

근로자들의 이직 의도에 영향을 주는 요인에 관한 실증연구: 공공 데이터베이스와 의사결정나무 기법을 중심으로

Empirical Analysis of Influential Factors Affecting Domestic Workers' Turnover Intention: Emphasis on Public Database and Decision Tree Method

고 건 우 (Geo Nu Ko) 경희대학교 무역학과 학부생

조 현 진 (Hyun Jin Jo) 경희대학교 수학과 학부생

이 건 창 (Kun Chang Lee) 성균관대학교 글로벌 경영학과/삼성융합의과학원(SAIHST) 융합의과학과 교수, 교신저자

요 약

본 연구는 한국고용정보원 공공디비 자료에 의사결정나무 기법을 적용하여 근로자들이 이직을 하려고 하는 요인을 분석하고자 한다. 이를 통하여 근로자들의 고용유지 강화방안을 제시한다. 분석대상은 근속기간이 4개월 이상으로 수습 기간이 지난 정규직 근로자를 대상으로 하였고, 분석목표는 해당 근로자들의 이직준비 여부에 영향을 주는 의사결정 규칙을 도출하는 것으로 하였다. 타겟 분석 데이터는 한국고용정보원에서 발간된 수집한 「2017년 대졸자 직업이동경로조사」를 사용하였다. 분석결과 ‘교육 수준과 일의 수준의 일치 정도’, ‘개인 발전 가능성’, ‘직무 관련 교육 및 훈련’, ‘승진 제도’, ‘임금 및 소득’, ‘일자리에 대한 사회적 평판’, ‘고용 안정성’ 등이 이직준비 여부에 영향을 주는 주요요인으로 나타났다.

키워드 : 데이터 마이닝, 의사결정나무, 이직, 고용유지, 근속기간

I. 서 론

기업이 경쟁우위를 확보하고 지속 가능한 이윤을 마련하기 위해서는, 기업 가치의 생산 주체인 인적자원에 대한 장기적인 유지·관리가 필요하다. 기업의 주체인 인적자원의 관리는 기업 경쟁력의 핵심으로 판단되며, 기업들은 인적자원의 역량 강화에 지속적인 투자를 하고 있다(김민경, 나

인강, 2012). 이직이란 근로자가 기업을 떠나거나, 근로의 조건으로 받을 임금이 상실했을 것을 의미한다(Mobley, 1982). 선행연구들의 결과, 기업의 이직률이 낮으면 기업의 조직성과를 긍정적으로 평가하고 있다(Hancock *et al.*, 2013; 이정현, 김동배, 2007). 기업성과에 부정적인 영향을 주는 이직이란 직원들 가운데 우수 인력들의 이직이다(Kwon and Rupp, 2013). 업무 능력이 우

수한 근무자들은 노동시장에서 넓은 선택권을 가지게 되어서 이직 의도가 높아질 수 있다. 반면, 기업에서 이직자가 발생하면 이직자의 업무 능력에 맞는 우수 인력을 확보하거나, 신입 사원을 채용해 교육하여 이직자 수준의 업무 능력으로 끌어 올려야 한다. 이때 인력 확보를 위한 구인과 교육에 비용이 발생하게 되고, 이직자에 대해 이미 실시한 교육 및 투자비용은 회수되지 않아서, 기업 경영에 부정적인 영향을 준다(박인규, 2010). 이에 따라 기업은 이직으로 발생하는 비용을 줄이고 교육 및 투자비용의 회수를 위하여, 이직에 대한 원인을 파악하고 대응할 필요가 있다. 한편 인적자원관리 분야의 연구들은 고용 안정성을 제공하는 조직 제도가 조직의 생존과 경쟁력을 유지하는 중요한 요인이라고 밝혔다(김진희, 2009a).

2019년 1월 기준의 국가통계포털 자료에 따르면, 전체 근로자의 이직률이 2018년 대비 18.36% 상승했다. 이직자 중 자발적 이직률도 전년 대비 12.72% 상승했다. 전체 이직자의 수는 137,207명 증가했으며 자발적 이직자는 33,061명 증가했다(통계청, 2019).

이직 의도에 관한 선행연구는 여러 분야에서 진행이 되었다. 경제학적 관점에서 근로자는 재직 중인 기업보다 다른 기업이 더욱 많은 경제적 이점을 제공할 경우, 이직을 준비하게 된다고 한다(Forrest *et al.*, 1977). 사회 심리학의 기대이론에 의하면, 기업 근로자가 기업에 대해 기대하는 바와 실재가 맞지 않으면 이직이 발생한다고 한다(Steers *et al.*, 1979). 직무특성이론에 따르면, 근로자들은 직무 적합성, 직무 중요성, 직무 정체성, 기술 다양성을 통해 직무에 대한 의미를 부여받고, 직무 자율성과 피드백을 통하여 업무와 책임감을 지각한다. 또한, 이러한 요인들은 이직 의도에 영향을 준다고 한다(Hackman and Oldham, 1975).

본 연구는 실질적 기업의 비용 및 투자비 회수에 대한 관점에서 「2017년 대졸자 직업이동경로조사」 데이터를 이용하여 근로자의 이직준비 여부의 결정규칙을 도출하고자 한다. 또한, 이를 이용

하여 이직준비 여부에 영향을 미치는 주요요인을 파악하고 근로자의 이직을 방지하여, 고용을 강화하는 방안을 제시하고자 한다. 데이터의 전처리에서 경제학적, 사회 심리학적, 직무특성 이론적 관점에 따라 이직 의도에 영향을 끼치는 독립변수 33개를 추출하였고, 근속 기간이 일반적 수습 기간인 3개월을 초과한 정규직을 통합 데이터로 구축하였다.

이직에 영향을 주는 여러 요인은 근속 기간에 따라 다르게 분석될 수 있다. 따라서 근로자의 이직준비 의도에 영향을 주는 요인을 파악할 때, 분석대상을 근속기간별로 구분한다면 근로자의 고용유지 강화와 조직이탈의 방지에 도움이 되는 방안들이 다양한 관점으로 개발될 수 있다. 즉, 근속 기간이 4개월 이상인 근로자에 관한 분석에서는 3개월의 수습 기간이 지난 정규직 근로자들의 전반적인 특성을 반영한 이직준비 요인들을 파악할 수 있다. 반면, 근속기간별로 구분된 근로자에 관한 분석에서는 기간별 특성을 반영한 이직준비 요인들을 파악할 수 있다. 본 연구에서는 근속기간별로 구분된 근로자의 이직준비 요인을 분석하기 위해 4~6개월, 7~12개월, 13~24개월, 25개월 이상으로 구분한 후에 연구를 진행하고, 근속기간별 구분된 고용유지 방안을 제시하고자 한다.

본 연구의 구성은 다음과 같다. 제Ⅱ장에서는 근로자들의 이직준비에 영향을 주는 요소들에 대해 분석한 기존 문헌을 살펴본다. 제Ⅲ장에서는 연구에서 실행한 연구방법과 데이터 마이닝 기법을 사용하여 이직준비 여부를 판단하는 예측모형이 형성되는 실험 과정에 관해 설명한다. 제Ⅳ장에서 연구결과를 설명하고, 제Ⅴ장에서 근로자의 고용을 유지하는 데 도움이 되는 방법들을 제시하고자 한다. 마지막 제Ⅵ장에서는 연구의 요약과 시사점과 한계점을 다룬다.

