



## 머신러닝 기법을 적용한 기업 구성원의 직무만족 및 조직몰입에 영향을 미치는 요인 탐색

박 소 영\*

### 요약

본 연구는 한국직업능력개발원의 인적자본기업패널(HCCP) 7차년도 데이터를 활용하여 기업 구성원의 직무만족 및 조직몰입에 영향을 미치는 요인을 탐색하는 데 목적이 있다. 이를 위해 본 연구는 잠재프로파일 분석을 통해 기업 구성원의 직무만족 및 조직몰입을 각각 이질적인 집단으로 구분하였으며, 직무만족 잠재프로파일 및 조직몰입 잠재프로파일 예측을 위해 LASSO, Ridge, Randomforest, Support Vector Machine 등 4가지 머신러닝 기법을 적용하였다. 분석결과, 첫째, 직무만족과 조직몰입은 각각 3개의 잠재프로파일로 구분되었다. 직무만족의 경우 각 잠재프로파일의 특징에 따라 '중간 직무만족 집단(계층 1)', '낮은 직무만족 집단(계층 2)', '높은 직무만족 집단(계층 3)'으로 명명하였다. 조직몰입도 '낮은 조직몰입 집단(계층 1)', '높은 조직몰입 집단(계층 2)', '중간 조직몰입 집단(계층 3)'으로 구분되었다. 둘째, 직무만족과 조직몰입 모두 Randomforest의 예측정확도가 각각 70.1%와 69.3%로 가장 높았다. 셋째, Randomforest로 직무만족과 조직몰입에 영향을 미치는 주요 변수를 탐색한 결과, 직무만족의 경우 근로자의 현재 상태 및 근무 상황, 기업과 근로자의 커뮤니케이션 및 신뢰관계, 스트레스, 기업문화, 인재우대, HR 부서활동, 팀 밖 회사차원의 공식훈련 등의 변수가 도출되었다. 조직몰입은 기업과 근로자의 커뮤니케이션 및 신뢰관계, 근로자의 현재 상태 및 근무 상황, 스트레스, 인재우대, 기업문화, HR 부서활동 등의 변수가 선택되었다. 마지막으로 본 연구결과를 바탕으로 직무만족 및 조직몰입 향상을 위한 다양한 시사점을 제공하였다.

### I 서론

기업의 재무적인 성과와 직무만족 및 조직몰입과의 관계에 대해서는 많은 연구에서 주목해왔으며, 직무만족 및 조직몰입이 기업의 성과뿐만 아니라 이직의도, 이직 감소에 까지 긍정적인 영향을 미칠 수 있다고 밝혀짐에 따라 기업 구성원의 직무만족 및 조직몰입을 높이기 위한 연구가 다각도로 진행되어 왔다(오인수, 김광현, Dornold, 황종오, 유

\* 충남대학교 교육학과 박사과정

태용, 박영아, 박량희, 2007). 심지어 직무만족 및 조직몰입을 기업에서 중요하게 달성해야 할 기업 성과의 일종으로 간주하여 매출액, 생산성 등의 재무적인 측면의 성과와 함께 직무만족, 조직몰입, 조직 스트레스 등을 비재무적인 측면의 성과로 보고 재무 및 비재무적 성과를 함께 달성하기 위한 기업의 노력이 촉구되고 있다(전현정, 정혜원, 2017).

선행연구에서는 성별, 연령, 최종학력 등 개인적인 요인(박성계, 2000; 오영미, 2004), 임금만족이나 임금(황매향, 최희철, 조흥순, 2013), 기업문화 혹은 조직문화(이용탁, 2017), 스트레스(이환범, 이수창, 2006), 교육훈련(전현정, 정혜원, 2017), 커뮤니케이션 및 신뢰관계(최영준, 심원술, 백승현, 2014), 공정성(민경호, 조국행, 2002), 기업에서의 인재우대(박진수, 이성희, 2019), HR 부서의 역할 및 활동(우성미, 최명빈, 2018), 전공 계열(문찬주, 남인혜, 이영선, 이지영, 한수현, 정동욱, 2019) 등 특정 변수가 직무만족 및 조직몰입에 어떠한 영향을 미치는지를 검증해왔다. 그러나 대부분의 선행연구들이 이론적 혹은 이전의 연구에 기반해 일부 변수만을 포함한 모형을 설정했으며, 또한 활용할 수 있는 변수의 갯수가 제한적인 구조방정식 모형, 회귀모형, 다층모형 등을 적용했다는 특징이 있다. 이러한 모형들은 통제변수와 같은 변수를 제대로 포함하는 것이 중요하며 어떠한 변수를 모형에 포함하는지에 따라 전혀 다른 결과가 도출될 수 있고, 많은 변수를 투입하게 되면 불안정한 모형이 될 수 있다는 한계가 있다(유진은, 노민정, 2017). 즉, 대부분의 연구들이 설정하고자 하는 연구목적에 맞춰 모형을 설정하기 때문에 결과를 일반화하는 데 어려움이 있다(최영준 외, 2014).

이처럼 일부 연구자의 관심 변수만으로 연구 모형 설정하거나 연구 방법 등의 한계로 인해 거시적인 연구가 불가능했던 것과는 달리, 최근 여러 분야에서는 빅데이터 분석을 위한 머신러닝(machine learning) 기법을 적용하여 가능한 모든 자료를 모형에 함께 투입할 수 있게 되었다. 특히 경영학, 교육학 등 사회과학분야에서는 인 한국직업능력개발원의 인적자본기업패널(HCCP), 한국교육고용패널조사(KEEP)와 같이 여러 연구기관에서 대규모 패널자료를 구축하고 있기 때문에 방대한 자료들을 최대한 활용하기 위한 방안으로 머신러닝 기법을 적용한 연구가 최근 시도되고 있으나 아직 미비한 실정이다(유진은, 2016). 머신러닝 기법을 적용한 연구에서는 이론적 배경이 아닌 데이터에 기반한 모형이 구축되기 때문에 기존에서 강조되어온 변수뿐만 아니라 기존에 고려하지 못했던 변수가 도출되어 새로운 시사점을 제공해 줄 수 있다. 이에 본 연구에서는 기업 구성원의 직무만족 및 조직몰입에 영향을 미치는 변수를 탐색하는 데 있어 주어진 패널자료를

최대한 활용할 수 있는 머신러닝 기법들을 적용하여 설명변수 중 가장 중요한 변수를 탐색하고 또한 직무만족 및 조직몰입의 연구가 활발히 진행되어 왔음에도 밝혀지지 않은 새로운 변수를 도출하고자 한다.

머신러닝 기법을 적용해 기업 구성원의 직무만족 및 조직몰입 영향요인을 탐색한 연구가 제한적으로 이루어졌다. 김정수와 김종우(2008)는 머신러닝 기법 중 의사결정나무 분석을 사용해 직무만족과 조직몰입에 영향을 미치는 요인을 탐색하였으며, 2005년에 수집된 인적자본기업패널 1차년도 자료를 활용하였다. 최영준 외(2014)도 마찬가지로 인적자본기업패널 자료를 사용했으며, 의사결정나무 분석을 보완한 CART 앙상블 방법인 Bagging과 Arcing 알고리즘을 적용하여 조직몰입에 영향을 미치는 주요 설명변수를 확인하였다. 그러나 최근에는 훨씬 다양한 머신러닝 기법들이 개별적으로 진행되고 있으며, 의사결정나무 분석보다 예측정확도가 높은 기법들을 활용하고 있다. 이에 본 연구에서는 첫째, 가장 최신자료인 인적자본기업패널의 7차년도 자료를 활용하고, 둘째, 선행연구(김정수, 김종우, 2003; 최영준 외, 2014)에서는 기법 간의 성능 차이를 비교하는 게 목적인 연구를 제외하고 결과변수의 영향요인을 탐색하기 위한 연구의 경우 한 두 개의 머신러닝 기법만을 적용하였지만, 본 연구에서는 4가지 머신러닝 기법을 비교하여 가장 예측정확도가 좋은 기법을 적용해 직무만족 및 조직몰입의 영향요인을 탐색하고자 한다.

이에 분류분석이 가능하며 기업 구성원의 직무만족 및 조직몰입에 영향을 미치는 설명변수를 선택하기 위해 대표적인 지도학습 알고리즘인 LASSO(Least Absolute Shrinkage and Selection Operator) regression, Ridge regression(능형회귀), Randomforest(랜덤 포레스트), SVM(Support Vector Machine 이하 서포트 벡터 머신) 등 4가지 머신러닝 기법을 적용하고자 한다. LASSO와 Ridge는 벌점화 회귀모형으로, 일반 회귀모형에 벌점(penalty)함수를 더하여 회귀계수를 변화시킨다. LASSO는 영향력이 없거나 혹은 낮은 설명변수를 0으로 만들어 변수를 선택할 수 있는 반면, Ridge는 영향력이 없거나 혹은 낮은 설명변수의 회귀계수를 축소해 예측오차를 낮추긴 하지만 모든 변수의 회귀계수를 추정하여 변수를 선택할 수 없다는 차이가 있다(노민정, 2018; 박민수, 김태현, 조은석, 김희발, 오희석, 2014; 송상윤, 2015). LASSO와 Ridge는 설명변수의 회귀계수를 직접적으로 해석가능하다는 장점이 있다. 즉, 남자에 비해 여자의 직무만족이 높다 혹은 남자에 비해 여자의 조직몰입이 낮다와 같은 회귀계수 부호를 확인할 수 있다.

Randomforest은 의사결정나무 분석을 확장한 앙상블 기법으로 설명변수가 많을 경

우에도 예측력이 높으며, 모형이 안정적이라는 장점이 있어 분류분석 시 자주 활용되는 기법이다(유진은, 2015). 그러나 중요도 지수 등을 통해 결과를 해석해야 하므로 LASSO와 Ridge와 같이 회귀계수 부호를 직접적으로 확인할 수 없어 결과에 대한 해석에 어려움이 있다. 사회과학분야 보다는 정보, 컴퓨터 분야 등에서 많이 활용되어 온 SVM은 Randomforest와 같이 예측력이 높으며, 과적합 등의 문제에서 자유로울 수 있으며 일반화 가능성이 높다는 장점이 있으나(오미애, 최현수, 김수현, 장준혁, 진재현, 천미경, 2017), 해석이 어렵고 학습과정에서 많은 시간이 소요된다(제홍모, 방승양, 2003). 이처럼 머신러닝 기법에 따라 장단점이 있으며, 자료 특성에 따라 머신러닝 기법의 성과가 다를 수 있기 때문에 여러 가지 머신러닝 기법을 적용해 본 자료를 가장 잘 분류하는 기법을 탐색할 필요가 있다.

또한 회귀분석이 아닌 분류분석을 위해 머신러닝 기법을 적용하는 경우 대다수의 선행연구에서 모형의 간명성, 분석시간 단축 등을 위해 결과변수를 두 개의 집단으로 구분하게 된다. 참여 혹은 미참여 등과 같이 이분형인 경우를 제외하고 연속형 결과변수를 두 개의 집단으로 구분할 때, 선행연구에서는 중앙값, 평균값 등 연구자가 임의적으로 설정한 기준에 따라 집단을 구분한다. 그러나 본 연구에서는 기업 구성원의 직무만족 및 조직몰입을 잠재프로파일 분석을 활용해 이질적인 잠재집단으로 구분하고자 한다. 잠재프로파일 분석을 활용하게 되면 기업 구성원의 직무만족 및 조직몰입 수준에서 존재하는 비관측된 이질적인 특징에 따라 집단을 구분하게 되고, 집단을 구분하는 데 있어 객관적인 모형 적합도를 고려할 수 있다. 특히 직무만족이나 조직몰입은 기업 구성원 간의 편차가 존재하기 때문에, 하나의 집단으로만 간주한다면 기업 구성원 간의 차이를 밝힐 수 없다. 따라서 직무만족 및 조직몰입 수준에 따라 이질적인 잠재프로파일로 구분하여 집단별 특징을 고려한 차별적인 비재무적 성과 향상 방안을 강구할 필요가 있다.

본 연구는 한국직업능력개발원의 인적자본기업패널(Human Capital Corporate Panel 이하 HCCP) 7차년도 자료를 활용해 기업 구성원의 직무만족 및 조직몰입을 이질적인 잠재프로파일로 구분하고자 한다. 다음으로 4가지 머신러닝 기법인 LASSO, Ridge, Randomforest, SVM을 적용해서 다양한 설명변수로 도출된 이질적인 잠재프로파일을 가장 잘 분류·예측하는 머신러닝 기법 선택한다. 마지막으로 최종적으로 선택된 기법을 통해 직무만족 및 조직몰입 잠재프로파일에 주요한 영향을 미치는 변수를 탐색하고자 한다.

