구의 결과는 특정 변수가 무용 경력 수준을 직접적으로 형성한다는 의미가 아니라, 입시 시점 무용 경력 수준과 함께 나타나는 경험적 패턴을 탐색적으로 확인한 결과로 이해되어야 한다.

본 연구는 실용무용 실기전형 지원자의 입시 준비 경험 구조를 다양한 경험이 축적되는 방식에 주목하여 분석하였다는 점에서 의의가 있다. 또한 머신러닝 기법을 활용하여 여러 변수가 함께 나타나는 양상을 탐색함으로써, 기존의 단일 변인 중심 접근을 보완할 수 있는 기초자료를 제시하였다.

이상의 결과를 바탕으로 본 연구의 제언은 다음과 같다.

첫째, 실용무용 입시 지도에서는 실기 능력이나 성취 결과만을 중심으로 지원자를 이해하기보다, 공연 참여와 입시 준비 경험이 어떠한 방식으로 축적되었는지를 함께 살펴볼 필요가 있다. 이는 지원자의 경험 특성과 준비 과정을 이해하기 위한 참고자료로 활용될 수 있다.

둘째, 시험장 이동 시간과 거주 지역 등 환경 관련 변수가 집단 구분 과정에서 일정한 관련성을 보인 만큼, 교육 현장에서는 학생들이 처한 이동 및 참여 조건을 함께 파악할 필요가 있다. 다만 본 연구의 시험장 이동 시간만으로 교육 접근성이나 참여 제약을 직접 판단하기 어려우므로, 향후에는 실제 교육기관 접근성, 이동 빈도, 교통 여건 및 지역별 참여 기회 등을 포함한 구체적인 조사가 이루어질 필요가 있다.

셋째, 본 연구에서 활용한 파생변수는 입시 준비 경험 간의 결합 구조를 탐색하기 위해서 연구자가 구성한 변수이다. 따라서 향후 연구에서는 다양한 변수 산출 방식을 적용하여 민감도 분석과 Permutation importance 등을 활용하여, 결과의 안정성과 일반화 가능성을 탐색할 필요가 있다.

본 연구는 다음과 같은 제한점을 지닌다.

첫째, 본 연구 대상은 특정 대학 실기전형 지원자를 대상으로 한 데이터이므로, 해당 대학의 전형 특성과 입시 준비 기간의 연습량 및 참여 특성이 반영되었을 가능성이 있다. 이에 본 연구 결과를 전체 실용무용 입시생 또는 모든 실용무용 전공자의 경험 구조로 확대하여 해석하는 데에는 한계가 있다.

둘째, 입시 시점에 수집된 관찰자료를 활용한 탐색적 연구이므로, 공연 횟수, 수상 경험, 입시 준비 기간 등 주요 변수와 무용 경력 수준 간 관계를 인과적으로 해석하는 데에는 한계가 있다.

셋째, 양적 자료를 중심으로 집단 구분 구조를 분석하였기 때문에 개별 지원자의 입시 준비 경험이 형성된 구체

적인 과정과 의미를 심층적으로 설명하는 데에는 제약이 있다.

넷째, 클래스 불균형을 보정하기 위해 SMOTE를 적용하였으나, 합성 자료가 변수 중요도와 분류 결과에 일부 영향을 미쳤을 가능성을 완전히 배제하기 어렵다.

다섯째, 본 연구에서는 여러 변수에 대해 통계검정을 수행하였으나, 일부 결과는 신중하게 해석할 필요가 있다. 향후 연구 다양한 검증 방법을 통해 결과의 안정성을 검증할 필요가 있다.

따라서 후속 연구에서는 다양한 대학과 전형, 지역의 실용무용 입시 지원자 자료를 포함하여 연구대상의 범위를 확장할 필요가 있다. 아울러 면담, 사례연구 등의 질적 접근을 병행하여 실용무용 실기전형 지원자의 입시 준비 경험이 형성되는 교육적·환경적 맥락을 구체적으로 탐색하고, 다양한 자료와 분석 방법을 활용한 반복 검증을 통해 결과의 안정성을 확인할 필요가 있다. 이를 통해 실용무용 입시 지원자의 경험 구조를 보다 체계적으로 이해하고, 교육 현장의 입시 지도를 지원할 수 있는 기초자료를 축적할 수 있을 것이다.