Ⅱ. 선행연구

이직 의도에 영향을 주는 요소에 관해 분석한

기존 연구들은 다음과 같다. 김경범 등(2016)은 패 널 조사 자료를 분석하여 청년 취업자 이직준비 영향요인을 직무 관련 특성, 인구 사회학적 특성 으로 나누어 연구하였다. 그 결과 인구 사회학적 특성에서 나이, 성별, 월 소득, 최종학력, 고용형태 가 유의하였으며, 직무 관련 특성에서 고용 안정 성, 교육수준 대비 업무 난이도, 업무와 전공 일치 도, 자기발전 추구, 일의 자율성, 일의 권한 그리고 직장 내 인간관계가 유의하다고 하였다. 여기에서 직무환경과 근무조건에 불만족할수록 이직준비 율이 높아진다고 하였다. 김보인(2017)은 이직 의 도가 조직분배공정성과 상사의 배려행동에는 음 의 상관관계를 지니고, 직무 모호성과 직무갈등에 는 양의 상관관계를 지닌다고 하였다. 반면, 상사 의 구조주도 행동과 조직 공식성은 이직 의도에 영향을 주지 않는다고 하였다. 또한, 이직 의도가 근로자의 특성과의 관계에서 조사 대상자의 국가 나 나이의 차이와는 유의하지 않지만, 성별과 직 급별 차이와는 유의하다고 하였다. 김진희(2009b) 는 제조업종 중소기업 인사담당 관리자와 종업원 을 조사하여, 인적자원관리가 조직구성원의 행동 에 주는 영향을 분석하였다. 여기에서 인사제도 내재화와 직무몰입은 이직준비율을 낮추지만 역 할 피로는 높인다고 하였다. 성지미, 안주엽(2016) 은 대졸 청년층을 대상으로 이직 의도와 일자리 만족도의 관계를 조사하였다. 그 결과 일자리나 일에 대한 만족도에 따라 이직준비 의사나 실제 이직률에 영향을 미친다고 하였다. 여기에서 조사 자의 답변이 일의 수준도, 요구 기술 수준 그리고 전공 일치도 항목에서 부정적일 경우에 이직 의도 와의 관계가 통계적으로 유의한 효과를 준다고 하 였다. 또한, 이직 의사가 실제로 이직할 확률을 크 게 높인다고 하였다. 이진복, 곽원준(2017)은 직장 상사의 이직 의도가 부하직원의 이직 의도에 어떠 한 영향을 주는가를 조사하여 상사의 이직 의도가 부하직원의 이직 의도에 영향을 미치는 것으로 확 인하였다.

다음으로 특정 직종으로 한정하여 이직 의도

영향요인을 분석한 연구는 다음과 같다. 이채호 (2019)는 보육교사의 이직 의도 영향요인을 분석 하여 직무만족도, 직무 스트레스, 직무환경, 조직 문화 등이 있다고 언급했다. 윤은자 등(2016)은 간 호사의 이직 의도 영향요인을 분석하였는데, 이직 준비에 가장 큰 영향을 미치는 요인이 정서적 소 진이라고 했다. 따라서 정서적 소진을 줄이는 중 재를 통해 이직을 줄일 수 있을 것이라 언급했다.

근속 기간이 고용유지에 미치는 영향 또는 서로 의 상관관계를 분석한 연구는 다음과 같다. 먼저 고용 안전성(employment security)이란 고용 안정 성(employment stability)의 측정지표인 근속 기간, 고용유지율에 근로자의 직관을 함께 측정하여 구 하거나(OECD, 1997), 경제 전반의 실직, 실직 이 후 고용 가능성, 첫 일자리를 찾는 어려움 등을 지표로 한다(Auer and Cazes, 2000). 근속 기간이 길수록 고용 안전성은 증가하지만, 고용이 오래 유지되어도 근로자는 고용상태를 안전하다고 생 각하지 않는다.(OECD, 1997). Auer and Cazes(2000) 는 1990년대 일본과 미국 및 유럽 국가 등 16개국 을 대상으로 평균 근속 기간·년수, 직장 유지율 의 변화를 보여준다. 미래에는 점차 장기 고용이 감소할 것으로 예측했고, 고용 불안정성이 성별, 연령 등의 개인적 속성에 영향을 받는다고 언급했 다. 이와 유사하게 근로자의 특성별로 고용 안정 성 변화를 분석한 연구로 Burgess and Rees(1998), Heisz(2005), Lee and Eckert(2002) 등이 있으며, 이 연구들은 성별, 연령, 교육수준별 평균 근속 기간 의 차이를 보여준다. 이를 통해 모든 근로자가 비 슷한 근속 기간을 가지지 않고, 개인적 특성에 따 라 다양한 근속 기간을 가진다는 것을 알 수 있다. OECD(1997)에 따르면 OECD 국가의 고용유지 기 간 차가 매우 크다. 고용유지율이 감소한 국가도 있고, 고용 안정성이 유지되어 고용유지율이 감소 하지 않는 국가도 존재한다. 즉, 모든 국가의 고용 안정성이 감소하는 것은 아님을 알 수 있다. Burgess et al.(1997)는 기업의 크기가 클수록 근속 기간이 따라서 증가하는 것은 상대적으로 대규모

기업이 직무 교육에 더 많은 투자하기 때문이라고 해석한다. 또한, 고용 보호 규제가 대규모 기업보다 소규모 기업에 더 관대하게 적용되는 것도 영향을 줄 수 있다고 하였다. Cazes(2003)는 근속 기간이 길수록 직업 만족도가 높아지지만, 낮은 보수와 지위를 가진 직업의 경우에는 직업 만족도가 낮다고 했다. 즉, 근속 기간과 직업 만족도는 양의 상관관계를 가지지만 필연적이지는 않다는 것을 의미한다. 또한, 상근직(full-time job)과 시간제 근로자(part-time job)에 상관없이 적절한 보수와 작업 환경 등은 직업 만족도에 큰 영향을 주며, 안전한 일자리가 곧바로 좋은 일자리로 해석할 수는 없다고 언급했다.

기존 연구와 본 연구의 차이점으로 첫째, 기존 연구들은 주로 특정 직종별로 이직에 영향을 미치는 요인을 분석하거나, 근속 기간 자체가 이직에 영향을 미치는지에 관한 연구들이 대부분이었다. 본 연구는 데이터를 근속기간별로 범주화하여 이직 의도 영향요인을 분석한다는 점에서 차이가 있다. 이를 통해 수습 기간이 지난 근로자들의 이직 의도 영향요인과 함께 근속기간별 영향요인을 비교한다는 차별점을 가진다. 둘째, 기존 연구에서는 주로 구조 방정식이나 다중회귀분석과 같은 통계분석법을 활용하였기 때문에 분석할 수 있는 변수의 개수가 한정적이었다. 본 연구는 의사결정나무 알고리즘을 활용한 데이터 마이닝 기법을 통하여 변수의 수를 제한하지 않고 모두 활용한 분석이 가능하고, 변수들의 가중치를 비교해 더 영향력이 있는 변수를 추출한다. 동시에 분석 결과를 바탕으로 근로자의 고용유지를 강화할 실용적인 방안을 제시한다는 점에서 차별성이 있다.

III. 연구설계 및 분석 방법

3.1 연구설계

<그림 1>과 <그림 2>는 데이터 마이닝 기법을 활용한 이직준비 여부에 대한 전반적인 결정규칙

이 도출되는 과정을 보여준다. 데이터 마이닝의 정의는 많은 양의 데이터로부터 유의미한 패턴이나 규칙을 발견하기 위해 데이터를 분석하는 것이다(Adriaans and Zantinge, 1996; Chawla, 2009). 향상된 데이터 분석의 목적은 데이터로부터 의사결정을 하는 것이다. 그 과정은 ‘데이터에서 정보, 정보에서 지식, 지식에서 의사결정’으로 설명되며, 데이터에서 의사결정까지 도출하는 과정의 주요 기술로서 데이터 마이닝 기법이 사용된다(Brethenoux *et al.*, 1996).

로지스틱 회귀 분석이나 구조모형 분석과 같은 데이터 마이닝의 연구방법은 목표변수에 상호작용하는 효과를 확인하기 어렵다(정석원, 2014). 또한, 추론에 사용되는 신뢰구간을 구하거나 가설검정의 해석에 있어 어려움이 있다. 반면, 의사결정나무 알고리즘은 결과를 도표화 하여 집단을 분류, 세분화하기 때문에 결과 분석에 있어 이해하기 쉽다. 본 연구에서 변수들의 척도 수준은 혼재해있기 때문에 결과의 신뢰성에 문제가 생길 수 있지만, 의사결정나무 알고리즘은 척도 수준이 혼재되어있어도 결과의 신뢰성에 결함이 발생하지 않는다(이철호 등, 2009). 또한, 의사결정나무 알고리즘은 선형성, 정규성이나 동분산성의 전제 조건이 필요하지 않기 때문에 본 연구에 적합한 방법이다(황정임 등, 2009).