## II 선행연구 고찰

### 1. 직무만족 영향요인

직무만족(job satisfaction)은 조직의 성과, 목표달성, 성공 등을 위한 중요 요소로, 기업 구성원이 현재 자신이 하고 있는 일, 임금, 직무 환경, 직무 조건 등에 대한 만족스러운 정도라고 할 수 있다(윤장열, 2010). 직무만족은 크게 직무에 대한 전반적인 만족감을 연구하는 전반적인 접근과 인간관계, 임금, 복리후생 등 직무 환경이나 조건과 관련된 단면적 접근으로 나눌 수 있는데(Spector, 1997; 이용탁, 2017, 재인용). 본 연구의 경우 현재 일, 현재 받는 임금, 현 직장의 인간관계에 대한 만족을 묻는 단면적 접근이라고 할 수 있다. 기업 구성원의 직무만족 수준을 높이기 위한 다양한 연구가 수행되어 오고 있는 만큼 연구자들은 인적자본기업패널 외에도 한국고용정보원의 고령화연구패널조사(KLoSA) 등 패널자료 혹은 연구자의 설문조사 등 여러 자료가 활용되고 있기에, 인적자본기업패널을 활용한 논문에 초점을 맞춰 선행연구를 고찰하고자 한다.

박광오, 이은령, 정대현(2018)은 기업 구성원 간의 신뢰관계가 형성되어 있으며, 커뮤니케이션이 활발하게 유지되는 것들이 직무만족에 긍정적인 영향을 미친다고 하였다. 또한 임금은 직무만족에 정적으로 유의미한 영향을 미쳤다. 전현정과 정혜원(2017)의 연구에서 교육훈련 참여는 오히려 직무만족에 부적인 영향을 미치는 것으로 나타났으며, 교육훈련에 대한 관심, 직무연관성 등 교육훈련의 내용이 직무만족에 정적인 영향을 미치는 것으로 나타났다. 성별, 직급, 최종학력에 따른 직무만족의 차이는 통계적으로 유의하지 않았다. 구조방정식을 이용한 이용탁(2017)의 연구에서는 기업문화 중 혁신문화, 관계문화, 위계문화, 합리문화가 직무만족에 정적으로 유의한 영향을 미치는 것으로 나타났다. 그러나 관료적 성격이 강한 위계적인 기업문화가 직무만족에 미치는 영향은 일관적이지 않게 보고되었는데, 위계적인 기업문화가 기업 구성원의 직무만족에 부적인 영향을 미치거나 혹은 유의한 영향을 미치지 않는다는 결과와(주영하, 선민정, 2018), 이용탁(2017)의 연구결과와 같이 직무만족에 긍정적인 영향을 미친다는 연구결과가 혼재되어 있다. 또한 이용탁(2007)의 연구에서는 다양한 HRD 제도(선·후배 및 동료간 상호작용을 통한 학습, 일을 통해 스스로 배우기, 멘토링 또는 코칭, 학습조직, 지식 마일리지 프로그램, 6-시그마 등) 참여가 직무능력에 미치는 효과는 직무만족에 정적인 영향을 미치는 것으로 나타났다. 김보인(2017)의 연구에서는 기업 규모와 관계없이 근무기

간, 커뮤니케이션과 신뢰관계, 교육훈련 다양성 등이 직무만족에 중요한 영향을 미치는 것으로 나타났다. 즉, 근무기간이 길수록, 회사 경영진과의 커뮤니케이션이 활발하고 신뢰가 두터울수록, 교육훈련에 다양하게 참여할수록 직무만족 높은 것으로 나타났다. 특히 대기업의 경우에는 직급, 최종학력, 노조가입 여부도 유의한 영향을 미쳤는데, 직급이 높고, 최종학력이 높고, 노조에 가입한 경우 직무만족이 높았다.

## 2. 조직몰입 영향요인

많은 연구에서 조직몰입과 직무만족을 함께 연구하고 있으며, 두 변인 간 상관성이 높다고 보지만 조직몰입과 직무만족은 구분되는 개념이다. 간단히 말하면, 직무만족은 기업 구성원이 현재 하고 있는 직무 혹은 그와 관련된 만족이나, 조직몰입은 기업 구성원이 속한 조직에 대한 태도나 신뢰라고 할 수 있다(이영면, 백승규, 2010). 조직몰입(organizational commitment)은 자신이 속한 조직의 비전 혹은 목표에 대해 신뢰를 가지며 조직에 대한 애착이 있는 상태로 조직의 구성원으로 현 조직에 대한 소속감, 귀속감이 강한 상태를 말한다(Mowday, Porter, Steers, 1982).

인적자본기업패널을 활용해 조직몰입에 대한 다양한 설명변인의 영향력을 검증한 연구가 다수 있었다. 전현정과 정혜원(2017)은 교육훈련의 내용과 관심이 조직몰입에 중요한 변인임을 밝혔으며, 여자에 비해 남자 근로자의 조직몰입이 높았으며, 직급이 높을수록 조직몰입이 높은 것으로 나타났다. 그러나 최종학력의 경우 조직몰입에 부적인 영향을 미쳤으며, 최종학력이 높을수록 조직몰입이 낮은 것으로 나타났다. 배성은과 이견만(2014)도 교육훈련의 내용이 직무와 연관성이 높을 때 조직몰입을 높일 수 있다고 하였다. 최영준, 심원술, 백승현(2014)의 연구에서는 신뢰관계와 커뮤니케이션, 혁신, 근속연수 등이 중요한 설명변인으로 도출되었다. 이용탁(2017)의 연구에서는 구조방정식을 사용해 기업문화 중 혁신문화, 관계문화, 합리문화가 조직몰입에 미치는 긍정적인 영향을 밝혀냈다. 다만 위계적인 기업문화는 조직몰입에 유의한 영향을 미치지 못하였다. 장혜정, 진인선(2014)은 인재우대와 교육훈련이 조직몰입에 정적인 영향을 미친다고 하였는데, 기업 내에서 인재가 우대받는 분위기가 조성될수록, 교육훈련이 기회가 균등하고, 선호하는 내용, 실무와 관련성 등 적절하게 이루어질 때 조직몰입이 높아진다고 하였다. 또한 이용탁(2007)의 연구에서는 다양한 HRD 제도(선·후배 및 동료간 상호작용을 통한 학습, 일을 통해 스스로 배우기, 멘토링 또는 코칭, 학습조직, 지식 마일리지 프로그램, 6-시그마 등)가 직무능력에 미치는 효과는 조직몰입에 정적인 영향을 미치는 것으로 나타났다.

### III 연구 방법

#### 1. 연구 대상

본 연구는 한국직업능력개발원에서 2005년부터 기업과 기업의 구성원을 2년 주기로 추적·조사한 인적자본기업패널(HCCP)에서 가장 최근에 조사된 7차년도(2017) 데이터를 활용하였다. 인적자본기업패널은 기업(본사) 조사와 기업의 구성원에 대한 조사로 구성되며, 본 연구에서는 기업 구성원을 조사한 자료를 활용하였다. 분석에 활용된 총 기업 구성원 수는 총 10,005명(남자 7,892명, 여자 2,113명)이다.

#### 2. 변수 설명

##### 가. 데이터 정리

본 연구에서는 기업 구성원의 직무만족 및 조직몰입에 영향을 미치는 변수를 탐색하기 이전에 모형에 투입할 수 없는 변인을 정리하였다. 우선 기업 ID, 사업장 ID, 팀(부서/반) ID, 팀원(응답자) ID 등 분석에 활용할 수 없는 ID 변수들을 삭제하였다. 둘째, 시기를 응답한 변수인 현직장 입사시기 년(W701\_01), 월(W701\_02), 현재직급 근무시기 월(W702\_01), 현재 업무 시작시기 년(W705\_01), 월(W705\_02), 출생년도 월(W7\_birthm)은 의미 있는 설명변수로 보기 어려워 제외하였으나, 단 현재직급 근무시기가 얼마나 되었는지 월이 아닌 년을 묻는 문항과 출생년도는 분석에 포함하였다. 셋째, 해당되는 코드를 입력해야하는 변수인 현재 업무(W704), 자격증명 코드(W719\_01, W719\_02, W719\_03), 취득년도(W719\_04)도 분석에서 제외하였다. 넷째, 범주형 변수 중 더미코딩으로 구분이 어려운 직무 관련성이 가장 높은 자격증의 종류(W719\_05), 직무 관련성이 가장 높은 자격증의 활용분야(W719\_05) 변수도 제외하였다. 다섯째, 분야별 인력 수준(W709\_01~W709\_08) 변수의 응답 항목 중 8(해당 인력 없음)과 회사의 교육훈련 참여 과정(W715)과 전공계열(W7\_affil)의 응답 항목 중 5(기타)를 응답한 사례는 결측으로 재코딩하였다. 마지막으로 본 연구에서는 다중대체법(multiple imputation)을 사용하여 결측치를 대체하였으나 응답률을 확인하여 결측률이 35%가 넘는 변수는 삭제한 후 대체를 실시하였다. 이에 80개의 변수가 삭제되었으며, 이 변수들의 결측률은 64~99%였다. 결측률로 인해 제외된 문항은 <표 1>에 제시하였다. 다중대체 시행에는 R 프로그램

(Ver, 3.3.4)의 mice 패키지(Ver, 3.4.0)(van Buuren & Groothuis-Oudshoorn, 2011)를 활용하였다.

〈표 1〉 결측률이 높은 삭제 문항

변수		결측률(%)
W707	HR업무 담당인력의 전문성 수준	80.37
W708_01	[인적자원역량의 경쟁력 수준] (1)인적자원의 전반적인 역량	80.37
W708_02	[인적자원역량의 경쟁력 수준] (2)우수인력의 확보	80.37
W708_03	[인적자원역량의 경쟁력 수준] (3)신뢰를 바탕으로 한 공동체 형성	80.37
W708_04	[인적자원역량의 경쟁력 수준] (4)경영진의 리더십	80.37
W708_05	[인적자원역량의 경쟁력 수준] (5)종업원 생산성	80.37
W708_06	[내부 프로세스 역량의 경쟁력 수준] (1)신제품(상품/서비스) 개발능력	80.37
W708_07	[내부 프로세스 역량의 경쟁력 수준] (2)업무절차의 효율화·간소화	80.37
W708_08	[내부 프로세스 역량의 경쟁력 수준] (3)원가절감을 통한 경쟁우위의 확보	80.37
W708_09	[내부 프로세스 역량의 경쟁력 수준] (4)제품(상품/서비스)의 품질	80.37
W708_10	[내부 프로세스 역량의 경쟁력 수준] (5)불량률 감소 및 생산수율 향상	85.34
W708_11	[대고객역량의 경쟁력 수준] (1)고객요구에 대한 신속한 대응능력	80.37
W708_12	[대고객역량의 경쟁력 수준] (2)제품(상품/서비스)의 다양성	80.37
W708_13	[대고객역량의 경쟁력 수준] (3)신규고객 확보율	80.37
W708_14	[대고객역량의 경쟁력 수준] (4)주요고객 유지율	80.37
W708_15	[대고객역량의 경쟁력 수준] (5)브랜드 이미지 향상 및 관리	80.37
W709_01	[분야별 인력 수준] (1)연구개발인력	81.64
W709_02	[분야별 인력 수준] (2)영업 및 서비스인력	80.37
W709_03	[분야별 인력 수준] (3)엔지니어(생산/제조기술)인력	85.34
W709_04	[분야별 인력 수준] (4)관리(경영지원)인력	80.37
W709_05	[분야별 인력 수준] (5)생산기능직인력	85.34
W709_06	[분야별 인력 수준] (6)상품개발 및 영업기획	98.73
W709_07	[분야별 인력 수준] (7)자금운용 및 투자인력	98.73
W709_08	[분야별 인력 수준] (8)핵심전문직인력	96.30
W710_04	[교육훈련 방법별 참여현황] (2)집체식 사외교육훈련 - 2)직무능력에 미친 효과	64.73
W710_06	[교육훈련 방법별 참여현황] (3)인터넷 학습 - 2)직무능력에 미친 효과	64.95
W710_08	[교육훈련 방법별 참여현황] (4)우편통신훈련 - 2)직무능력에 미친 효과	90.67
W710_10	[교육훈련 방법별 참여현황] (5)국내연수 - 2)직무능력에 미친 효과	96.76
W710_12	[교육훈련 방법별 참여현황] (6)해외연수 - 2)직무능력에 미친 효과	98.60
W710_14	[교육훈련 방법별 참여현황] (7)외부업체로부터 기술지도 받기 - 2)직무능력에 미친 효과	94.29
W713_04	[HRD 제도별 참여현황] (3)경력개발제도 - 2)직무능력에 미친 효과	92.88
W713_06	[HRD 제도별 참여현황] (4)교육훈련 휴가제 - 2)직무능력에 미친 효과	94.52
W713_08	[HRD 제도별 참여현황] (5)멘토링 또는 코칭 - 2)직무능력에 미친 효과	77.90
W713_10	[HRD 제도별 참여현황] (6)학습조직 - 2)직무능력에 미친 효과	90.31
W713_12	[HRD 제도별 참여현황] (7)OJT 프로그램 - 2)직무능력에 미친 효과	72.41
W713_14	[HRD 제도별 참여현황] (8)직무순환 - 2)직무능력에 미친 효과	84.98
W713_16	[HRD 제도별 참여현황] (9)제안제도 - 2)직무능력에 미친 효과	65.65
W713_18	[HRD 제도별 참여현황] (10)지식 마일리지 프로그램 - 2)직무능력에 미친 효과	96.78

