

데이터 마이닝 기법을 활용한 근로자의 고용유지 강화 방안 개발

Enhancing Workers' Job Tenure Using Directions Derived from Data Mining Techniques

안민욱, 김태운, 유동희
경상대학교 경영정보학과

Minuk An(naru_mine@naver.com), Taeun Kim(dmldz@naver.com),
Donghee Yoo(dhyoo@gnu.ac.kr)

요약

본 연구에서는 데이터 마이닝 기법을 활용하여 근로자의 이직준비 여부에 관한 예측모형을 구축하는 실험을 진행하였다. 이를 위해, 한국고용정보원 주관으로 수집된 「2015년 대졸자 직업 이동경로조사」 데이터를 사용하였다. 이직준비 여부 예측모형에는 의사결정나무, 베이즈넷, 인공신경망 알고리즘이 사용되었다. 전체 직종을 대상으로 한 분석에서는 의사결정나무 기반 예측모형에서 최고 예측률을 기록하였으며, 이직준비 여부에 영향을 주는 요인은 '근로시간 형태', '종사상 지위', '정규직 여부', '주당 정규 근로시간', '주당 정규 근로일', '개인의 발전가능성'으로 나타났다. 의사결정나무 기반 예측모형의 결과를 활용하여 근로자 전반에 관한 12개의 이직준비 여부 규칙을 최종 도출하였고, 도출된 규칙을 바탕으로 근로자의 고용유지 강화에 도움을 주는 방안들을 제안하였다. 또한 직종별 영향 요인을 분석하기 위해 직종을 사무, 문화예술, 건설, 정보기술 분야로 구분하여 실험을 진행하였다. 그 결과 사무 분야는 10개, 문화예술 분야는 9개, 건설 분야는 4개, 그리고 정보기술 분야는 6개의 이직준비 규칙이 도출되었고 이를 토대로 직종별 맞춤형 고용유지 강화 방안을 제시하였다.

■ 중심어 : | 고용유지 | 이직 | 데이터마이닝 | 예측모형 | 의사결정나무 | 베이즈넷 | 인공신경망 |

Abstract

This study conducted an experiment using data mining techniques to develop prediction models of worker job turnover. The experiment used data from the "2015 Graduate Occupational Mobility Survey" by the Korea Employment Information Service. We developed the prediction models using a decision tree, Bayes net, and artificial neural network. We found that the decision tree-based prediction model reported the best accuracy. We also found that the six influential factors affecting employees' turnover intention are *type of working time*, *job status*, *full-time or not full-time*, *regular working hours per week*, *regular working days per week*, and *personal development opportunities*. From the decision tree-based prediction model, we derived 12 rules of employee turnover for all job types. Using the derived rules, we proposed helpful directions for enhancing workers' job tenure. In addition, we analyzed the influential factors affecting employees' job turnover intention according to four job types and derived rules for each: office (ten rules), culture and art (nine rules), construction (four rules), and information technology (six rules). Using the derived rules, we proposed customized directions for improving the job tenure for each group.

■ keyword : | Job Tenure | Job Turnover | Prediction Model | Decision Tree | Bayes Net | Artificial Neural Network |

I. 서론

급변하는 환경과 치열한 경쟁 속에서 기업이 지속적인 이윤을 창출하고 경쟁우위를 확보하여 생존하기 위해서는 기업 가치 생산의 주체인 인력에 관한 체계적인 유지관리가 필요하다. 2018년 1월 기준 국가통계포털의 자료에 의하면 전체 근로자의 이직률은 전년 대비 17.2% 증가하였으며, 그 중 자발적 이직률에 대한 비율은 29.2% 증가한 것으로 분석되었다. 이는 전체 이직자의 수가 144,413명 증가한 것이며, 자발적 이직자의 수도 74,726명 증가한 수치이다[1]. 여기에서 이직이란 조직의 구성원이 조직을 떠나는 것과 조직으로부터 노동의 대가로 보상을 받는 자격이 종결되는 것을 의미한다[2]. 이처럼 근로자의 이직이 발생할 경우 기업에서는 추가로 우수 인력을 확보하는 노력이 필요하며, 동시에 신규 인력을 선별하고 교육시키는 일에 막대한 비용이 발생하게 된다. 여기에서 근로자의 이직 의도나 이직준비 여부가 실제 이직으로 이어질 수 있는 가능성이 매우 높으며[3], 기업의 이직률이 높아질 경우 이직한 인력에 투자했던 비용을 회수하는 것이 어려워져 기업 경영에 부정적인 영향을 미치게 된다[4]. 따라서 기업은 이직 여부에 영향을 주는 요인들을 분석하여 근로자의 고용을 지속적으로 유지할 수 있는 실용적인 방안들을 마련해야 한다.