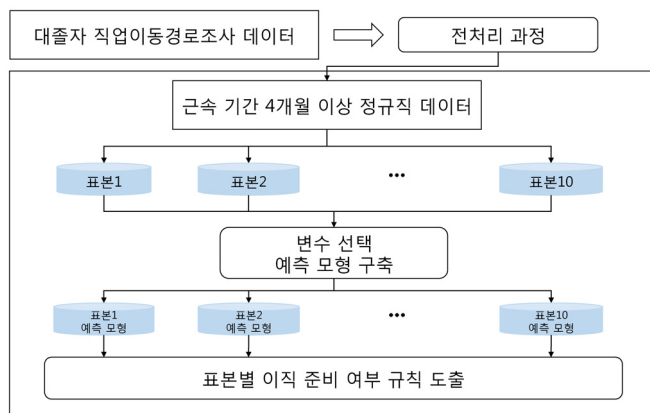
연구 자료로 한국고용정보원 주관으로 수집된 「2017년 대졸자 직업이동경로조사」를 데이터로 사용하였다. 데이터 마이닝 분석을 위해 수집된 자료를 결측값 제거 등의 전처리 과정 후 연구에 적합하게 데이터를 변형한 뒤, 연구에 이용될 초기 독립변수를 선별하였다. 이때, 통합 데이터에는 근속 기간이 4개월 이상인 정규직 근로자의 데이터로 구성된다.

<그림 1>에서 ‘실험 1’과 같이 근속 기간이 4개월 이상인 정규직 근로자를 대상으로 하는 데이터에서 목표변수에 대한 데이터 균형을 위하여 10개 표본을 추출하여 검토하였다. 변수 선택(feature selection) 과정으로 목표변수에 영향력을 미치는

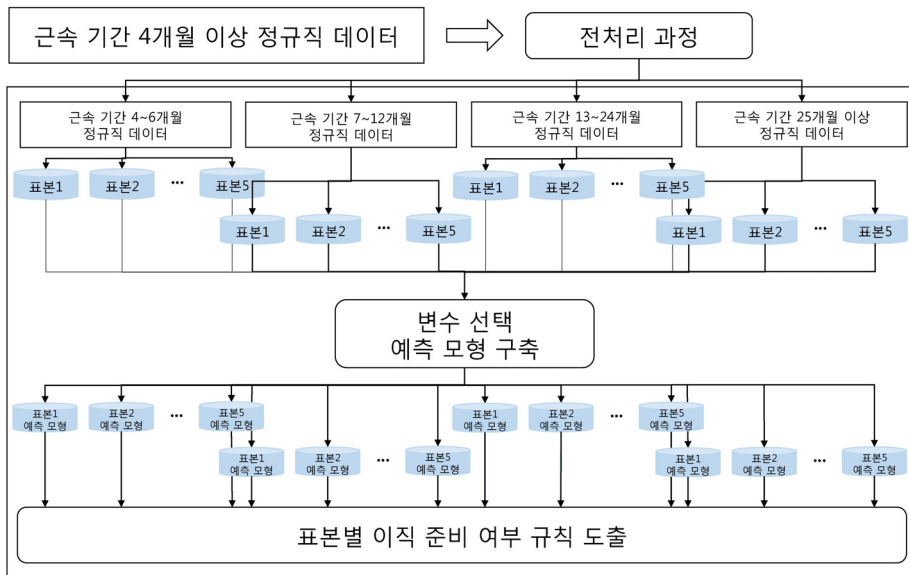
독립변수들을 최종적으로 선택하였고, 분석 방법으로는 R ver 3.6.1에서 의사결정나무(classification and regression trees, CART) 알고리즘을 채택하여 표본별로 예측모형을 만들었다. 알고리즘 내에 복잡성 매개 변수(complexity parameter, CP) 속성값을 조정하면 의사결정나무의 크기를 제어하고 가지치기를 통한 최적의 모형 크기를 선택할 수 있다. 이를 통해 의사결정나무의 깊이나 leaf node의

최소 샘플 수는 자동적으로 결정되며, 표본별로 과적합을 피한 최적의 모형을 도출할 수 있다. 구축된 모형 중 최고예측률을 보인 모형을 통해 근속 기간이 4개월 이상인 정규직 근로자 전반에 대한 이직준비 여부 결정규칙을 도출했다.

다음으로, <그림 2>의 '실험 2'에서 볼 수 있듯이 근속기간별 이직준비에 대한 특성을 파악하기 위해 구간을 '4~6개월', '7~12개월', '13~24개월',



<그림 1> 실험 1 진행 과정



<그림 2> 실험 2 진행 과정

‘25개월 이상’ 4가지로 구분하였다. 근속기간별 근로자의 데이터를 추출하고 ‘실험 1’에서 사용하는 방법으로 근속기간별 근로자의 이직준비 여부의 결정규칙을 도출했다. ‘실험 2’의 예측모형을 만들 때는 ‘실험 1’에서 사용한 의사결정나무(CART) 알고리즘을 활용하였다.

3.2 분석 데이터

한국고용정보원이 주관하여 조사한 「2017년 대졸자 직업이동경로조사」 설문 데이터 18,199건을 분석 데이터로 수집하였다. 수집된 통합 데이터는 2017년 9월 1일 기준으로 설문자의 구직활동 여부, 이직준비 여부, 현재 일자리, 현재 경제활동 상황 등에 관한 정보들을 포함한다. 그 후에 전처리 과정을 통해 7,962건의 데이터를 선별하였다. 데이터를 살펴보면 1,598명이 이직을 준비하였고 나머지 6,364명은 이직을 준비하지 않았다. 여기에서 근속기간이 4개월 이상인 정규직 근로자의 이직준비율은 20.07%임을 알 수 있다. 또한, 근속기간별 데이터의 경우, 4~6개월(915건), 7~12개월(1,588건), 13~24개월(4,351건), 25개월 이상(1,108건) 중 이직준비의사가 있는 사람은 4~6개월(167건), 7~12개월(327건), 13~24개월(881건), 25개월 이상(223건)이며, 이직준비를 하지 않은 사람은 4~6개월(748건), 7~12개월(1,261건), 13~24개월(3,470건), 25개월 이상(885건)이었다. 근속기간별 이직준비율은 4~6개월은 18.25%, 7~12개월은 20.59%, 13~24개월은 20.25%, 그리고 25개월 이상은 20.13%로 분석되었다.

3.3 실험 변수 처리

본 실험을 위해 「2017년 대졸자 직업이동경로조사」의 1,297개의 변수 중 이직준비 여부를 종속변수로 설정했다. 이직준비 여부는 ‘준비함’과 ‘준비 안 함’으로 나뉜다. 나머지 1,296개의 변수 중 실험에 사용될 독립변수를 33개로 추출했다. 이때 인구통계학적, 경제학적, 사회 심리학적, 직무특

성 이론적 기준들을 통해 이직준비 여부에 영향을 미칠 것으로 판단되는 독립변수를 우선으로 추출했다. <표 1>을 통해 33개의 독립변수와 1개의 종속변수를 확인할 수 있다. 변수들의 데이터 값에서 무응답 데이터 혹은 실험에 사용될 수 없는 데이터들은 삭제 처리했다.

데이터 마이닝 기법은 분석에 투입되는 데이터의 유형에 따라 민감하게 반응하여 결과가 나온다. 특히 종속변수의 특정 클래스에 데이터가 집중적으로 분포하면 그 클래스에 대해 학습이 집중되어 특정 클래스의 정확도만 올라간다(Chawla, 2009). 따라서 데이터 구성을 할 때 클래스별 데이터 분포를 고르게 한 후 데이터 마이닝 기법을 적용해야 한다. 이처럼 더 적은 클래스의 수에 맞추어 더 많은 클래스에서 무작위로 샘플링하는 기법을 언더샘플링이라 한다. 본 연구의 종속변수인 이직준비 여부는 ‘준비함’과 ‘준비 안 함’ 2개의 클래스로 구분된다. 전처리를 통해 도출된 근속기간이 4개월 이상인 정규직에 대한 데이터에서, 1,598명의 응답자는 이직준비 여부에 대한 ‘준비함’이라 대답했고 6,364명의 응답자는 이직준비 여부에 ‘준비 안 함’이라 응답했다. ‘준비 안 함’의 클래스의 응답이 ‘준비함’의 약 4배인 것을 확인할 수 있다. 즉, 전체 데이터는 ‘준비 안 함’ 클래스에 편향되어 있다. 이런 문제를 해결하기 위해 상대적으로 응답자 수가 적은 ‘준비함’의 응답 수에 맞춰 ‘준비 안 함’ 클래스에서 10개의 표본에 대한 언더샘플링을 진행했다. 그 결과, 표본별 3,196명(준비함: 1,598명, 준비 안 함: 1,598명)으로 구성된 데이터셋을 연구에 사용할 최종 데이터셋으로 하였다.