변수		결측률(%)
W713_20	[HRD 제도별 참여현황] (11)QC(품질분임조) - 2)직무능력에 미친 효과	85.84
W713_22	[HRD 제도별 참여현황] (12)전사적 품질관리 - 2)직무능력에 미친 효과	85.56
W713_24	[HRD 제도별 참여현황] (13)6-시그마 - 2)직무능력에 미친 효과	93.03
W714_02	[자기주도적학습 지원제도] (1)학원 수강료 지원제도 - 2)직무능력에 미친 효과	82.60
W714_04	[자기주도적학습 지원제도] (2)국내 대학 등록금 지원제도 - 2)직무능력에 미친 효과	99.80
W714_06	[자기주도적학습 지원제도] (3)국내 대학원 등록금 지원제도 - 2)직무능력에 미친 효과	99.88
W714_08	[자기주도적학습 지원제도] (4)해외 대학원 학위과정 지원제도 - 2)직무능력에 미친 효과	99.94
W716	근무한 반수	64.83
W717_01	[숙련수준] (1)입사당시	64.83
W717_02	[숙련수준] (2)현재	64.83
W719_01	[직무 관련성이 가장 높은 자격증] (1)자격증명 코드1	79.07
W719_02	[직무 관련성이 가장 높은 자격증] (1)자격증명 코드2	79.07
W719_03	[직무 관련성이 가장 높은 자격증] (1)자격증명 코드3	79.07
W719_04	[직무 관련성이 가장 높은 자격증] (2)취득년도	79.07
W719_05	직무 관련성이 가장 높은 자격증의 종류	79.07
W719_06	직무 관련성이 가장 높은 자격증 취득의 업무능력 향상 정도	79.07
W719_07	직무 관련성이 가장 높은 자격증의 활용분야	79.07
W723_01	[팀원에 대한 견해] (1)팀의 문제해결 및 의사결정에 적극적 참여	80.37
W723_02	[팀원에 대한 견해] (2)팀 내에서 다양한 경험을 쌓을 수 있음	80.37
W723_03	[팀원에 대한 견해] (3)업무를 주도할 충분한 자율성 부여	80.37
W723_04	[팀원에 대한 견해] (4)팀 내 협동 및 신뢰관계 구축	80.37
W723_05	[팀원에 대한 견해] (5)이질적 업무 수행을 위한 역량과 경험 보유	80.37
W723_06	[팀원에 대한 견해] (6)팀 내 의사소통 원활	80.37
W724_01	[반의 자율권] (1)작업업무량 결정	94.03
W724_02	[반의 자율권] (2)업무수행방식 결정	94.03
W724_03	[반의 자율권] (3)업무수행속도 결정	94.03
W724_04	[반의 자율권] (4)작업시간(초과근무 등) 결정	94.03
W724_05	[반의 자율권] (5)작업단위내 업무배분	94.03
W724_06	[반의 자율권] (6)작업단위내 로테이션	94.03
W724_07	[반의 자율권] (7)작업단위별 훈련	94.03
W724_08	[반의 자율권] (8)작업반원 채용	94.03
W730_01	[생산직 근로자 숙련수준 분포] (1)창의적, 혁신적수준	97.54
W730_02	[생산직 근로자 숙련수준 분포] (2)다른 사람을 가르치는 수준	97.54
W730_03	[생산직 근로자 숙련수준 분포] (3)독자 수행하는 수준	97.54
W730_04	[생산직 근로자 숙련수준 분포] (4)독자 수행에 미흡한 수준	97.54
W731_01	[생산라인별 수준] (1)1번째 라인 - 2)근로자 숙련수준	97.54
W731_02	[생산라인별 수준] (1)1번째 라인 - 3)생산설비 수준	97.54
W731_03	[생산라인별 수준] (2)2번째 라인 - 2)근로자 숙련수준	97.59
W731_04	[생산라인별 수준] (2)2번째 라인 - 3)생산설비 수준	97.59
W731_05	[생산라인별 수준] (3)3번째 라인 - 2)근로자 숙련수준	97.81
W731_06	[생산라인별 수준] (3)3번째 라인 - 3)생산설비 수준	97.81
W736_02	2016년 월평균 근로소득	88.75

## 나. 설명변수

데이터 정리 과정을 통해 분석에 포함하기 어려운 변수들을 제외하고 최종적으로 총 88개 설명변수를 활용하였으며, 88개의 설명변수 중에서 범주형 변인의 경우 더미코딩하여 분석에 포함하였다. 범주형 변인 중 3개 이상의 항목으로 구분되어 있는 경우에는 0과 1로 더미코딩을 하기 위해 항목들을 유사한 의미로 구분하였으며, 응답 항목별 빈도도 함께 고려하였다. 팀 구분(W7\_team) 변수는 '1(관리직)', '2(생산관리직)'을 '1'로, '3(생산직)', '4(서비스직)', '5(핵심전문직)'을 '0'으로 더미코딩하여 다른 팀과 관리직을 구분하였다. 산업 대분류(W7\_ind1) 변수의 경우 전체 사례 중 제조업에 재직 중인 경우가 80.7%로 나타나, '1(제조업)'을 '1'로, 그 외 '2(금융업)'과 '3(비금융업)'을 '0'으로 변환하였다. 팀장/팀원여부(W7\_posit)는 '1(팀장)'과 '2(생산직 감독자)'를 '1'로, '3(팀원)'과 '4(생산직 근로자)'를 '0'으로 재코딩하여 더미변수 처리하였다. 입사당시 직급(W703\_01)과 현재 직급(W703\_02)을 묻는 변수의 경우 가장 많은 구성원이 해당되는 사원급을 기준 구분하였으며, '1(사원급)'외 사원급이 아닌 '2(주임/계장)', '3(대리급)', '4(과장급)', '5(차장급)', '6(부장급)', '7(임원급)' 및 '8(생산직 반장급)'을 '0'으로 변환하였다. 또한 교육훈련 방법별 참여현황, HRD 제도별 참여현황, 자기주도적학습 지원제도와 같이 참여여부를 묻는 변수의 경우 '2(참여않음)'을 '0'으로 변환하였다. 직무 관련 자격증 보유(W718) 변수에 '2(없음)'에 응답한 경우와 정규직 여부(W732), 수습(견습)사원 여부(W733), 노동조합원 여부(W734)를 묻는 문항에 '2(아니오)'로 응답한 경우도 '0'으로 변환하였다. 성별(W7\_sex)은 '1'은 '남자', '0'을 '여자'로 재코딩하였으며, 혼인상태(W7\_marr)는 '1(미혼)'과 '3(이혼 또는 사별)'을 '1'로, '2(기혼(재혼 포함))'을 '0'으로 하였다. 최종학력(W7\_edu)은 4년제 대졸이 48.7%로 나타나 비율을 유사하도록 4년제 대졸을 기준으로 구분하였다. 즉, 4년제 대졸 이상인 '6(4년제 대졸)', '6(석사졸)', '7(박사졸)'을 '1'로, '1(중졸 이하)', '2(인문고졸)', '3(공고졸)', '4(기타 실업고졸)' 및 '5(전문대(2·3년제 대학)졸)'을 '0'으로 변환하였다. 전공계열(W7\_affil)은 문과와 이과로 구분하였으며, 이과계열인 '3(공학계열)', '4(자연계열)', '5(이과계열)'을 '1'로, 문과계열인 '1(인문계열)', '2(사회계열)', '6(교육(사범)계열)', '7(예술/체육계열)'을 '0'으로 변환하여 분석에 포함하였다.

## 다. 결과변수

본 연구에서는 기업 구성원의 직무만족과 조직몰입을 결과변수로 활용하였다. 직무만족의 경우 현재 일에 대한 만족, 현재 받는 임금에 대한 만족, 현 직장의 인간관계에 대한 만족을 묻고 있으며, '1(전혀 그렇지 않음)'에서 '5(전적으로 그러함)'까지 5점 리커트 척도로 측정되었다. 즉, 점수가 높을수록 직무만족이 높음을 의미하며, 3문항의 평균을 사용하였다. 조직몰입은 4문항이 5점 리커트 척도로 측정되었으며, 한 문항(더 좋은 조건을 제시할 경우 이직을 고려함)의 경우 역문항이므로 역채점하였다. 역채점 후 조직몰입도 4문항의 평균을 분석에 사용하였다. 직무만족 3문항의 신뢰도는 .737이며, 조직몰입 4문항의 신뢰도는 .736으로 양호한 신뢰도를 보였다. 직무만족과 조직몰입에 대한 자세한 문항은 <표 2>와 같다

<표 2> 결과변인 문항

변인	문항	척도	신뢰도
직무 만족 (3문항)	(1)현재 일에 만족	5점 리커트 척도 (1=전혀 그렇지 않음, 2= 그렇지 않은 편임, 3=그저 그러함, 4=그런편임, 5=전적으로 그러함)	.737
	(2)현재 받는 임금에 만족		
	(3)현 직장의 인간관계 만족		
조직 몰입 (4문항)	(1)더 좋은 조건을 제시할 경우 이직을 고려함(역문항)		.736
	(2)회사의 문제를 내 문제처럼 느낌		
	(3)회사를 떠난다면 많은 것을 잃게 될 것임		
	(4)회사에 충성할 만한 가치가 있음		

## 3. 연구 방법

### 가. 분석 절차

본 연구의 구체적인 분석 절차는 다음과 같다. 크게는 잠재프로파일 분석 단계와 머신러닝 기법을 적용한 분석 단계로 구분할 수 있다. 우선, 잠재프로파일 분석 단계에서는 기업 구성원의 직무만족 및 조직몰입을 이질적인 잠재집단으로 유형화 하였다. 이를 위해 첫째, 정보지수인 AIC, BIC, SABIC, 분류의 질인 Entropy, 5%로 미만인 집단이 존재하는 지 여부, 해석의 용이성 등을 판단하여 직무만족 및 조직몰입을 각각 몇 개의 이질적인 잠재프로파일로 구분할 것인지 결정한다. 둘째, 잠재프로파일의 수를 결정하고 나

면, 각 잠재프로파일별 직무만족 및 조직몰입 수준을 확인하여 특징에 따라 집단을 명명해준다.