참고문헌

- 고정민, 안병순(2025). 실용무용 전공 대학생의 진로 탐색, **영남춤학회지**, 13(3), 25-55.
- 김나현(2023). **실용무용을 통한 순수무용의 대중화발전방안**, 한양대학교 대학원, 석사학위논문. 서울.
- 김세준(2024). **머신러닝을 활용한 조종사 공간정위상실(SD) 회복력 비교분석에 관한 연구 : 조종사 비행 경력, 자격 및 재교육 여부를 중심으로**, 한서대학교 대학원, 박사학위논문. 충청남도.
- 김수산나, 안귀호(2023). 교수 감성리더십이 실용무용전공 학생의 몰입 및 학업지속에 미치는 영향, **무용예술학연구**, 91(2), 21-36.
- 김은영(2025). 실용무용 전공자가 현대무용 수업 만족도에 따른 잠재적 태도와 진로교육에 미치는 영향, **대한무용학회논문집**, 83(4), 117-135.
- 김정숙(2008). 무용수의 신체역학 프로파일 분석, **무용예술학연구**, 23, 1-12.
- 김정숙(2009). 무용수가 인지하는 객체화 신체역학 분석, **무용예술학연구**, 26, 51-63.
- 노정식, 홍혜전 (2011). 무용전공 대학생의 자기관리가 완벽성향과 공연몰입에 미치는 영향, **한국무용과학회지**, 24, 1-17.
- 문다예(2022). **실용무용전공자의 무용몰입경험이 무용능력성취 및 무용지속의도에 미치는 영향**, 조선대학교 교육대학원, 석사학위논문. 광주.
- 문재웅(2022). **머신러닝을 활용한 악성코드 이미지 분류 기법 및 탐지에 관한 연구**, 세종대학교 대학원, 박사학위논문. 서울.
- 박수현(2021). **실용무용전공생의 자아존중감이 사회성과 학교생활만족도에 미치는 영향**, 광주교육대학교 교육대학원, 석사학위논문. 광주.
- 박영하(2015). 실용무용의 흐름과 경향 분석, **무용예술학연구**, 52(1), 35-52.
- 박한솔(2023). 실용무용 전공자의 중도포기 문제에 대한 고찰, **한국실용무용학회지**, 1(2), 3-15.
- 백세룡(2022). **머신러닝을 이용한 후방추돌시 탑승자의 경추상해 예측모델 개발**, 가천대학교 대학원, 박사학위논문. 경기도.
- 신숙경 (2008). **무용전공자의 몰입경험이 공연만족도에 미치는 영향에 관한 조사 연구**. 이화여자대학교 대학원. 석사학위논문. 서울.
- 이보람, 권소영(2016). 무용전공학생의 진로결정수준이 진로결정 자기효능감에 미치는 영향 - 진로교육의 매개효과를 중심으로 -, **한국무용연구**, 34(4), 175-198.
- 이주영, 박현옥(2017). 스트리트 댄스의 배틀문화적 특성, **한국무용교육학회지**, 28(2), 157-169
- 이준모(2019). 현대무용 전공생의 전공만족, 대학생활적응, 진로준비행동의 관계, **한국무용과학회지**, 36(3), 1-14.
- 임승준(2023). **머신러닝을 활용한 자국 브랜드의 중고차 가격 예측 모델에 관한 연구 : 국내 중고차 시장을 중심으로**, 홍익대학교 대학원, 박사학위논문. 서울.
- 정한결(2020). 실용무용교육의 재개념화를 위한 탐색적 연구 : 실용무용의 연구동향과 실용음악의 사례를 통해, **한국무용학회지**, 20(2), 69-82.
- 정윤영(2015). 무용공연체형 프로그램이 보호관찰 청소년들의 자기효능감과 사회적 적응에 미치는 영향, **한국무용과학회지**, 32(1),

151-164.