근속기간별 데이터도 앞서 진행한 방법과 같이 기간별로 5개의 표본을 추출할 때에 언더샘플링을 실시하였다. 그 결과, 표본별로 4~6개월은 334명(준비함: 167명, 준비 안 함: 167명), 7~12개월은 654명(준비함: 327명, 준비 안 함: 327명), 13~24개월은 1,762명(준비함: 881명, 준비 안 함: 881명) 그리고 25개월 이상은 446명(준비함: 223명, 준비 안 함: 223명)으로 최종 데이터셋을 구성하였다.

〈표 1〉 실험에 사용한 변수 정보

변수명	설명	변숫값	변수명	설명	변숫값		
인구 통계학적	A1	성별	남성/여성	근로 만족도	C1	임금 또는 소득	만족, 보통, 불만족
	A2	연령	숫자 값		C2	고용 안정성	만족, 보통, 불만족
	A3	혼인 여부	미혼, 기혼, 이혼, 사별		C3	근무 환경	만족, 보통, 불만족
	A4	부양 자녀 수	숫자 값		C4	근로시간 및 근무 일수	만족, 보통, 불만족
근로 관련	B1	근로계약 기간 지정 여부	예, 아니요		C5	개인 발전 가능성	만족, 보통, 불만족
	B2	근로시간 형태	전일제, 시간제		C6	인간관계	만족, 보통, 불만족
	B3	주당 정규 근로일	숫자 값		C7	복지 후생 제도	만족, 보통, 불만족
	B4	주당 정규 근로시간	숫자 값		C8	인사체계(승진 제도)	만족, 보통, 불만족
	B5	월평균 휴일 근로	숫자 값		C9	일에 대한 사회적 평판	만족, 보통, 불만족
	B6	월평균 근로소득_만원	숫자 값		C10	일의 자율성과 권한	만족, 보통, 불만족
근로 자아 정체	D1	교육수준 일 수준 일치 정도	낮음, 알맞음, 높음		C11	일자리에 대한 사회적 평판	만족, 보통, 불만족
	D2	일의 수준 대비 본인의 수준 일치 정도	낮음, 알맞음, 높음		C12	하는 일과 자신의 적성, 흥미의 일치	만족, 보통, 불만족
	D3	일과 전공 일치 정도	맞지 않음, 보통, 맞음		C13	직무 관련 훈련 또는 교육 훈련	만족, 보통, 불만족
	D4	전공 지식의 업무 도움 정도	도움 안 됨, 보통, 도움 됨	4대 보험 가입 여부	E1	국민연금 가입 여부	가입 안 됨, 가입됨
	D5	외국어 능력의 필요 정도	요구되지 않음, 보통, 요구됨		E2	건강보험 가입 여부	가입 안 됨, 가입됨
구직 경험	F1	구직활동을 경험 여부	예, 아니요		E3	고용보험 가입 여부	가입 안 됨, 가입됨
	이직 의도	Target	현재 이직준비 여부		준비함, 준비 안함	E4	산재보험 가입 여부

3.4 변수 정렬

선정한 33개의 독립변수 중 예측모형을 구축하는 데 영향을 주지 못하는 변수가 존재할 수 있다. 목표변수의 예측에 도움이 되지 않음에도 예측모형이 만들어질 때 이용되는 변수가 있다면, 예측모형의 정확성을 떨어트릴 수 있다. 따라서 예측모형을 도출할 때 불필요한 독립변수는 제거해야 한다(Dash and Liu, 1997). 변수 선택이란, 이처럼 목표변수의 예측에 영향을 주지 못해 예측모형의

정확성을 떨어트리는 독립변수를 제거하는 과정을 의미한다.

본 연구에서는 변수를 선택하기 위해 정보의 이득비(gain ratio) 알고리즘을 채택하여 독립변수들의 중요도를 평가하였고, 목표변수의 예측에 전혀 영향을 주지 않는 변수는 삭제 처리한 뒤 예측모형을 구축하였다. 또한, 예측모형을 만드는 데 이용될 독립변수들을 래퍼(wrapper) 방법의 하나인 역방향 제거를 통해 선택하였다. 역방향 제거란, 예측모형을 구축하는 때회 가장 가중치가 낮

은 변수를 순서대로 제거하면서 예측모형을 만드는 방식으로, 가중치가 가장 높은 변수 단 한 개가 남을 때까지 예측모형을 반복해 구축하는 방법이다(Witten *et al.*, 2016). 이 과정을 통해서 33개보다 적은 수의 예측모형이 구축되며, 구축된 예측모형 중에서 최고의 정확도를 보여준 예측모형에서의 변수들을 최종으로 선별하였다.

IV. 연구결과

4.1 통합 데이터의 예측모형과 결과

본 연구는 통합 데이터에서 이직 의도에 관한 예측모형을 만들기 위하여 오픈 소스 데이터 마이닝 도구인 R ver 3.6.1을 이용했으며, 의사결정나무 알고리즘으로 예측모형을 구축하였다.

<그림 1>에서는 통합 데이터에서 언더샘플링

방식으로 10개의 표본을 추출하였다. 추출된 표본들을 각각 의사결정나무 알고리즘에서 예측모형의 학습에 사용될 66%의 학습데이터와 만들어진 예측모형을 평가할 34%의 검증데이터 비율로 나누어 실험을 진행하였다. 선정된 33개의 독립변수 중 목표변수에 영향을 주지 않는 변수는 이득비 알고리즘으로 삭제 처리한 뒤, 중요도가 낮은 변수를 하나씩 제거해가며 예측모형의 예측률을 계산하였다. 그 결과 ‘실험 1’에서 표본 10개에 대해 최종적으로 204개의 예측모형이 구축되었다. <표 2>에서는 10개 표본의 평균 예측률을 보여준다. 의사결정나무 기반으로 구축된 예측모형의 평균 예측률이 68.186%로 나타났다.

<표 2> 통합 데이터 예측모형 결과

예측모형	평균 예측률
DT	68.186%

<표 3> 통합 데이터 표본별 예측모형 예측률

순서	표본 1		표본 7		표본 9		표본 6	
	변수명	DT(%)	변수명	DT(%)	변수명	DT(%)	변수명	DT(%)
1	D3	68.93	B6	67.56	B6	67.20	B5	67.03
2	B6	68.56	D3	67.48	D3	68.23	D3	67.84
3	F1	70.78	D4	68.52	F1	67.90	B6	69.02
4	D4	68.65	F1	69.50	D4	68.96	E1	70.52
5	C10	69.12	E1	69.12	E1	68.88	F1	70.60
6	C6	69.58	C6	68.57	C10	68.35	D4	70.03
7	C4	70.21	C4	67.48	C6	68.26	C10	69.06
8	C3	70.13	C9	67.01	C4	68.04	C6	69.00
9	C12	69.11	C3	67.65	C12	70.43	C4	69.81
10	C9	70.34	C10	69.15	C9	69.46	C7	70.12
11	C7	72.17	C7	69.00	C3	69.59	C3	70.31
12	C2	72.52	C12	69.91	C11	69.32	C1	70.17
13	C11	70.86	C2	69.30	C2	70.01	C9	70.16
14	C1	71.05	C1	70.33	C1	70.22	C12	70.55
15	C8	68.77	C11	68.84	C7	70.76	C2	70.59
16	D2	71.01	C13	69.76	C13	70.91	C13	69.91
17	C13	70.13	C8	68.53	C8	69.57	C11	70.94
18	C5	66.21	C5	66.70	D2	67.93	C8	69.46
19	D1	64.66	D1	67.12	C5	66.83	D2	67.83
20			D2	62.94	D1	63.00	C5	68.64
21							D1	68.19

<표 3>에서는 10개 표본 가운데 최고예측률을 보인 예측모형을 지닌 표본 4개에 관한 정보를 보여준다. 여기에서 독립변수 순서가 낮을수록 목표 변수 'Target'에 대한 가중치가 낮은 변수이다. 1번 변수부터 마지막 변수까지 변수가 제거되며 기록된 예측모형의 예측률을 보여준다. 예를 들어, 표본 1에서는 8개의 변수를 사용했을 때 72.52%, 표본 7에서는 7개의 변수를 사용했을 때 70.33%의 예측률을 기록하였다. 표본 9와 표본 6에서는 5개의 변수를 사용했을 때 70.91%와 70.94%의 예측률을 기록하였다. 앞서 분석된 결과를 통해, 표본 1에서 근로자의 이직준비 여부의 예측률이 가장 높은 것을 확인할 수 있다.