다음으로 머신러닝 기법을 적용하여 설명변수 중 중요한 변수를 탐색하는 단계이다. 구체적으로 본 연구에서는 설명변수를 활용해 LASSO, Ridge, Randomforest, SVM 등 4가지 머신러닝 기법 중 직무만족 및 조직몰입 잠재프로파일을 가장 잘 분류하는 기법을 선택하고, 선택한 기법을 적용한 모형에서 중요한 변수를 탐색하는 과정을 거친다. 이를 위해 첫째, 설명변수를 정리하는 과정을 진행한다. 둘째, 자료를 훈련용(train) 70%와 검증용(test) 30%로 구분한 후 분석 및 모형 평가를 실시하였는데, 10005명 중 7,045명이 훈련용 자료에, 2,960명이 검증용 자료에 포함되었다. 훈련용 자료에 LASSO, Ridge, Randomforest, SVM 등 4가지 머신러닝 기법을 적용해 설명변수로 직무만족 및 조직몰입 잠재프로파일을 예측한다. 단, 대부분의 연구들이 집단을 0, 1(예를들어, 중앙값을 기준으로 조직만족이 낮은 집단과 높은 집단을 구분)로 분류한 후 이 집단을 예측하였지만, 본 연구의 경우 중앙값, 평균값 등의 기준을 적용해 단순히 두 개의 집단으로 구분한 것이 아니라 잠재프로파일 분석을 실시하여 기업 구성원의 직무만족 및 조직몰입의 여러 개의 이질적인 잠재집단을 확인하였으므로, 이항(binomial)모델이 아닌 다항(multinomial)모델을 구축하였다. 셋째, 모형의 성과를 평가하기 위해 검증용 자료를 활용해 4가지 머신러닝 기법 중 예측정확도(accuracy)가 가장 높은 모형을 선택하였다. 넷째, 선택한 머신러닝 기법을 활용해 직무만족 및 조직몰입 잠재프로파일 분류에 중요한 영향요인을 도출하였다.

## 나. 잠재프로파일 분석

잠재프로파일 분석(Latent Profile Analysis 이하 LPA)은 관측된 자료로부터 모집단에서 존재하는 비관측된 이질적인 특성을 확인하는 것으로, 집단 내에서도 수준이나 특징이 다른 하위 집단으로 구분하게 된다(신태수, 2010). 즉, 잠재프로파일 분석은 혼합모형 분석의 한 종류라고 할 수 있으며, 연구대상이 응답한 문항 패턴을 통해 집단 내 존재하는 이질적인 잠재집단을 분류하는 것이다(Muthén, 2004).

군집분석 등 다른 집단 분류 방법과 잠재프로파일 분석의 차이는 잠재프로파일 수를 결정하는 방법에 있다(박민지, 이신혜, 2017). 잠재프로파일 분석에서는 다양한 기준을 적용하여 최종 잠재계층 수를 결정하게 되는데, 대표적으로 정보지수인 AIC(Akaike Information Criterion), BIC(Baysian Information Criterion), SABIC(Sample-size

Adjust BIC)와, 분류의 질을 확인할 수 있는 Entropy, LMRT(Lo-Mendell-Rubin Adjusted LRT Test) 등의 모형적합도 지수와 집단 분류율 및 해석의 가능성과 용이성 등을 종합적으로 고려해야 한다. 정보지수인 AIC, BIC, SABIC는 작을수록, Entropy는 0.8이상 높을수록 좋은 결정 조건이 된다. 또한 LMRT는 통계적 유의성을 확인하게 되는데, 유의하다면 1개의 잠재프로파일을 증가시킨 모형이 전 모형보다 적합하다는 것을 의미하며, 유의하지 않다면 잠재프로파일 수를 늘린 모형보다 이전의 모형이 더 적합함을 의미한다. 본 연구는 직무만족 및 조직몰입의 잠재프로파일 분석을 위해 Mplus 8.3(Muthén & Muthén, 1998-2019)를 사용하였다.

## 다. 머신러닝 기법

### 1) LASSO

LASSO는 벌점을 부과해 변수를 선택하는 벌점화 회귀모형 중 하나이다(Tibshirani, 1996). LASSO는 일부 중요한 변수를 제외하고 영향력이 없거나 낮은 변수의 회귀계수를 0으로 만들어 변수를 선택하게 되는 방법인데, 조율모수( $\lambda$ )에 따라 0으로 되는 변수의 수가 달라진다(유진은, 2016). 그러므로 LASSO에서 조율모수( $\lambda$ )를 결정하는 것이 매우 중요한데, 10-fold 교차타당화(Cross-Validation, Hastie, Tibshirani, & Friedman, 2009)를 사용해 조율모수( $\lambda$ )를 선택하였다. 교차타당화 결과, 직무만족의 조율모수( $\lambda$ )는 0.01126508, 조직몰입의 조율모수( $\lambda$ )는 0.01246315로 나타났다. LASSO 분석을 위해 R 3.3.4의 패키지 glmnet(Ver, 2.0-16)(Friedman, Hastie, & Tibshirani, 2019)을 사용하였다.

### 2) Ridge

Ridge regression(능형회귀)도 LASSO와 마찬가지로 벌점화 회귀모형 중 하나이다(Hoerl & Kennard, 1971). 단, LASSO는 영향력이 낮은 변수의 회귀계수를 0으로 만드는데 비해, Ridge는 회귀계수에 벌점을 부과해 회귀계수를 축소하여 평균제곱오차를 줄인다(김은령, 2010). 즉, LASSO는 영향력이 큰 일부 설명변수의 회귀계수만 확인할 수 있는 반면, Ridge는 모든 설명변수의 회귀계수를 보여주기 때문에 결과 해석이 어려울 수 있다. Ridge에서도 교차타당화 과정을 통해 조율모수( $\lambda$ )를 결정해야 하는데, 조율모

수( $\lambda$ )가 클수록 회귀계수가 0에 가까워지게 된다. 교차타당화 결과 Ridge를 적용했을 때 직무만족의 조율모수( $\lambda$ )는 0.1421492, 조직몰입의 조율모수( $\lambda$ )는 0.1572672로 나타났다. Ridge 분석에도 R 3.3.4의 패키지 glmnet(Ver, 2.0-16)(Friedman, Hastie, & Tibshirani, 2019)을 사용하였다.

### 3) Randomforest

Randomforest(RF 일명 랜덤포레스트) 기법은 의사결정나무(decision tree) 분석에서 확장된 분석 기법으로, 수백 개 이상의 의사결정나무를 생성(본 연구에서는 500개의 트리를 생성)하여 더 안정적이고 정확한 분류를 하게 된다. 의사결정나무 분석의 경우 구축된 모형을 검증용 자료와 같이 새로운 자료에 적용하면 모형 자체가 상당히 변화할 수 있고, 예측정확도가 낮다는 등의 문제점이 지적되어 왔기 때문에, 본 연구에서는 의사결정나무 분석에 비해 예측정확도가 높고, 새로운 자료에 적용해도 모형이 상대적으로 안정적인 Randomforest를 적용하여 분류를 수행하고자 한다(유진은, 2015). 다만 의사결정나무 분석에서는 나무 모양의 그림이 제공되어 가시적인 결과를 확인할 수 있으나, 랜덤포레스트의 경우 중요도 지수와 같이 설명변수 중 상대적으로 중요한 정도와 관련된 정보가 제시된다. Randomforest 분석을 위해 R 3.3.4의 패키지 randomForest(Ver, 4.6-14)(Breiman, Cutler, Liaw, Wiener, 2018)를 사용하였다.

### 4) Support Vector Machine(SVM)

SVM 기법은 자료를 가장 잘 분류해주는 최적의 초평면을 찾는 것이다(Cortes & Vapnik, 1995). 즉, 집단 사이의 마진(margin)을 최대로 하는 경계면을 찾는 것이다(박영재, 김계영, 장석우, 2013). SVM은 최적의 gamma와 c 값을 찾는 과정에 많은 시간이 소요되고, 이러한 모수를 찾는 과정이 모형 예측에 중요한 영향을 미치게 된다(제홍모, 방승양, 2003). 모수 튜닝(parameter tuning)과정을 통해 직무만족은 gamma는 0.03125, c는 2로 선택되었으며, 조직몰입은 gamma가 0.03125, c는 1로 결정되었다. 본 연구는 SVM 분석을 위해 R 3.3.4의 패키지 e1071(Ver, 1.7-1)(Meyer, Dimitriadou, Hornik, Weingessel, Leisch, 2019)을 사용하였다.

## IV 연구 결과

### 1. 잠재프로파일 분석

#### 가. 잠재프로파일 수 결정

##### 1) 직무만족

직무만족의 잠재프로파일 수를 결정하기 위해 잠재프로파일 분석을 실시하였으며, 잠재프로파일 수를 2개에서 5개까지 하나씩 증가시키면서 모형 적합도를 비교한 결과를 <표 3>에 결과를 제시하였다. <표 3>를 보면, 2개부터 5개까지 잠재프로파일 수를 추가 하면서 AIC, BIC, SABIC, Entropy, LMRT 및 집단 비율 등을 종합적으로 고려하였다. 우선 분석결과, 4계층과 5계층의 경우 2와 3계층에 비해 AIC, BIC, SABIC가 작고 Entropy가 높았지만 5% 미만이 속하는 집단이 존재하고 결과가 수렴하지 않는 것으로 나타났다. 그러므로 4계층과 5계층은 제외하였으며, 2계층에 비해 정보지수(AIC, BIC, SABIC)가 작고 Entropy가 0.8이상인 3계층을 최종적으로 결정하였다.

**<표 3> 잠재프로파일 수에 따른 모형 적합도 비교(직무만족) (n=10,005)**

	2계층	3계층	4계층	5계층
AIC	65425.227	61961.107	38572.092	37765.033
BIC	65497.335	62062.059	38701.887	37923.672
SABIC	65465.556	62017.569	38644.686	37853.759
Entropy	0.750	0.992	1.000	0.960
LMRT	5337.098 <sup>***</sup>	3380.370 <sup>***</sup>	21183.125	802.610 <sup>***</sup>
계층 1(%)	3,687(36.9%)	5,346(53.4%)	5,515(55.1%)	372(3.7%)
계층 2(%)	6,318(63.1%)	3,713(37.1%)	3,008(30.1%)	429(4.3%)
계층 3(%)		946(9.5%)	372(3.7%)	3,008(30.1%)
계층 4(%)			1,110(11.1%)	5,086(50.8%)
계층 5(%)				1,110(11.1%)

\*\*\* $p < .001$

## 2) 조직몰입

조직몰입의 잠재프로파일 수를 결정하기 위해 잠재프로파일 분석을 실시하였으며, 잠재프로파일 수를 2개에서 5개까지 하나씩 증가시키면서 모형 적합도를 비교한 결과를 <표 4>에 결과를 제시하였다. 2개부터 5개까지 잠재프로파일 수를 추가하면서 AIC, BIC, SABIC, Entropy, LMRT 및 집단 비율 등을 종합적으로 고려한 결과, 조직몰입의 경우에도 4계층과 5계층의 결과가 수렴하지 않는 것으로 나타나 제외하였다. 이에 2계층에 비해 정보지수(AIC, BIC, SABIC)가 작고 Entropy가 0.991로 매우 높은 3계층을 최종적으로 결정하였다. 3계층으로 구분했을 때, 5% 미만이 속하는 집단은 없는 것으로 나타났다.

<표 4> 잠재프로파일 수에 따른 모형 적합도 비교(조직몰입)

(n=10,005)

	2계층	3계층	4계층	5계층
AIC	97181.777	92824.568	80489.695	80363.116
BIC	97275.518	92954.363	80655.544	80565.020
SABIC	97234.206	92897.162	80582.453	80476.040
Entropy	0.682	0.991	1.000	0.996
LMRT	6145.620***	4274.397***	12082.519***	134.750
계층 1(%)	5,408(54.1%)	968(9.9%)	987(9.9%)	987(9.9%)
계층 2(%)	4,597(45.9%)	4,370(43.7%)	3,739(37.4%)	4,648(46.4%)
계층 3(%)		4,649(46.4%)	4,648(46.4%)	3,706(37.1%)
계층 4(%)			631(6.3%)	631(6.3%)
계층 5(%)				33(0.3%)

\*\*\* $p < .001$

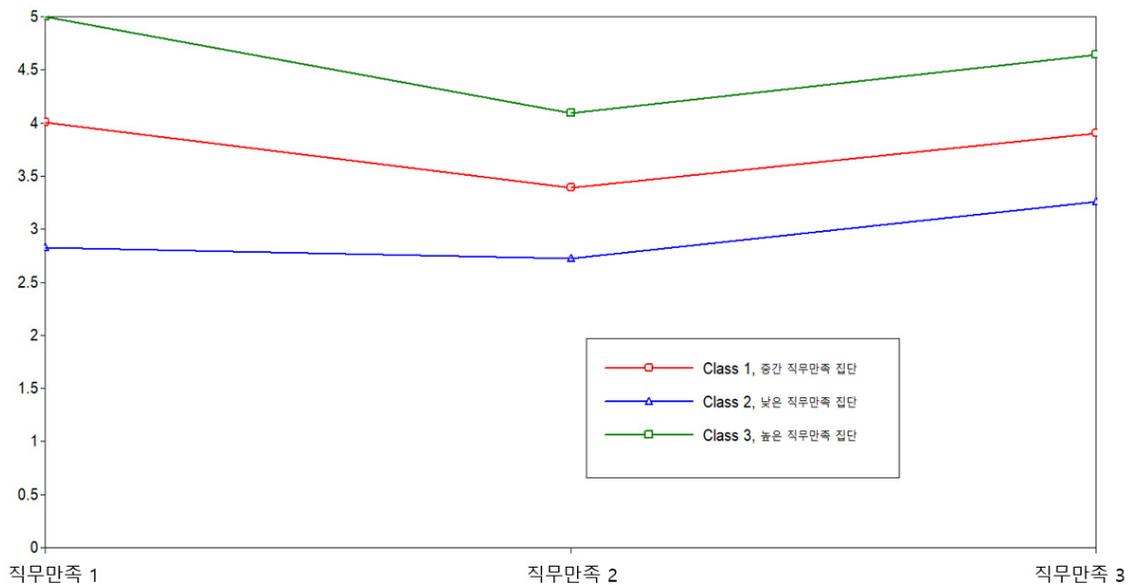
## 나. 잠재프로파일 유형

### 1) 직무만족

기업 구성원의 직무만족에 따라 최종적으로 3개의 잠재프로파일을 결정했으며, 이질적인 각 잠재프로파일의 특징을 살펴보기 위해 도출된 잠재프로파일의 직무만족 수준을 [그림 1]에 제시하였다. 또한 3개 잠재프로파일의 직무만족 수준에 유의미한 차이가 나

타나는지를 확인하기 위해 평균, 표준편차 및 일원변량 분산분석 결과를 <표 5>에 제시하였다.