- 조유연(2025). **무용 전공대학생의 공연 활동 참여가 무용 지속성 및 재미에 미치는 영향**, 숙명여자대학교 교육대학원, 석사학위논문. 서울.
- 조준희(2025). 춤의 경계를 넘어서: 실용무용 전공자의 순수무용 경험과 해석, **한국무용과학회지**, 42(3), 81-107.
- 최용원(2020). 평생교육원 실용무용 교육서비스품질과 진로교육이 전공자의 학업지속의향, 지속적 대회참여 미치는 영향, **한국체육학회지**, 59(6), 433-448.
- 최정은, 박희승, 권순민, 황성우(2015). 스트리트 댄서의 성격과 배틀 참여 시 발생하는 불안 및 심리기술 간의 관계, **코칭능력개발지**, 17(4), 73-83.
- Domingos, P. (2012). A Few Useful Things to Know About Machine Learning. *Communications of the ACM*, 55(10), 78-87.
- Halat, D. H., Abdel-Salam, A. G., Bensaid, A., Soltani, A., Alsarraj, L., Dalli, R., & Malki, A. (2023). Use of machine learning to assess factors affecting progression, retention, and graduation in first-year health professions students in Qatar: A longitudinal study. *BMC Medical Education*, 23(1), 909.
- Nam, S. M., Park, H. Y., & Kim, M. J. (2024). Exploring the experiences of dancers who have achieved peak performance: On-stage, pre-stage, and post-stage. *Frontiers in Psychology*, 15, 1392242.

ABSTRACT

Exploring Experiential Factors Related to Dance Experience Levels among Practical Dance Admission Applicants: A Machine Learning-Based Study

Hyein Kim* · Taehoon Kang** Sejong University

This study aimed to explore admission preparation behaviors and experiential factors that distinguish dance experience levels at the time of entrance examinations among applicants for a practical dance performance-based admission screening. Survey data from 807 applicants ($N = 807$) to the practical dance program at University A in the Seoul metropolitan area, collected between September and December 2025, were used as secondary data. Based on career classification criteria suggested in previous studies, applicants were categorized into a Short group (three years or less of dance experience) and a Long group (four years or more of dance experience). Prior to analysis, data preprocessing procedures were conducted, including the removal of unnecessary variables, reconstruction of family composition variables, generation of derived variables, one-hot encoding of categorical variables, and standardization of numerical variables. To prevent data leakage, the dance experience duration variable used for group classification was excluded from the predictor variables. The dataset was divided into training and validation sets using a 7:3 stratified split, and SMOTE was applied only to the training data to address class imbalance. Machine learning-based supervised learning methods were employed, and Random Forest feature importance analysis was conducted alongside the Mann-Whitney U test and chi-square test. The results indicated that the number of performances, travel time to the examination site, mock interview experience, entrance examination preparation period, and competition award experience were major variables distinguishing dance experience level groups at the time of entrance examinations. Furthermore, analyses including derived variables identified Mobility_Challenge and Preparation_Performance as important variables, suggesting that group classification may be associated with combined structures of performance participation, mobility characteristics, and admission preparation experiences rather than with single experiences alone. As this study was an exploratory analysis based on applicants to a single university, caution is required when generalizing the findings to all practical dance majors, and the results should be interpreted as associations rather than causal relationships. Nevertheless, this study is meaningful in that it explored the structure of admission preparation experiences prior to entering a practical dance major using a machine learning-based approach.

Key words : Practical Dance, Dance Experience Level, Admission Preparation Experience, Machine Learning, Random Forest, Exploratory Data Analysis (EDA)

논문투고일: 2026.05.30

논문심사일: 2026.07.06

심사완료일: 2026.07.23

* Ph.D. Student, Department of Dance, Sejong University, First Author

** Ph.D. Student, Department of Dance, Sejong University, Corresponding Author