4.2 근속기간별 데이터의 예측모형 결과

근속기간별 데이터로부터 근속기간별 이직준비 여부에 관한 실험 과정을 요약하면 <그림 2>와 같다. 먼저 '4~6개월', '7~12개월', '13~24개월', '25개월 이상'으로 기간을 분류한 뒤, 근속기간별로 5개의 표본을 언더샘플링 방식으로 생성하였다.

'실험 2'는 '실험 1'과 같이 의사결정나무 알고리즘을 사용하여 예측모형을 만들었고, 역방향 제거를 이용하여 각 표본당 33개 이하의 예측모형을 구축하였다. 근속 기간마다 예측모형을 만들어 '4~6개월'은 59개, '7~12개월'은 80개, '13~24개월'은 90개 그리고 '25개월 이상'은 73개로 최종 302개의 예측모형이 구축되었다. <표 4>는 예측모형이 근속기간별 표본에서의 최고예측률을 보여준다.

<표 4> 근속기간별 표본에서의 최고예측률

기간	표본	최고예측률 (%)
4~6개월	표본 1	67.56
	표본 2	65.06
	표본 3	67.57
	표본 4	70.64
	표본 5	67.76
7~12개월	표본 1	69.65
	표본 2	69.37
	표본 3	69.69
	표본 4	66.53
	표본 5	70.65
13~24개월	표본 1	68.71
	표본 2	69.33
	표본 3	69.83
	표본 4	70.21
	표본 5	71.16
25개월 이상	표본 1	69.21
	표본 2	69.29
	표본 3	69.78
	표본 4	66.09
	표본 5	70.32

<표 5>는 근속기간별로 최고의 예측률을 기록한 표본 결과를 보여준다. '4~6개월' 구간에서는 표본 4에서 8개의 변수를 사용하였을 때 70.64%의 예측률을 보여주었다. '7~12개월' 구간에서는 표본 5에서 5개의 변수를 사용하였을 때 70.65%, '13~24개월' 구간에서는 표본 5에서 8개의 변수를 사용하였을 때 71.16% 그리고 '25개월 이상' 구간에서 표본 5에서 5개의 변수를 사용하였을 때 70.32%의 예측률을 기록하였다.

<표 5> 근속기간별 최고예측률을 기록한 표본 결과

순서	4~6개월		7~12개월		13~24개월		25개월 이상	
	표본 4		표본 5		표본 5		표본 5	
	변수명	DT(%)	변수명	DT(%)	변수명	DT(%)	변수명	DT(%)
1	C3	67.54	D3	62.53	D3	63.52	C12	61.98
2	C13	63.23	D4	61.61	B6	66.53	C11	63.08
3	C12	64.08	C2	60.72	F1	66.83	C2	59.54

<표 5> 근속기간별 최고예측률을 기록한 표본 결과(계속)

순서	4~6개월		7~12개월		13~24개월		25개월 이상	
	표본 4		표본 5		표본 5		표본 5	
	변수명	DT(%)	변수명	DT(%)	변수명	DT(%)	변수명	DT(%)
4	C8	66.70	C4	63.56	D4	66.33	C9	61.24
5	D3	70.64	C11	61.23	C6	66.16	A3	61.24
6	D1	65.06	C9	59.39	C10	66.00	A3	62.50
7	C7	64.94	C6	60.27	C4	68.33	A2	60.56
8	C1	60.56	C3	66.59	C3	64.70	C4	60.66
9	C6	64.92	D2	63.84	C7	65.17	D2	60.89
10	D2	65.79	C1	68.30	C13	65.03	C5	64.56
11	C2	69.37	D1	67.03	C2	69.33	C7	67.86
12	C5	61.03	C13	65.18	C9	71.16	D1	70.32
13			C10	70.65	C8	67.17	C3	66.05
14			C7	67.92	C12	69.21	C1	63.56
15			C8	62.16	C11	67.67	C8	63.63
16			C5	65.18	C1	64.75	C13	59.88
17			C12	59.40	D2	64.53		
18					D1	63.87		
19					C5	57.68		

이런 결과들을 살펴보면 ‘4~6개월’ 구간을 제외한 나머지 구간들의 예측모형에서 4개월 이상의 전 구간 예측모형의 평균 예측률인 68,186%보다 높은 예측률을 보였다. 이를 통해 예측모형은 4개월 이상의 전 구간 데이터를 사용하는 것보다 분화된 데이터인 근속기간별 데이터를 사용하였을 때, 예측의 성능이 높아지는 것을 알 수 있다.

V. 고용유지 강화방안

5.1 근속 기간이 4개월 이상인 정규직 근로자에 관한 고용유지 방안 도출

본 절에서는 ‘실험 1’의 의사결정나무를 통해 나온 예측모형의 결과를 활용하여 수습 기간이 지난 정규직 근로자에 관한 이직 의도에 관한 결정규칙을 구하고, 정규직 근로자의 고용을 유지하는 강화방안을 제시하려고 한다. 의사결정나무 알고

리즘을 통해 예측모형을 도출할 때, 의사결정나무는 분류모형이 복잡하지 않게 구성되어 분류모형을 통한 결정규칙 해석이 쉽다는 장점이 있다 (Witten *et al.*, 2016).

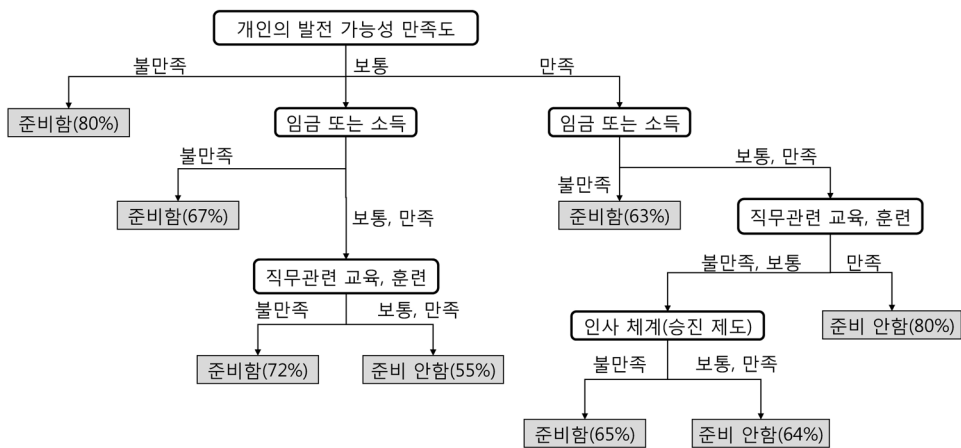
<표 3>을 통해 변수 8개를 사용한 표본 1에서 예측률이 75.52%로 모든 표본 중 가장 높은 것을 확인할 수 있다. 이때 변수는 C2(고용 안정성), C11(일자리에 대한 사회적 평판), C1(임금 또는 소득), C8(인사 체계), D2(일의 수준 대비 본인의 수준 일치 정도), C13(직무관련 훈련 또는 교육), C5(개인 발전 가능성), D3(교육 수준 일 수준 일치 정도)가 사용되었다. <그림 3>은 75.52%의 예측률을 보인 표본 1의 8개 변수 사용한 예측모형의 결정규칙을 도식화한 것이다. 이 예측모형은 결정규칙에 따라 5개의 ‘준비함’과 3개의 ‘준비 안 함’으로 나뉜다. <표 6>은 <그림 3>의 8개의 결정규칙을 ‘준비함’과 ‘준비 안 함’의 순서대로 높은 비율부터 보여준다. 이를 통해 8개의 결정규칙 비교

및 이직 의도에 영향을 주는 요인 분석을 통해 고용유지 방안을 제시하고자 한다.

첫 번째, 규칙 1을 통해 이직준비에 가장 많은 영향을 주는 요인이 개인 발전 가능성에 대한 만족도임을 알 수 있다. 개인 발전 가능성에 대해 만족하지 않은 응답자의 80%는 이직을 준비했다. 이는 전체의 23%의 수준이다. 이직을 준비하는 응답자의 46%가량이 개인 발전 가능성에 만족하

지 않았다. 따라서 기업은 근로자들에게 창의적인 업무부여, 학사 및 석·박사 학위 획득 지원 등의 방법을 통해 근로자들의 개인 발전 가능성 만족도를 높여야 한다.