3개의 잠재프로파일 특성을 살펴보면, 계층 1은 직무만족 3문항(일, 임금, 인간관계에 대한 만족)에서 계층 2, 계층 3과 비교해 중간 수준의 직무만족을 보였으며, 과반수가 넘는 53.4%(5,346명)의 기업 구성원이 계층 1에 속하는 것으로 나타났다. 직무만족 수준이 다른 집단에 비해 중간 수준인 계층 1을 ‘중간 직무만족 집단’이라고 명명하였다. 계층 2는 계층 1과 계층 3에 비해 직무만족 수준이 낮은 특징을 보이고 있으며, 계층 2에는 37.1%(3,713명)가 속한 것으로 나타났다. 계층 2는 일, 임금, 인간관계에 대한 만족도가 상대적으로 낮은 수준을 보이고 있어, 이 집단을 ‘낮은 직무만족 집단’으로 명명하였다. 마지막으로 계층 3은 다른 집단에 비해 직무만족이 높은 것으로 나타나, 이를 ‘높은 직무만족 집단’으로 명명하였다. 특히 이 집단의 경우 현재 일에 대한 만족 문항(직무만족 1)의 평균이 5점 만점에 5점으로 매우 높은 수준을 보였다. ‘높은 직무만족 집단’은 다른 집단에 비해 가장 비율이 낮은 9.5%(946명)만이 속하는 것으로 나타났다.



[그림 1] 직무만족의 잠재프로파일 형태

〈표 5〉 잠재프로파일별 직무만족 수준

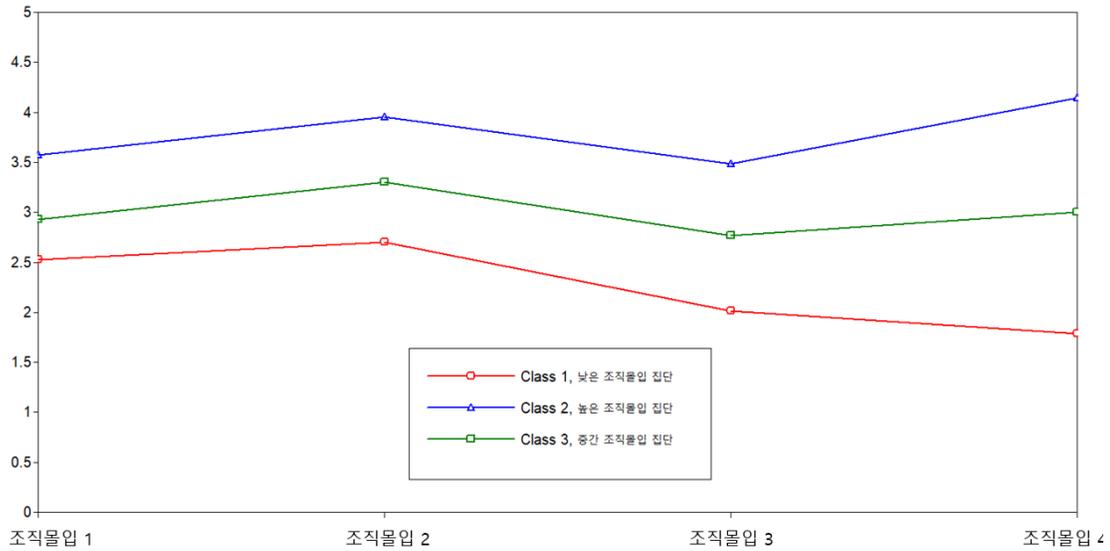
	계층1 중간 직무만족 집단 ( $n_1=5,346$ )	계층2 낮은 직무만족 집단 ( $n_2=3,713$ )	계층3 높은 직무만족 집단 ( $n_3=946$ )	<i>F</i>	<i>df</i>	사후 검증 (Tukey)
	<i>M(SD)</i>	<i>M(SD)</i>	<i>M(SD)</i>			
직무만족 1	4.001(.024)	2.827(.430)	5.000(.000)	35474.887***	2	2<1<3
직무만족 2	3.389(.798)	2.723(.790)	4.081(.957)	1344.598***	2	2<1<3
직무만족 3	3.899(.512)	3.259(.681)	4.630(.513)	2561.998***	2	2<1<3

\*\*\* $p < .001$

## 2) 조직몰입

기업 구성원의 조직몰입에 따라 최종적으로 3개의 잠재프로파일을 결정했으며, 이질적인 각 잠재프로파일의 특징을 살펴보기 위해 도출된 잠재프로파일의 조직몰입 수준을 [그림 2]에 제시하였다. 또한 도출된 3개 잠재프로파일의 조직몰입 수준이 유의미하게 차이가 있는지를 확인하기 위해 평균, 표준편차 및 일원변량 분산분석 결과를 〈표 6〉에 제시하였다.

3개의 잠재프로파일 특성을 살펴보면, 계층 1은 조직몰입 4문항에서 모두 계층 2, 계층 3과 비교해 낮은 조직몰입 수준을 보였다. 또한 전체 기업 구성원 중 9.9%(968명)만이 계층 1에 속하는 것으로 나타났다. 조직몰입이 다른 두 집단에 비해 낮은 수준을 보이는 계층 1을 ‘낮은 조직몰입 집단’으로 명명하였다. 반면 계층 2는 조직몰입이 다른 집단에 비해 높은 수준인 것으로 나타났다. 조직몰입 4문항의 평균이 3.568~4.144로, 계층 1과 계층 2와 비교해서는 조직에 몰입하는 정도가 높은 것으로 나타나 이 계층을 ‘높은 조직몰입 집단’으로 명명하였다. ‘높은 조직몰입 집단’인 계층 2에는 43.7%(4,370명)이 속하는 것으로 나타났다. 계층 3은 다른 집단과 비교해 조직몰입이 중간 수준으로, 이에 이 계층을 ‘중간 조직몰입 집단’으로 명명하였다. ‘중간 조직몰입 집단’은 전체 기업 구성원 중 46.4%(4,649명)가 속해 가장 높은 비율을 보였다.



[그림 2] 조직몰입의 잠재프로파일 형태

<표 6> 잠재프로파일별 조직몰입 수준

	계층1 낮은 조직몰입 집단 (n <sub>1</sub> =986)	계층2 높은 조직몰입 집단 (n <sub>2</sub> =4,370)	계층3 중간 조직몰입 집단 (n <sub>3</sub> =4,649)	F	df	사후 검증 (Tukey)
	M(SD)	M(SD)	M(SD)			
조직몰입 1	2.527(1.100)	3.568(.984)	2.928(.861)	773.771***	2	1<3<2
조직몰입 2	2.709(.926)	3.950(.638)	3.302(.654)	1820.115***	2	1<3<2
조직몰입 3	2.024(.859)	3.481(.854)	2.766(.758)	1675.827***	2	1<3<2
조직몰입 4	1.788(.409)	4.144(.352)	3.000(.015)	40572.503***	2	1<3<2

\*\*\*p<.001

## 2. 직무만족 및 조직몰입 예측 및 영향요인 탐색

### 가. 기법별 예측정확도 비교

#### 1) 직무만족

4가지 머신러닝 기법을 사용하여 88개의 설명변수로 기업 근로자의 직무만족 잠재프로파일을 예측 분류한 결과는 <표 7>와 같다. <표 7>은 기법별 최종 모형이 새로운 자료에서도 얼마나 정확한 예측률을 보이는지 검증용 자료를 통해 분석한 결과이다. 검증용

자료 총 2,960명(계층 1(중간 직무만족 집단): 1,598명, 계층 2(낮은 직무만족 집단): 1,098명, 계층 3(높은 직무만족 집단): 264명)을 제대로 예측한 빈도를 제시한 오분류표를 보면, LASSO는 2,024명(계층 1(중간 직무만족 집단): 1,323명, 계층 2(낮은 직무만족 집단): 658명, 계층 3(높은 직무만족 집단): 43명), Ridge는 2,028명(계층 1(중간 직무만족 집단): 1,312명, 계층 2(낮은 직무만족 집단): 675명, 계층 3(높은 직무만족 집단): 41명), Randomforest는 2,075명(계층 1(중간 직무만족 집단): 1,271명, 계층 2(낮은 직무만족 집단): 736명, 계층 3(높은 직무만족 집단): 68명), SVM은 1,700명(계층 1(중간 직무만족 집단): 1,389명, 계층 2(낮은 직무만족 집단): 311명, 계층 3(높은 직무만족 집단): 0명)을 정확하게 분류하였다. LASSO, Ridge와 근소한 차이지만 Randomforest가 직무만족 잠재프로파일을 정확히 분류한 빈도가 높았으며, SVM의 경우에는 가장 많은 근로자들이 포함된 계층 1(중간 직무만족 집단)은 다른 기법에 비해 정확히 분류하였으나 가장 적은 근로자들이 속하는 계층 3(높은 직무만족 집단)을 제대로 분류한 빈도가 0으로 나타났다.

분류모형을 평가하기 위해 기법별 예측정확도를 비교해보면, Randomforest가 70.1%로 가장 높은 예측정확도를 보였으며, LASSO, Ridge는 예측정확도가 각각 68.4%, 68.5%로 매우 근사한 것으로 나타났다. 반면, SVM의 경우 57.4%로 다른 3가지 기법에 비해 예측정확도가 떨어지는 것으로 나타났다. 이에 본 연구에서는 검증용 자료로 평가한 모형의 예측정확도를 고려하여 Randomforest를 최종적으로 선택하였다.