두 번째, <표 6>의 규칙 2와 규칙 3을 비교했을 때 근로자들은 개인 발전 가능성 만족도가 보통 혹은 만족이더라도, 임금 만족도가 불만족이면 이직을 준비했다. 이에 대한 예측률은 각각 67%,



<그림 3> 의사결정나무로 도출된 이직준비 여부에 관한 규칙

<표 6> 근속기간이 4개월 이상인 근로자에 관한 이직준비 여부 규칙

규칙	내용	이직 예측	예측률	비율
1	개인 발전 가능성 = 불만족	준비함	80%	23%
2	개인 발전 가능성 = 보통 & 임금 또는 소득 = 불만족	준비함	67%	10%
3	개인 발전 가능성 = 만족 & 임금 또는 소득 = 불만족	준비함	63%	8%
4	개인 발전 가능성 = 보통 & 임금 또는 소득 = 보통, 만족 & 직무관련 교육, 훈련 만족도 = 불만족	준비함	72%	2%
5	개인 발전 가능성 = 만족 & 임금 또는 소득 = 보통, 만족 & 직무관련 교육, 훈련 = 보통, 불만족 & 인사 체계(승진 제도) = 불만족	준비함	65%	1%
6	개인 발전 가능성 = 만족 & 임금 또는 소득 = 보통, 만족 & 직무관련 교육, 훈련 = 만족	준비 안 함	80%	28%
7	개인 발전 가능성 = 보통 & 임금 또는 소득 = 보통, 만족 & 직무관련 교육, 훈련 만족도 = 보통, 만족	준비 안 함	55%	20%
8	개인 발전 가능성 = 만족 & 임금 또는 소득 = 보통, 만족 & 직무관련 교육, 훈련 = 보통, 불만족 & 인사 체계(승진 제도) = 보통, 만족	준비 안 함	64%	8%

63%다. 이 둘의 비율은 합치면 전체 중 18%를 차지한다. 따라서 기업은 근로자들의 임금 만족도를 높이기 위해 그들의 임금을 주관적인 만족 수준까지 높이거나 임금에 대한 타당성을 납득시켜야 한다.

세 번째, 규칙 4와 규칙 7을 비교했을 때 개인 발전 가능성 만족도가 보통이고 임금 만족도가 보통 또는 만족일 때 직무관련 교육, 훈련 만족도에 의해 이직 의도가 나뉘었다. 또한, 규칙 6에 따르면 개인 발전 가능성 만족도가 만족이고 임금 만족도가 보통 또는 만족일 때 직무관련 교육, 훈련 만족도가 만족이면 이직준비를 하지 않는 것으로 나타났다. 규칙 6과 규칙 7은 각각 표본 전체 응답자의 28%, 20%를 차지했다. 즉, 전체 48% 정도인 대부분의 이직준비를 하지 않는 응답자들은 이 범주에 속했다. 대부분의 이직준비를 하지 않는 근로자들은 개인 발전 가능성 만족도, 임금 만족도가 비슷한 조건일 경우 직무관련 교육, 훈련 만족도에 영향을 받는다. 따라서 기업은 근로자들은 이직준비를 하지 않는 범주에 더 많이 포함시키기 위해선 개인 특성에 맞는 직무관련 교육, 훈련 프로그램을 구성해야 한다.

네 번째, 규칙 5와 규칙 8은 개인 발전 가능성 만족도, 임금만족도, 직무관련 교육, 훈련 만족도가 같은 조건일 때 승진 제도 만족도에 따라 이직준비 여부가 달라짐을 보여준다. 개인 발전 가능성 만족도는 만족, 임금 만족도는 보통 혹은 만족, 직무관련 교육 훈련 만족도는 불만족 혹은 보통일 때 승진 제도에 대해 불만족한 응답자 중 65%는 이직을 준비했고, 같은 질문에 대해 보통 혹은 만족이라 응답한 인원 중 64%는 이직을 준비하지 않았다. 따라서 기업은 이직준비 의도를 낮추기 위해서 근로자들이 승진 제도에 대해 납득할 수 있게 해야 한다.

5.2 근속기간별 근로자의 고용유지 방안 도출

본 절에서는 근속기간별 데이터 분석을 통해 정규직 근로자의 근속기간별 차별화된 이직준비

의사결정 요인들을 분석하고, 고용유지 방안을 제안하려 한다. <표 5>에서 보이는 의사결정나무를 통해 구해진 예측모형의 기간별 최고예측률 결과를 활용하며, ‘준비함’ 규칙을 중심으로 기간별 근로자의 이직 의사결정에 대한 요인을 분석하여 고용유지 방안을 도출하고자 한다.

<표 7>은 근속기간별로 분류한 정규직 근무자들의 이직준비 여부 결정규칙을 보여준다. ‘실험 1’의 정규직 근로자의 결정규칙과 마찬가지로, 근속 기간이 4~6개월, 7~12개월, 13~24개월인 정규직 근로자들의 이직 의사결정에 주요한 요인으로 개인 발전 가능성 만족도와 임금 또는 소득이 나타났다. 또한, 모든 근속 기간에서 공통으로 임금 또는 소득 만족도가 이직 의사결정에 영향을 주는 것으로 나타났다. 하지만 기간별로 차별화된 이직준비 의사결정 요인 또한 나타났다.

근속 기간이 4~6개월의 정규직 근로자의 이직준비 여부 결정규칙 2를 보면, 개인 발전 가능성에 대해 만족하더라도 고용 안정성에 만족하지 못하면 이직준비를 하는 것을 알 수 있다. 또한, 결정규칙 4에서 개인 발전 가능성과 고용 안정성에 대해 만족하더라도 인간관계와 기업의 복리후생 제도에 만족하지 못할 때 이직을 생각하는 것을 알 수 있다. 상대적으로 근속 기간이 짧은 근로자들은 고용 안정성, 직장 내 인간관계, 복리후생 제도가 큰 영향을 주는 것으로 나타났다. 따라서 기업은 근속 기간이 짧은 근로자에 대해서 고용 안정성을 보장하며, 원활한 인간관계를 유지할 친목 도모를 위한 비 업무적인 활동을 제공해야 한다. 또한, 근로자들이 만족할 수 있는 복리후생 제도를 마련해야 한다.

근속 기간이 13~24개월인 근로자들의 이직준비 여부 규칙 4를 보면, 개인 발전 가능성과 임금 또는 소득에 만족하더라도, 업무와 자신의 적성, 흥미 일치 만족도와 업무에 대한 사회적 평판에 만족하지 못하면 이직준비를 하는 것을 알 수 있다. 근속 기간이 13~24개월의 근로자들은 기업 내부의 제도적인 부분보다 개인 적성과 일의 일치 정도와 업무에 대한 사회적 평판이 이직 의사결정

에 더 영향을 미치는 것으로 보인다. 기업은 근속 기간이 13~24개월인 근로자들에 대해서 근로자들의 적성과 흥미에 맞는 업무와 사회적인 평판이 좋은 업무를 지시해야 한다.