〈표 7〉 기법별 오분류표 및 예측정확도 비교(직무만족)

	계층 1	계층 2	계층 3
계층 1	1323	438	206
계층 2	270	658	15
계층 3	5	2	43
Accuracy	0.6838		

(1) LASSO

	계층 1	계층 2	계층 3
계층 1	1312	422	209
계층 2	279	675	14
계층 3	7	1	41
Accuracy	0.6851		

(2) Ridge

	계층 1	계층 2	계층 3
계층 1	1271	360	181
계층 2	309	736	15
계층 3	18	2	68
Accuracy	0.7007		

(3) Randomforest

	계층 1	계층 2	계층 3
계층 1	1389	786	253
계층 2	207	311	11
계층 3	2	1	0
Accuracy	0.5743		

(4) SVM

## 2) 조직몰입

4가지 머신러닝 기법을 사용하여 88개의 설명변수로 기업 근로자의 조직몰입 잠재프로파일을 예측 분류한 결과는 <표 8>과 같다. 검증용 자료 총 2,960명(계층 1(낮은 조직몰입 집단): 285명, 계층 2(높은 조직몰입 집단): 1,266명, 계층 3(중간 조직몰입 집단): 1,409명)을 제대로 예측한 빈도를 제시한 오분류표를 보면, LASSO는 2,014명(계층 1(낮은 조직몰입 집단): 59명, 계층 2(높은 조직몰입 집단): 1,287명, 계층 3(중간 조직몰입 집단): 1,614명), Ridge는 2,005명(계층 1(낮은 조직몰입 집단): 58명, 계층 2(높은 조직몰입 집단): 1,286명, 계층 3(중간 조직몰입 집단): 1,616명), Randomforest는 2,050명(계층 1(낮은 조직몰입 집단): 104명, 계층 2(높은 조직몰입 집단): 1,252명, 계층 3(중간 조직몰입 집단): 1,604명), SVM은 1,710명(계층 1(낮은 조직몰입 집단): 2명, 계층 2(높은 조직몰입 집단): 1,065명, 계층 3(중간 조직몰입 집단): 1,893명)을 정확하게 분류하였다. 조직몰입에서도 앞선 직무만족과 마찬가지로 Randomforest가 조직몰입 잠재프로파일을 정확히 분류한 빈도가 높았으며, 다음으로 LASSO, Ridge, SVM 순으로 높았다. SVM의 경우 가장 많은 근로자들이 포함된 계층 3(중간 조직몰입 집단)은 다른 기법에 비해 많은 근로자를 정확히 분류하였으나, 가장 적은 근로자들이 속하는 계층 1(낮은 조직몰입 집단)을 정확히 분류한 빈도가 2명밖에 없는 것으로 나타났다.

분류모형을 평가하기 위해 기법별 예측정확도를 비교해보면, LASSO, Ridge 및 Randomforest의 예측정확도는 67~69%정도로 유사하게 나타났다. Randomforest가 69.3%로 가장 높은 예측정확도를 보였으며, LASSO는 68.0%, Ridge는 67.7%로 나타났다. 반면, SVM의 경우 57.7%로 다른 3가지 기법에 비해 예측정확도가 많이 떨어지는 것으로 나타났다. 이에 본 연구에서는 검증용 자료로 평가한 모형의 예측정확도를 고려하여 조직몰입 잠재프로파일을 가장 정확히 분류한 Randomforest를 최종적으로 선택하였다.

<표 8> 기법별 분류표 및 예측정확도 비교(조직몰입)

	계층 1	계층 2	계층 3
계층 1	41	1	17
계층 2	22	923	342
계층 3	222	342	1050
Accuracy	0.6804		

(1) LASSO

	계층 1	계층 2	계층 3
계층 1	42	1	15
계층 2	19	918	349
계층 3	224	347	1045
Accuracy	0.6774		

(2) Ridge

	계층 1	계층 2	계층 3
계층 1	69	4	31
계층 2	21	917	314
계층 3	195	345	1064
Accuracy	0.6926		

(3) Randomforest

	계층 1	계층 2	계층 3
계층 1	2	0	0
계층 2	38	663	364
계층 3	245	603	1045
Accuracy	0.5777		

(4) SVM

## 나. 직무만족 및 조직몰입 영향요인

앞서 4가지 머신러닝 기법의 예측성능을 비교한 결과, Randomforest가 직무만족 및 조직몰입 잠재프로파일을 가장 잘 분류하는 것으로 나타났다. 이에 따라 Randomforest를 활용해 직무만족 및 조직몰입 잠재프로파일 예측에 어떠한 설명변수가 상대적으로 중요한지를 탐색하였으며, 훈련용 자료로 모형을 구축한 결과이다. Randomforest에서는 Mean Decrease Accuracy(MDA)와 Mean Decrease Gini(MDG)로 상대적인 중요도 지수를 제시하는데, MDA에 비해 MDG가 안정적이고 강력한 결과를 제공하므로(Calle & Urrea, 2010), MDG를 기준으로 중요한 변수를 확인하였다.

### 1) 직무만족

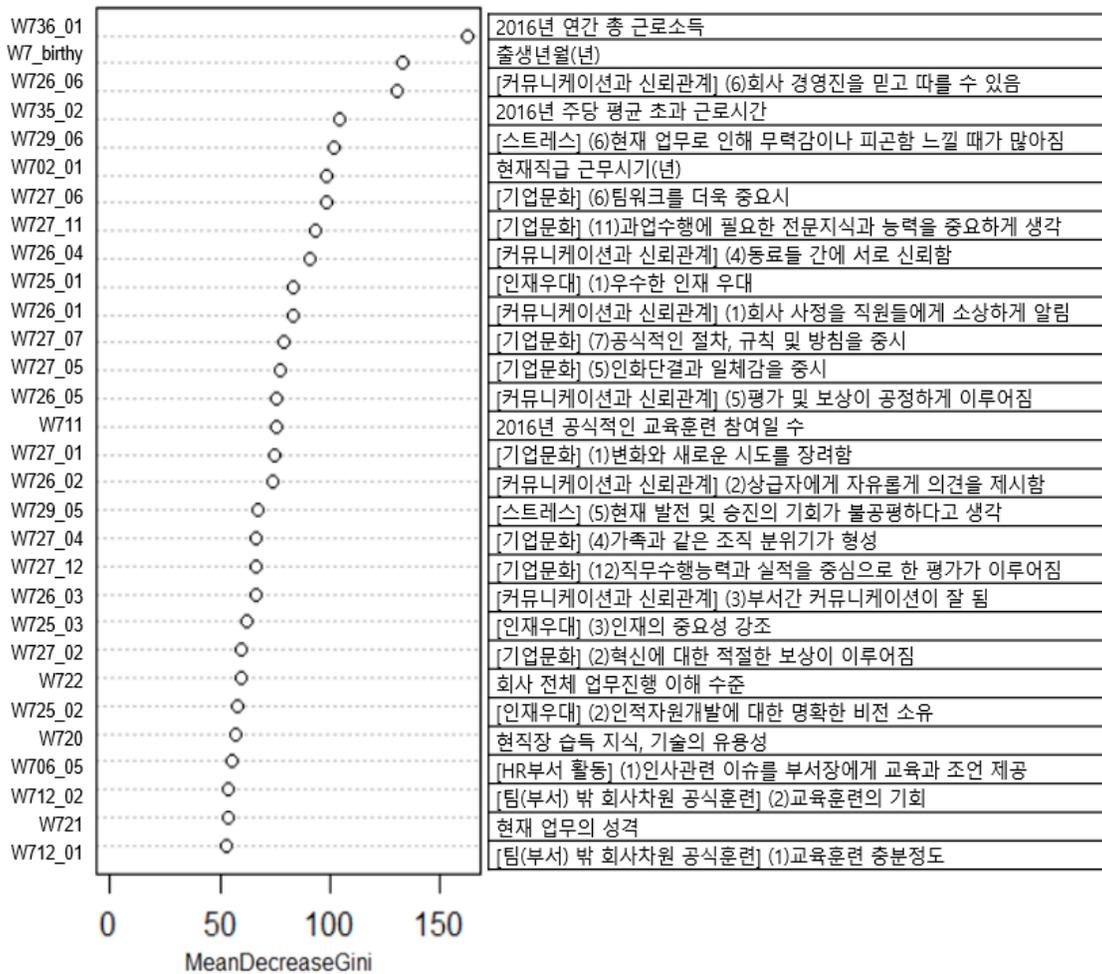
직무만족 수준에 따른 잠재프로파일 분류에 있어 중요도 지수가 높은 30개를 [그림 3]에 제시하였다. 설명변수 중 ‘2016년 연간 총 근로소득(W736\_01)’의 중요도가 가장 높은 것으로 나타났으며, 다음으로 ‘출생년월(년)(W7\_birthy)’, ‘[커뮤니케이션과 신뢰관계] (6)회사 경영진을 믿고 따를 수 있음(W726\_06)’, ‘2016년 주당 평균 초과 근로시간(W735\_02)’, ‘[스트레스] (6)현재 업무로 인해 무력감이나 피곤함 느낄 때가 많아짐(W729\_06)’ 순으로 중요도 지수가 높았다. 도출된 30개의 변수는 크게 근로자의 현재 상태 및 근무 상황, 기업과 근로자의 커뮤니케이션 및 신뢰관계, 스트레스, 기업문화, 인재우대, HR 부서활동, 팀 밖 회사차원의 공식훈련과 관련된 것으로 구분할 수 있다.

우선 근로자의 현재 상태 및 근무 상황과 관련된 변수로는 ‘2016년 연간 총 근로소득(W736\_01)’, ‘출생년월(년)(W7\_birthy)’, ‘2016년 주당 평균 초과 근로시간(W735\_02)’과 함께 ‘현재직급 근무시기(년)(W702\_01)’, ‘2016년 공식적인 교육훈련 참여일 수(W711)’, ‘회사 전체 업무진행 이해 수준(W722)’, ‘현직장 습득 지식, 기술의 유용성

(W720)', '현재 업무의 성격(W721)'이 중요한 변수로 나타났다. 커뮤니케이션 및 신뢰관계와 관련해서는 '[커뮤니케이션과 신뢰관계] (6)회사 경영진을 믿고 따를 수 있음(W726\_06)', '[커뮤니케이션과 신뢰관계] (4)동료들 간에 서로 신뢰함(W726\_04)', '[커뮤니케이션과 신뢰관계] (5)평가 및 보상이 공정하게 이루어짐(W726\_05)', '[커뮤니케이션과 신뢰관계] (2)상급자에게 자유롭게 의견을 제시함(W726\_02)', '[커뮤니케이션과 신뢰관계] (3)부서 간 커뮤니케이션이 잘 됨(W726\_03)' 등 5개의 변수가 중요하게 나타났다. 또한 근로자의 스트레스와 관련된 '[스트레스] (6)현재 업무로 인해 무력감이나 피곤함 느낄 때가 많아짐(W729\_06)', '[스트레스] (5)현재 발전 및 승진의 기회가 불공평하다고 생각(W729\_05)' 변수도 선택되었다.

기업의 문화와 관련된 변수도 중요도 지수가 높았는데, '[기업문화] (6)팀워크를 더욱 중요시(W727\_06)', '[기업문화] (11)과업수행에 필요한 전문지식과 능력을 중요하게 생각(W727\_11)', '[기업문화] (7)공식적인 절차, 규칙 및 방침을 중시(W727\_07)', '[기업문화] (5)인화단결과 일체감을 중시(W727\_05)', '[기업문화] (1)변화와 새로운 시도를 장려함(W727\_01)', '[기업문화] (4)가족과 같은 조직 분위기가 형성(W727\_04)', '[기업문화] (12)직무수행능력과 실적을 중심으로 한 평가가 이루어짐(W727\_12)', '[기업문화] (2)혁신에 대한 적절한 보상이 이루어짐(W727\_02)' 등 기업문화와 관련된 많은 변수들이 중요하게 나타났다. 인재우대와 관련된 변수인 '[인재우대] (1)우수한 인재 우대(W725\_01)', '[인재우대] (3)인재의 중요성 강조(W725\_03)', '[인재우대] (2)인적자원개발에 대한 명확한 비전 소유(W725\_02)' 등도 선택되었다. 그 외 HR부서 활동 중 '[HR부서 활동] (1)인사관련 이슈를 부서장에게 교육과 조언 제공(W706\_05)'과 팀 밖 회사차원의 공식훈련에 관한 '[팀(부서) 밖 회사차원 공식훈련] (2)교육훈련의 기회(W712\_02)', '[팀(부서) 밖 회사차원 공식훈련] (1)교육훈련 충분정도(W712\_01)' 등이 주요 변수로 도출되었다.

[그림 3] 랜덤포레스트 분석 결과(직무만족)



## 2) 조직몰입

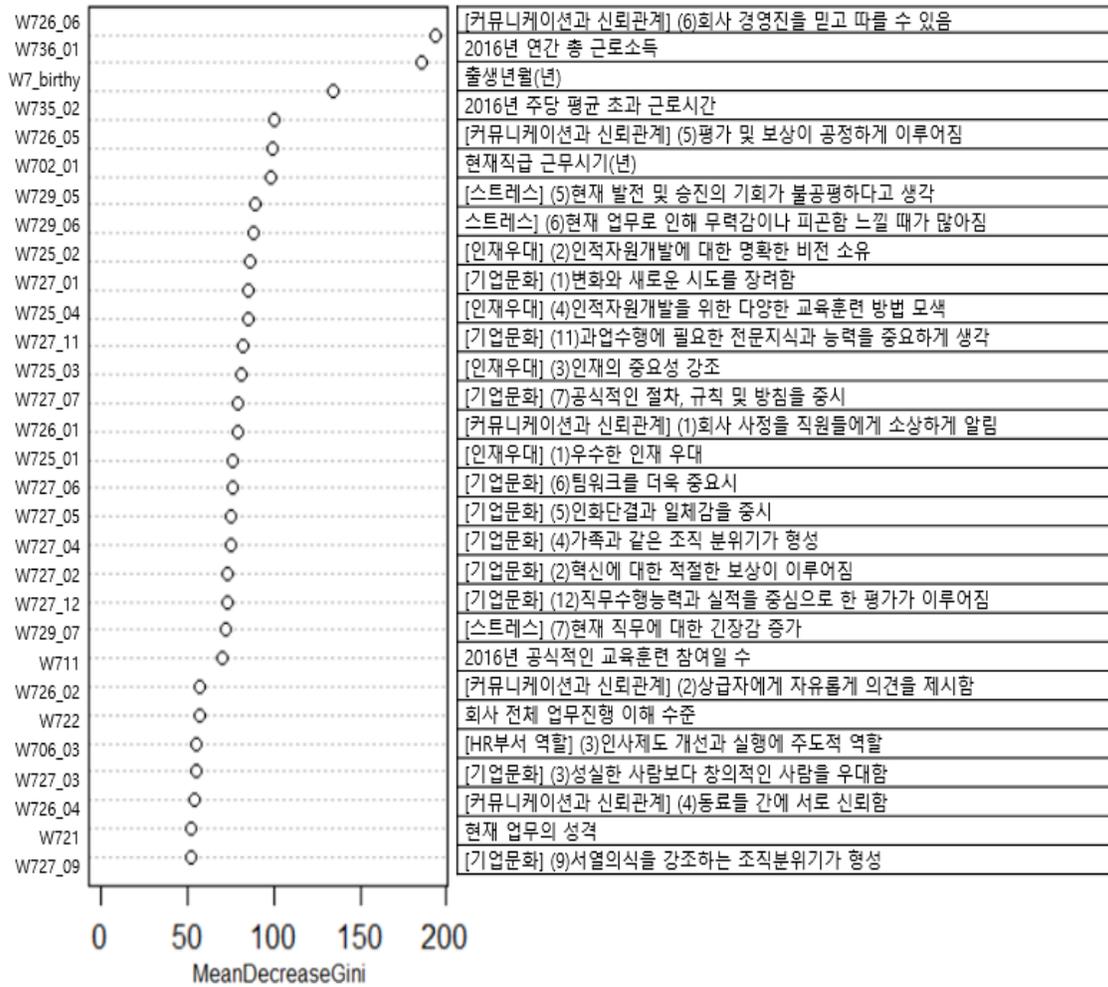
조직몰입 수준에 따른 잠재프로파일을 분류하는데 있어 중요도 지수가 높은 1위에서 30위까지를 [그림 4]에 제시하였다. 다양한 설명변수 중 ‘[커뮤니케이션과 신뢰관계] (6) 회사 경영진을 믿고 따를 수 있음(W726\_06)’의 중요도가 가장 높은 것으로 나타났으며, ‘2016년 연간 총 근로소득(W736\_01)’, ‘출생년월(년)(W7\_birthy)’, ‘2016년 주당 평균 초과 근로시간(W735\_02)’, ‘[커뮤니케이션과 신뢰관계] (5)평가 및 보상이 공정하게 이루어짐(W726\_05)’ 순으로 중요도 지수가 높았다. 도출된 30개의 변수는 크게 기업과 근로자의 커뮤니케이션 및 신뢰관계, 근로자의 현재 상태 및 근무 상황, 스트레스, 인재 우대, 기업문화, HR 부서활동 요인이 중요하게 선택되었다.

커뮤니케이션 및 신뢰관계와 관련된 변수 중 1위(W726\_06), 5위(W726\_05)외에도 ‘[커뮤니케이션과 신뢰관계] (1)회사 사정을 직원들에게 소상하게 알림(W726\_01)’,

‘[커뮤니케이션과 신뢰관계] (2)상급자에게 자유롭게 의견을 제시함(W726\_02)’, ‘[커뮤니케이션과 신뢰관계] (4)동료들 간에 서로 신뢰함(W726\_04)’ 등이 선택되었다. 근로자의 현재 상태 및 근무 상황과 관련된 변수는 ‘2016년 연간 총 근로소득(W736\_01)’, ‘출생년월(년)(W7\_birthy)’, ‘2016년 주당 평균 초과 근로시간(W735\_02)’과 함께 ‘2016년 공식적인 교육훈련 참여일 수(W711)’, ‘회사 전체 업무진행 이해 수준(W722)’, ‘현재 업무의 성격(W721)’이 도출되었다.

스트레스 변수인 ‘[스트레스] (5)현재 발전 및 승진의 기회가 불공평하다고 생각(W729\_05)’, ‘[스트레스] (6)현재 업무로 인해 무력감이나 피곤함 느낄 때가 많아짐(W729\_06)’도 중요도 지수가 높게 나타났다. 또한 인재우대에 대한 ‘[인재우대] (2)인적자원개발에 대한 명확한 비전 소유(W725\_02)’, ‘[인재우대] (4)인적자원개발을 위한 다양한 교육훈련 방법 모색(W725\_04)’, ‘[인재우대] (3)인재의 중요성 강조(W725\_03)’, ‘[인재우대] (1)우수한 인재 우대(W725\_01)’ 등의 변수가 선택되었다.

10위에서 30위 사이에는 기업문화에 대한 10개의 변수가 포함되어 있었는데, ‘[기업문화] (1)변화와 새로운 시도를 장려함(W727\_01)’, ‘[기업문화] (11)과업수행에 필요한 전문지식과 능력을 중요하게 생각(W727\_11)’, ‘[기업문화] (7)공식적인 절차, 규칙 및 방침을 중시(W727\_07)’, ‘[기업문화] (6)팀워크를 더욱 중요시(W727\_06)’, ‘[기업문화] (5)인화단결과 일체감을 중시(W727\_05)’, ‘[기업문화] (4)가족과 같은 조직 분위기가 형성(W727\_04)’, ‘[기업문화] (2)혁신에 대한 적절한 보상이 이루어짐(W727\_02)’, ‘[기업문화] (12)직무수행능력과 실적을 중심으로 한 평가가 이루어짐(W727\_12)’, ‘[기업문화] (3)성실한 사람보다 창의적인 사람을 우대함(W727\_03)’, ‘[기업문화] (9)서열의식을 강조하는 조직분위기가 형성(W727\_09)’ 등이다. 이 외 ‘[HR부서 역할] (3)인사제도 개선과 실행에 주도적 역할’도 선택되었다.



[그림 4] 랜덤포레스트 분석 결과(조직몰입)

## V 논의 및 결론

본 연구는 가장 최근에 수집된 인적자본기업패널 7차년도 자료를 활용해 기업 구성원의 직무만족 및 조직몰입을 이질적인 잠재프로파일로 구분하고, 머신러닝 기법을 적용해 직무만족 및 조직몰입 잠재프로파일을 예측하고 중요한 영향을 미치는 요인을 탐색하였다. 특히 본 연구는 4가지 머신러닝 기법(LASSO, Ridge, Randomforest, SVM)을 모두 활용해 직무만족 및 조직몰입 잠재프로파일을 예측하고 각 기법의 예측정확도를 비교하여 모형 성과가 좋은 기법을 찾기 위한 시도를 하였다. 또한 예측정확도가 가장 높은 머신러닝 모형에서 상대적으로 주요한 설명변수를 탐색하였다. 주요 연구결과를

제시하면 다음과 같다.

첫째, 잠재프로파일 분석을 통해 기업 구성원의 직무만족 및 조직몰입 잠재프로파일을 분류한 결과, 직무만족과 조직몰입 각각 3개의 이질적인 잠재프로파일로 구분되었다. 도출된 3개의 직무만족 잠재프로파일의 특징에 따라 '중간 직무만족 집단(계층 1)', '낮은 직무만족 집단(계층 2)', '높은 직무만족 집단(계층 3)'으로 명명하였다. 조직몰입도 마찬가지로 3개 잠재프로파일의 특징을 파악하여 '낮은 조직몰입 집단(계층 1)', '높은 조직몰입 집단(계층 2)', '중간 조직몰입 집단(계층 3)'으로 명명하였다. 조직몰입에서 '높은 조직몰입 집단(계층 2)'에 속하는 비율은 43.7%로 높은 편이었으나, 직무만족의 '높은 직무만족 집단(계층 3)'에는 9.5%만이 속했다.

또한 직무만족이 높은 집단의 경우 직무만족 3문항 평균이 4.081~5.000으로 상당히 높은 편이었으나, 조직몰입이 높은 집단은 조직몰입 4문항의 평균이 3.568~4.144로 상대적으로 높은 수준은 아니었다.

둘째, 잠재프로파일 간 직무만족, 조직몰입 정도에는 통계적으로 유의미한 차이가 있는 것으로 나타났다. 이러한 점을 비추어보 때, 기업 구성원의 직무만족, 조직몰입에 대한 연구를 진행할 때 평균을 사용해 한 집단으로 묶는 것 보다 이질적인 집단으로 분류하여 집단에 따른 차별적인 방안을 강구해야 될 필요가 있음을 보여준다.

셋째, LASSO, Ridge, Randomforest, SVM을 활용해 직무만족 및 조직몰입 잠재프로파일을 예측한 결과, 직무만족과 조직몰입 모두 Randomforest의 예측정확도가 가장 높은 것으로 나타났다. 또한 LASSO, Ridge와 Randomforest의 예측정확도 차이는 크지 않았으나, SVM은 다른 기법에 비해 예측정확도가 낮았다. 특히 SVM은 비율이 높은 집단은 가장 잘 예측했으나, 비율이 약 10% 이하로 낮은 집단은 거의 예측하지 못하는 것으로 나타났다. 이는 SVM의 경우 분류 정확도가 높아 많이 사용되고 있지만, 집단 별 비율이 상당히 차이가 나는 불균형 데이터를 예측할 때는 비율이 높은 집단으로 편향된 분류를 하는 경향이 있어 비율이 낮은 집단을 분류하는 정확도가 매우 낮아진다는 단점으로 비롯한 것이라고 볼 수 있다(김은경, 전명식, 방성완, 2015). 그러므로 집단 별 비율의 차이가 큰 불균형 데이터일 경우, 상대적으로 정확하게 예측한 Randomforest, LASSO 등을 적용해볼 필요가 있으며, 김은경 외(2015)가 제안한  $WL_1$ -norm SVM 등 발전된 SVM을 적용해볼 수 있다. 이처럼 자료에 따라 머신러닝 기법의 모형 성과가 다를 수 있기 때문에, 머신러닝 기법을 활용하고자 하는 연구자들은 여러 기법 중 결과변수를 가장 잘 예측하는 기법을 탐색하는 단계를 권장하는 바이다.

넷째, Randomforest를 적용해 직무만족 및 조직몰입 잠재프로파일 예측에 어떠한 설명변수가 상대적으로 중요한지를 각각 탐색한 결과, 직무만족의 경우 임금과 관련된 변수(2016년 연간 총 근로소득)가, 조직몰입은 조직과의 신뢰관계와 관련된 변수(회사 경영진을 믿고 따를 수 있음)가 가장 중요한 영향요인으로 선택되었다. 또한 순위에서는 차이가 있으나 직무만족 및 조직몰입의 중요도 지수 4위 안에는 동일한 변수가 선택되었다.

임금 수준은 직무만족에 영향을 미치는 변수 중 결정적인 요소로 간주되어 온 만큼 임금으로 인해 직무만족이 낮아지거나 혹은 조직몰입에 방해되거나 이직까지 이어지는 것을 방지하기 위해 기업차원에서도 관심을 가질 필요가 있다(주사총, 이정언, 2016). 예를 들어, 유연한 성과급 제도를 활용하여 개인의 성과나 노력에 따라 임금이나 추가적인 보너스 등을 받을 수 있는 장치를 마련할 수 있다.