근속 기간이 25개월 이상인 근로자들의 이직준비 여부 규칙에서 이전 근로자들과 전반적으로 다른 경향을 보인다. 근속 기간이 25개월 이상인 근로자들의 결정규칙 1을 보면, 이들 근로자의 이직

〈표 7〉 근속기간별 근로자에 관한 이직준비 여부 규칙

기간	규칙	내용	이직 예측	예측률	비율
4~6개월	1	개인 발전 가능성 = 불만족	준비함	81%	26%
	2	개인 발전 가능성 = 보통, 만족 & 고용 안정성 = 불만족, 보통	준비함	62%	20%
	3	개인 발전 가능성 = 만족 & 임금 또는 소득 = 불만족	준비함	64%	5%
	4	개인 발전 가능성 = 보통, 만족 & 고용 안정성 = 만족 & 인간관계 = 보통, 불만족 & 복리후생 제도 = 보통, 불만족	준비함	78%	4%
7~12개월	1	개인 발전 가능성 = 불만족	준비함	87%	24%
	2	개인 발전 가능성 = 보통, 만족 & 임금 또는 소득 = 불만족	준비함	54%	12%
	3	개인 발전 가능성 = 만족 & 임금 또는 소득 = 불만족	준비함	77%	5%
13~24개월	1	개인 발전 가능성 만족도 = 불만족	준비함	79%	23%
	2	개인 발전 가능성 = 보통 & 임금 또는 소득 = 불만족	준비함	71%	9%
	3	개인 발전 가능성 = 만족 & 임금 또는 소득 = 불만족	준비함	60%	9%
	4	개인 발전 가능성 = 보통, 만족 & 임금 또는 소득 = 보통, 만족 & 하는 일과 자신의 적성, 흥미의 일치 = 보통, 불만족 & 하는 일에 대한 사회적 평판 = 보통, 불만족	준비함	62%	9%
25개월 이상	1	직무관련 훈련 또는 교육 = 불만족	준비함	92%	17%
	2	직무관련 훈련 또는 교육 = 보통, 만족 & 임금 또는 소득 = 불만족 & 인사체계(승진제도) = 불만족	준비함	88%	6%
	3	직무관련 훈련 또는 교육 = 보통 & 임금 또는 소득 = 불만족 & 인사체계(승진제도) = 보통, 만족	준비함	72%	6%
	4	직무관련 훈련 또는 교육 = 보통, 만족 & 임금 또는 소득 = 만족 & 교육수준 일 수준 일치 정도 = 불만족	준비함	83%	4%
	5	직무관련 훈련 또는 교육 = 보통, 만족 & 임금 또는 소득 = 보통 & 교육수준 일 수준 일치 정도 = 보통, 만족 & 인사체계(승진제도) = 불만족	준비함	70%	4%

의사결정에 있어서 가장 중요한 요인은 직무 관련 훈련 또는 교육으로 나타났다. 또한, 규칙 2와 규칙 5에서 이들 근로자는 인사 체계(승진 제도)에 만족하지 못한 근로자들은 이직을 준비했다. 규칙 4에서 근로자들은 직무 관련 훈련 또는 교육 그리고 임금 또는 소득에 만족하더라도, 교육수준과 일의 수준의 일치 정도에 불만족하다면 이직준비를 하는 것을 알 수 있다. 근속 기간이 25개월 이상인 장기적인 근로자들은 직무관련 훈련 또는 교육, 인사 체계(승진 제도), 교육수준과 일의 수준의 일치 정도의 만족도에 따라 이직을 생각하는 것을 알 수 있다. 따라서 기업은 근속 기간이 25개월 이상인 근로자들의 고용유지를 강화하기 위하여 직무 관련 교육을 제공해야 하며, 공정한 인사 체계를 마련하고 이들의 교육수준에 맞는 업무를 부여해야 한다.

VI. 결 론

본 연구에서 기업의 지속적인 고용유지를 위해, 근속 기간 4개월 이상의 정규직 근로자와 근속기간별로 나뉜 근로자들을 대상으로 각각 이직 의도에 영향을 주는 요인들을 분석하여, 근로자들의 조직 이탈 방지와 고용유지 강화를 위한 실용적인 방안을 제안하고자 했다. 의사결정나무 알고리즘을 사용하여 실용적 방안 제시를 위한 결정규칙을 도출하고, 근속 기간이 4개월 이상인 근로자들과 근속기간별로 구분된 근로자들의 이직준비 여부 예측모형을 구축하였다. 이렇게 구축된 예측모형을 통해 이직준비 여부에 관한 결정규칙을 도출하고, 이를 활용하여 고용유지 강화방안을 제시하였다.

주요 내용을 요약하면 다음과 같다. 근속 기간 4개월 이상의 정규직 근로자들의 경우 ‘교육수준과 일의 수준의 일치 정도’, ‘개인 발전 가능성’, ‘직무관련 교육 및 훈련’, ‘일의 수준 대비 본인의 수준 일치 정도’, ‘인사 체계(승진 제도)’, ‘임금 및 소득’, ‘일자리에 대한 사회적 평판’, ‘고용 안정성’ 요인이 이직 의도에 영향을 미치는 것으로 분

석되었다. 주요요인을 통하여 도출된 8가지의 이직 의도에 관한 규칙을 활용하여, 개인 발전 가능성 증진 기회 확대, 근로자의 주관적 임금 기준과 임금 수준 일치화, 직무 관련 교육, 훈련 프로그램 제공, 공정한 인사 체계 마련 등 근로자의 고용유지 강화방안을 제시하였다.

근속기간별로 구분한 연구에서, 전반적으로 근속 기간이 4개월 이상인 근로자의 결과와 비슷하게 ‘개인 발전 가능성’, ‘임금 및 소득’ 요인에서 이직준비에 많은 영향을 끼친 것으로 확인되었다. 하지만 근속기간별로 특징적인 요인들도 나타났다. 근속 기간이 비교적 짧은 4~6개월 근로자들이 이직준비에 영향을 주는 요인으로는 ‘고용 안정성’, ‘직장 내 인간관계’, ‘복리 후생제도’가 나타났다. 근속 기간이 13~24개월인 근로자들은 기업 내부의 제도적인 요인보다는 ‘업무와 자신의 적성, 흥미 일치 만족도’, ‘업무에 대한 사회적 평판 만족도’와 같은 요인이 이직 의사결정에 더 영향을 주는 것으로 나타났다. 그리고 근속 기간이 25개월 이상인 장기적인 근로자들은 이전 기간으로 구분된 근로자들과는 다르게 이직준비에 가장 큰 영향을 주는 요인으로 ‘직무관련 훈련 또는 교육’, ‘인사 체계(승진 제도)’, ‘교육수준 일 수준 일치 정도’가 분석되었다.

본 연구는 기존 연구들과는 다른 두 가지의 차별점을 두고 있다. 첫째, 기존 문헌들은 근속기관과 이직 의도에 관한 연구를 할 때, 대부분의 연구에서 이직 의도를 종속변수로 근속 기간은 독립변수로 설정했다. 하지만 본 연구를 데이터를 근속 기간으로 범주화하여 기간별 근로자들의 서로 다른 이직 의도에 영향을 주는 요인을 발견했다. 둘째, 기존 연구들은 분석 방법론은 주로 구조 방정식이나 다중회귀분석을 이용했다. 이러한 방법론은 독립변수를 연구자가 제한해야 하고, 다양한 독립변수를 한 번에 비교할 수 없다. 하지만 본 연구에서는 데이터 마이닝 기법을 활용하여 이직 의도에 영향을 줄 것으로 예상되는 많은 독립변수를 한 번에 비교했다. 이를 통해 분석된

결과를 통해 실용적인 근로자의 고용유지 방안을 제시했다.

본 연구는 3가지의 한계점을 가진다. 첫 번째, 본 연구의 분석 데이터는 언더샘플링 과정을 통해 구성된다. 전체 데이터에서 이직준비 의사가 없는 응답자의 수에 맞춰 이직준비 의사가 있는 응답에서 샘플링을 했다. 전체 응답자 중 이직 의사가 없는 응답자의 수가 현저히 적기 때문에 언더샘플링 후의 표본 수가 적어 연구 결과가 일반화될 수 있다는 점에서 한계를 가진다. 두 번째, 근속기간 별 데이터에서 근속 기간이 24개월 이상의 데이터는 하나로 범주화했다. 연구 자료로 사용한 데이터의 구조상 근속 기간이 24개월 이상인 데이터는 많지 않아서 한 기업에 장기적으로 근속 중인 근로자에 대한 세밀한 분석이 이루어지지 못했다. 세 번째, 의사결정 나무를 구축할 때 복잡성 매개 변수를 통제하여 자식 노드의 최소 비율을 정함으로써 데이터의 세분화된 분석이 이루어지지 않았다. 명료한 분석을 위해 자식 노드의 최소 비율을 정함에 따라 분석되지 못한 독립변수들이 존재했다.