다섯째, 출생년도, 주당 평균 초과 근로시간도 직무만족과 조직몰입에서 모두 중요도 지수가 상위인 것으로 나타났다. 출생년도 즉, 연령에 따라 직무만족이나 조직몰입 수준이 다르다는 것은 선행연구에서도 밝혀져 왔으며 대체적으로 연령이 높을수록 직무만족이나 조직몰입 수준이 높다고 하였다(정효채, 석진홍, 박우성, 2013). 그러나 연령에 따라 직무만족이나 조직몰입 수준 차이가 존재하는 지를 파악하는 게 중요한 것이 아니라, 직무만족 및 조직몰입 향상을 위한 교육 및 프로그램 등이 기업 구성원의 연령에 따라 차별적인 내용이 추가적으로 구성될 필요가 있다는 것이다. 연령에 따라 현재 관심사나 정서적인 반응을 이끌어 내는 주제 등이 다를 수 있기 때문에 체계적으로 연령별 직무만족 및 조직몰입 프로그램이 진행되어야 할 것이다. 주 52시간 초과 근로시간을 법으로 제재하고 있지만 집배원이 과로사를 하고(노컷뉴스, 2019. 09. 12), 최저임금이 높아지자 특성화고 학생들을 현장실습, 조기취업 등을 통해 초과근무를 강요받는 현실이다(오마이뉴스, 2019. 8. 14). 법 제정보다 중요한 것이 기업 혹은 조직에서 이를 적용하는 것이며, 기업 구성원의 근로시간을 효율적이고 체계적으로 관리하여 근로시간을 초과하지 않는 선에서 생산성을 높일 수 있는 방안을 모색해야 할 것이다(배규식, 2013).

여섯째, 커뮤니케이션과 신뢰관계와 관련된 변수는 조직몰입과 직무만족에 중요한 변수로 선택되었다. 특히 회사 경영진에 대한 신뢰, 평가나 보상에 대한 공정성, 구성원과 회사 사정에 대한 공유와 같이 회사의 투명성과 공정성은 기업 구성원이 조직을 신뢰하고 몰입할 수 있도록 해준다. 이처럼 기업과 기업 구성원 간에는 투명하고 공정한 기업 운영을 통해 신뢰관계를 유지하고, 기업과 구성원, 구성원 간의 커뮤니케이션이 활발하고 자유롭게 이루어질 수 있는 환경을 조성할 필요가 있다. 또한 기업의 문화도 중요한

변수로 선택되었는데, 무려 직무만족은 8개, 조직몰입은 10개의 기업문화 변수가 도출되었다. 기업마다 달성하고자 하는 목표, 비전 등이 다를 수 있으므로 기업의 특성, 구성원 요구 등을 고려하여 문화를 함께 형성해가는 것이 중요하다.

일곱째, 기업 구성원이 느끼는 스트레스는 직무만족과 조직몰입에도 주요한 영향을 미치는 것으로 나타났다. 직무(업무)를 수행하면서 느끼는 피곤함이나 무력감으로 인한 스트레스뿐만 아니라 조직 내에서 겪는 부당한 대우로 인한 스트레스는 직무만족과 조직몰입 향상을 위해 중요하게 해소해야 할 부분이다. 특히 발전 및 승진의 기회에 대한 불공평을 경험하는 것은 조직에 대한 신뢰를 저해할 수 있기 때문에 기업은 기업 구성원의 승진 기회나 혜택을 투명하게 제공해야 한다. 이 외에도 임금, 평가, 인사이동 등 여러 부분에서 부당한 문제가 발생하고, 이는 직접적인 스트레스 유발 요인이 될 수 있기 때문에 이와 관련된 의사결정 과정에서 절차상의 공정성이 유지되어야 한다(민경호, 조국행, 2002).

여덟째, 임금, 커뮤니케이션과 신뢰관계, 인재우대, 기업문화 등 요인은 직무만족과 조직몰입에 공통적으로 중요한 요인이었으며, 이와 관련된 다수의 변인이 선택되었다. CART 알고리즘을 활용해 조직몰입에 영향을 미치는 요인을 분석한 최영준 외(2014)의 연구에서도 본 연구와 동일한 요인들이 중요하게 도출되었다. 그러므로 후속연구에서 특정 변수가 직무만족 및 조직몰입에 영향을 미치는 요인을 검증하고자 할 때, 임금, 커뮤니케이션과 신뢰관계, 인재우대, 기업문화 등을 통제할 필요가 있음을 보여준다.

본 연구는 기업 구성원의 직무만족과 조직몰입의 수준에 따라 집단을 나눈다고 했을 때, 이를 임의적으로 구분하지 않고 잠재프로파일 분석을 활용하여 기준에 따라 이질적인 잠재집단을 도출해냈다는 데 의의가 있다. 또한 머신러닝 기법별 장단점이 있고, 각 자료를 잘 구분하는 머신러닝 기법은 다를 수 있기 때문에 다양한 머신러닝 기법을 비교하여 잠재집단 분류를 잘 수행하는 기법을 탐색하고, 그 중 가장 예측력이 높은 기법을 활용해 영향요인을 탐색했다는 데 의의가 있다. 이는 사회과학 분야에서도 머신러닝 기법 적용 시 의미 있는 분석을 위해 잠재프로파일 분석 등을 함께 활용하고, 다양한 머신러닝 기법을 적용해 연구자의 자료, 표본에 가장 적합한 기법은 무엇인지 탐색하는 과정이 필요하다는 것을 시사한다.

## 참고문헌

- 김보인(2017). 인적자본기업패널 자료 분석을 통한 한국기업종업원의 만족도 분석. 사회과학연구, 36(1), 1-19.
- 김은경, 전명식, 방성원(2015). 불균형 자료의 분류분석을 위한 가중 L1-norm SVM. 응용통계 연구, 28(1), 21-33.
- 노컷뉴스(2019. 09. 12). '주 90시간' 일하는 집배원들은 추석이 설레지 않는다. 2019년 9월 15일 인출.
- 문찬주, 남인혜, 이영선, 이지영, 한수현, 정동욱(2019). 기업 내 근로자의 전공 다양성이 기업성과에 미치는 영향 분석. 직업능력개발연구, 22(2), 91-127.
- 민경호, 조국행(2002). 공정성이 조직몰입 및 직무만족에 미치는 영향에 관한 실증연구. 인사관리연구, 26(3), 79-100.
- 박광오, 이은령, 정대현(2018). 조직특성과 정보시스템특성이 직무만족에 미치는 영향-비전공유의 조절효과 분석-. 경영과 정보연구, 37(3), 115-130.
- 박민수, 김태현, 조은석, 김희발, & 오히석. (2014). R 을 이용한 별점화 축소추정 기법 비교연구: 요크셔 돼지 산자수와 SNP 에 대한 적용 사례. 농업생명과학연구, 48(3), 147-155.
- 박민지, 이신혜(2017). 잠재프로파일 분석을 통한 청소년 직업가치 유형분류 및 영향요인 검증. 한국청소년연구, 28(1), 235-265.
- 박성계(2000). 스포츠산업 경영/스포츠 센터내 커뮤니케이션이 조직유효성에 미치는 영향. 한국체육학회지-인문사회과학, 39(1), 716-726.
- 박영재, 김계영, 장석우(2013). 다중 클래스 SVM을 이용한 트래픽의 이상패턴 검출. 한국 산학기술학회논문지, 14(4), 1942-1950.
- 박진수, 이성희(2019). 인재우대 활동과 직무 스트레스가 조직몰입에 미치는 영향: 제조업 관리직 팀장의 인식을 중심으로. 상품학연구, 37(2), 157-163.
- 배규식(2013). 한국의 장시간 노동과 근로시간 단축. 노동리뷰, 10, 7-18.
- 배성은, 이건만(2014). 교육훈련내용의 직무연관성이 조직몰입에 미치는 영향에 관한 구조적 분석. 미래교육연구, 27(2), 195-226.
- 송상윤(2015). 예대금리차 결정요인 모형의 예측력 비교 연구-Ridge, LASSO 및 Elastic Net 방법론을 중심으로-. 금융지식연구, 13(3), 41-65.
- 신태수(2010). 종단프로파일분석과 군집분석을 이용한 잠재집단 연구. 교육평가연구,

- 23(3), 641-664.
- 오마이뉴스(2019. 08. 14). 제조업 대기업, 좋은 일자리 맞습니까?. 2019년 09월 15일 인출.
- 오미애, 최현수, 김수현, 장준혁, 진재현, 천미경(2017), 기계학습(Machine Learning)기반 사회보장 빅데이터 분석 및 예측모형 연구(연구보고서 2017-46). 세종:한국보건사회연구원.
- 오영미(2004). 비서의 조직 커뮤니케이션 만족이 직무만족과 조직몰입에 미치는 영향. 이화여자대학교 정보과학대학원 석사학위 청구논문
- 오인수, 김광현, Dornold, T. C., 황종오, 유태용, 박영아, 박량희(2007). 직무만족, 조직몰입, 성과, 이직의도 간의 관련성: 문헌고찰 및 메타분석. 인사·조직연구, 15(4), 43-86.
- 우성미, 최명빈(2018). HR부서의 기능과 조직성과의 관계에서 교육훈련 질의 종단적 매개효과. 제7회 인적자본기업패널 학술대회 자료집, 549-570.
- 유진은(2015). 랜덤 포레스트: 의사결정나무의 대안으로서의 데이터 마이닝 기법. 교육평가연구, 28(2), 427-448.
- 윤장열(2010). 멘토링이 직원의 조직몰입과 직무만족 및 이직의도에 미치는 영향에 관한 연구: 여행사 규모 및 판매형태의 조절효과 검증. 세종대학교 일반대학원 박사학위 청구논문.
- 이영면, 백승규(2010). 조직구성원의 여가만족이 직무만족 및 조직몰입에 미치는 영향. 조직과 인사관리연구, 34(1), 25-62.
- 이용탁(2007). 인적자원개발을 통한 직무능력향상이 직무만족 및 조직몰입에 미치는 영향. 인적자원관리연구, 14(2), 143-160.
- 이용탁(2017). ICT 기업의 조직문화가 조직구성원 직무만족과 조직몰입에 미치는 영향에 관한 연구. 경영과 정보연구, 36(4), 149-166.
- 장혜정, 진인선(2014). IT 기업 근로자 숙련도, 교육훈련 적절성, 인재우대, 조직몰입의 인적자원경쟁력 예측 연구. 학습과학연구, 8(1), 81-99.
- 전현정, 정혜원(2017). 경향점수 매칭을 활용한 기업의 교육훈련이 근로자의 직무만족 및 조직몰입에 미치는 영향 분석. 교육학연구, 55(1), 97-123.
- 정효채, 석진홍, 박우성(2013). 연령이 직무만족에 미치는 영향- 잠재성장모형을 이용한 종단연구. 노동정책연구, 13(2), 67-93.
- 주사총, 이정언(2016). 임금만족이 직무만족과 이직의도에 미치는 영향. 한국콘텐츠학회

- 논문지, 16(10), 693-700.
- 주영하, 선민정(2018). 조직문화가 종사자의 직무만족에 미치는 영향 : 의사소통의 매개 효과를 중심으로. 디지털융복합연구, 16(7), 153-164.
- 최영준, 심원술, 백승현(2014). 데이터마이닝을 활용한 이직의도와 조직몰입의 결정요인에 대한 연구. 한국시뮬레이션학회논문지, 23(1), 21-31.
- 황매향, 최희철, 조흥순(2013). 임금, 조직몰입, 직무만족의 구조관계에서 조직몰입의 매개 효과: 다집단 분석. HRD 연구 (구 인력개발연구), 15(4), 157-182.
- Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2018). randomForest: Breiman and Cutler's Random Forests for Classification and Regression. R package version, 4.6-14.
- Calle, M. L., & Urrea, V. (2010). Letter to the editor: stability of random forest importance measures. Briefings in bioinformatics, 12(1), 86-89.
- Friedman, J. H., Hastie, T. J., & Tibshirani, R. J. (2019). glmnet: lasso and elastic-net regularized generalized linear models, 2010b. R package version, 2.0-16.
- Meyer, D., Dimitriadou, e., Hornik, K., Weingessel, A., & Leisch, F. (2019). e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R package version, 1.7-1.
- Mowday, R. T., Porter, L. W. and Steers., R. M.(1982). Employee-organization linkages: The psychology of commitment, absenteeism, and turnover. New York: Academic Press.
- Muthén, B. (2004). Latent variable analysis. The Sage handbook of quantitative methodology for the social sciences, 345(368), 106-109.
- Muthén, L. K., & Muthén, B. O. (1998-2019). Mplus User's Guide(8th ed.). Los Angeles, CA: Muthén & Muthén.
- van Buuren, S., & Groothuis-Oudshoorn, K. (2011) Mice: multivariate imputation by chained equations in R. J Stat Softw, 45(3), 1-67.