참 고 문 헌

- [1] 김경범, 이주현, 노진원, 권영대, “청년 취업자의 이직 준비 관련 요인: 청년패널 2007-2013 자료 분석”, *한국콘텐츠학회논문지*, 제16권, 제12호, 2016, pp. 480-491.
- [2] 김민경, 나인강, “교육훈련이 기업성장에 미치는 영향: 교육훈련전이의 매개효과를 중심으로”, *대한경영학회지*, 제25권, 제4호, 2012, pp. 2047-2064.
- [3] 김보인, “근로자의 이직의도에 영향을 미치는 요인에 대한 연구: 근로자 구성특성 별 비교”, *인적자원관리연구*, 제24권, 제5호, 2017, pp. 31-47.
- [4] 김진희, “기업의 고용유지정책이 종업원의 고용안정성 인식을 통해 정서적 몰입과 유지적 몰입에 미치는 효과”, *HRD 연구(구 인력개발연구)*, 제11권, 제3호, 2009a, pp. 232-251.
- [5] 김진희, “고몰입 인적자원관리가 종업원의 인사제도 내재화, 조직몰입 및 품질경영활동에 미치는 효과: 제조업종을 대상으로”, *노동정책연구*, 제9권, 제3호, 2009b, pp. 81-111.
- [6] 박인규, “호텔 종사원의 직무배태성이 이직 의도에 미치는 영향에 관한 연구: 임금만족을 조절변수로”, *관광연구*, 제25권, 제3호, 2010, pp. 61-79.
- [7] 성지미, 안주엽, “일자리 만족도와 이직의사 및 이직-청년층을 중심으로”, *산업노동연구*, 제22권, 제2호, 2016, pp. 135-179.
- [8] 윤은자, 권영미, 전미순, 안정화, “간호사의 이직의도 영향요인”, *한국콘텐츠학회논문지*, 제16권, 제1호, 2016, pp. 94-106.
- [9] 이정현, 김동배, “고성과조직시스템의 효과와 영향요인”, *산업관계연구*, 제17권, 제1호, 2007, pp. 1-38.
- [10] 이진복, 박원준, “상사의 이직의도가 부하의 이직의도에 미치는 영향: 이직의도, LMX, POS의 상호작용 효과를 중심으로”, *기업경영연구(구 독립경영연구)*, 제73권, 2017, pp. 117-136.
- [11] 이철호, 정정숙, 이민규, “부모-청소년 의사소통이 청소년의 우울에 미치는 영향”, *한국심리학회지: 건강*, 제14권, 제2호, 2009, pp. 433-448.
- [12] 이채호, “보육교사의 이직의도 관련 연구동향 분석”, *학습자중심교과교육연구*, 제19권, 제3호, 2019, pp. 153-168.
- [13] 정석원, *SPSS seminar*, 2014.
- [14] 통계청, “국가통계포털”, 2019, Available at <http://kosis.kr/index/index.do>.
- [15] 황정임, 김은자, 이상영, “농업, 농촌의 공익기능에 대한 사회적 인식 연구”, *농촌지도와 개발*, 제16권, 2009, pp. 967-992.
- [16] Adriaans, P. and D. Zantinge, *Data Mining Addison Wesley Longman Limited*, Edinburgh Gate, Harlow, CM20 2JE, England, 1996.
- [17] Auer, P. and S. Cazes, “Resilience of the long-term

- employment relationship: Evidence from the industrialized countries”, *The International Labour Review*, Vol.139, No.4, 2000, pp. 379-408.
- [18] Brethenoux, E., H. Dresner, K. Strange, and J. Block, “Data warehouse, data mining and business intelligence: The hype stops here”, *Gartner Group Strategic Analysis Report*, 1996, pp. 300-105.
- [19] Burgess, S., L. Pacelli, and H. Rees, “Job tenure and labour market regulation: A comparison of Britain and Italy using micro data”, *CEPR Discussion Papers*, No.1712, 1997,
- [20] Burgess, S. and H. Rees, “A disaggregate analysis of the evolution of job tenure in Britain, 1975-1993”, *British Journal of Industrial Relations*, Vol.36, No.4, 1998, pp. 629-655.
- [21] Cazes, S., “Employment stability in an age of flexibility: Evidence from industrialized countries”, *International Labour Organization*, 2003.
- [22] Chawla, N. V., “Data mining for imbalanced datasets: An overview”, *Data Mining and Knowledge Discovery Handbook*, Springer, Boston, MA, 2009, pp. 875-886.
- [23] Dash, M., and H. Liu, “Feature selection for classification”, *Intelligent Data Analysis*, Vol.1, No.1-4, 1997, pp. 131-156.
- [24] Heisz, A., “The evolution of job stability in Canada: Trends and comparisons with US results”, *Canadian Journal of Economics/Revue Canadienne D’économique*, Vol.38, No.1, 2005, pp. 105-127.
- [25] Forrest, C. R., L. L. Cummings, and A. C. Johnson, “Organizational participation: A critique and model”, *Academy of Management Review*, Vol.2, No.4, 1977, pp. 586-601.
- [26] Hackman, J. R. and G. R. Oldham, “Development of the job diagnostic survey”, *Journal of Applied Psychology*, Vol.60, No.2, 1975, p. 159.
- [27] Hancock, J. I., D. G. Allen, F. A. Bosco, K. R. McDaniel, and C. A. Pierce, “Meta-analytic review of employee turnover as a predictor of firm performance”, *Journal of Management*, Vol.39, No.3, 2013, pp. 573-603.
- [28] Kwon, K. and D. E. Rupp, “High-performer turnover and firm performance: The moderating role of human capital investment and firm reputation”, *Journal of Organizational Behavior*, Vo.34, No.1, 2013, pp. 129-150.
- [29] Lee, R. G. and P. J. Eckert, “Establishment size and employment stability in logging and sawmilling: A comparative analysis”, *Canadian journal of forest research*, Vol.32, No.1, 2002, pp. 67-80.
- [30] Mobley, W. H., *Employee Turnover: Causes, Consequences, and Control*, Addison-Wesley, 1982.
- [31] OECD, “Is Job Insecurity on the Increase in OECD Countries?”, *Chapter 5, OECD Employment Outlook 1997*, OECD Publishing, Paris, 1997.
- [32] Steers, R. M., R. T. Mowday, and L. W. Porter, *Employee Turnover and Post Decision Accommodation Processes*, No. TR-22. Oregon Univ Eugene Graduate School of Management and Business, 1979.
- [33] Witten, I. H., E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann, 2016.

Empirical Analysis of Influential Factors Affecting Domestic Workers' Turnover Intention: Emphasis on Public Database and Decision Tree Method

Geo Nu Ko* · Hyun Jin Jo** · Kun Chang Lee***

Abstract

This study addresses the issue of which factors make domestic workers have turnover intention. To pursue this research issue, we utilized a public database “2017 Occupational Migration Path Survey”, administered by Korea Employment Information Service (KEIS). Decision tree method was applied to extract crucial factors influencing workers' turnover intention. They include ‘the degree of matching the level of education with the level of work’, ‘the possibility of individual development’, ‘the job-related education and training’, ‘the promotion system’, ‘wage and income’, ‘social reputation for work’ and ‘the stability of employment’.

Keywords: *Data Mining, Decision Tree, Turn over, Employment Maintenance, Service Period*

* College Student, Department of International Business & Trade, Kyung Hee University

** College Student, Department of Mathematics, Kyung Hee University

*** Corresponding Author, Professor, Global Business Administration / Department of Health Sciences & Technology, SAIHST(Samsung Advanced Institute for Health Sciences & Technology), Sungkyunkwan University

◎ 저 자 소 개 ◎



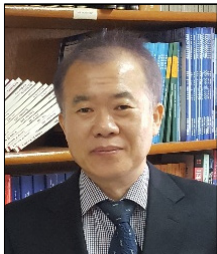
고 건 우 (rhrjsdn1214@gmail.com)

경희대학교 무역학과에 재학 중이며, 관심 분야는 데이터 마이닝, 텍스트 마이닝이다.



조 현 진 (rkwhrdb95@gmail.com)

경희대학교 수학과를 졸업했으며, 관심 분야는 인공지능, 빅데이터 분석, 텍스트 마이닝이다.



이 건 창 (kunchanglee@gmail.com)

한국과학기술원 경영과학과에서 석사 및 박사학위를 취득하였고, 현재 성균관대학교 경영대학 글로벌 경영학과/삼성융합의과학원(SAIHST) 융합의과학과 교수로 재직하고 있다. Decision Support Systems, Computers in Human Behavior, Journal of MIS, IEEE Transactions on Engineering Management, Frontiers in Psychology 등 다수의 국외 저널에 논문을 게재하였으며, 주요 연구분야는 인공지능, 빅데이터 분석, 헬스인포매틱스, 감성분석 등이다.

논문접수일 : 2020년 01월 15일

게재확정일 : 2020년 06월 08일

1차 수정일 : 2020년 04월 15일