이직 의도에 관한 연구는 여러 학문 분야에서 오랫동안 연구되어온 주제이다. 사회심리학의 기대이론에서는 조직 구성원들이 가지고 있는 조직에 대한 기대가 충족되지 않을 경우 이직이 발생한다고 언급하였다[5]. 경제학에서는 현 직장보다 이직할 직장에서 얻을 수 있는 경제적 이익이 클 경우 이직이 발생한다고 설명하였다[6]. 또한 직무특성이론에서는 조직원들은 직무정체성, 기술다양성, 직무중요성을 통해 직무에 관한 의미를 부여 받게 되고, 직무자율성과 피드백을 통해 책임감과 업무를 지각하게 되는데 이러한 요인들이 이직 의도에 영향을 준다고 하였다[7].

여기에서 이직에 영향을 주는 여러 요인들은 산업별 특성이나 종사자들의 특징에 따라 달라질 수 있다. 예를 들어, 금융기관 종사자[8], 외식업체 종사원[9], 간호

사[10], 정보기술 분야 종사자[11], 치과위생사[12] 등과 같이 특정 직종에 속한 종사자들의 이직 요인들을 조사한 연구들의 결과를 살펴보면 각 연구에서 언급된 영향 요인들이 서로 상이한 것을 확인할 수 있다.

따라서 근로자의 이직 의도에 영향을 주는 요인들을 파악할 때 분석 대상을 전체 직종과 특정 직종으로 구분하여 분석을 진행할 경우, 근로자의 조직이탈 방지와 고용유지 강화에 도움을 주는 방안들이 좀 더 다양한 관점에서 개발될 수 있다. 즉, 전체 직종에 대한 분석을 통하여 노동시장의 전반적인 특성이 반영된 이직 요인들과 산업 전반에 걸쳐 이직에 영향을 주는 요인들의 우선순위를 파악할 수 있다. 또한 특정 직종에 대한 분석을 통해서 직종 별 근로자의 특성이 반영된 영향 요인들을 알 수 있으며, 각 직종에 맞춤형 고용유지 강화 방안을 개발할 수 있다.

본 연구에서는 데이터 마이닝 기법을 활용하여 근로자의 이직에 영향을 미치는 요인들을 분석하고 근로자의 고용유지 강화에 도움을 주는 실용적인 방안들을 전체 직종의 관점과 특정 직종의 관점으로 구분하여 제안하고자 한다. 실험을 위해 한국고용정보원 주관으로 수집된 「2015년 대졸자 직업 이동경로조사」 데이터를 사용하고자 한다. 전체 직종에 속한 근로자들의 이직준비 여부를 예측하는 모형은 의사결정나무(decision tree: DT), 베이즈넷(Bayes net: BN), 인공신경망(artificial neural network: ANN) 알고리즘을 활용하여 구축하고자 한다. 여기에서 가장 좋은 성능을 보인 예측모형을 통해 전체 직종에 관한 이직준비 여부 규칙을 도출하고자 하며, 도출된 규칙을 활용하여 근로자의 고용유지 강화에 도움을 주는 일반적인 방안들을 제안하고자 한다. 또한 특정 직종에 속한 근로자들의 이직준비에 영향을 주는 요인들을 분석하기 위해 직종을 사무, 문화예술, 건설, 정보기술 분야로 구분하여 실험을 진행하고자 하며, 각 분야에 맞춤형 고용유지 강화 방안을 제시하고자 한다.

본 연구의 구성은 다음과 같다. 2장에서는 근로자의 이직준비에 영향을 주는 요인들을 분석한 기존 문헌들에 대해 살펴본다. 3장에서는 본 연구에서 제안한 연구 방법과 데이터 마이닝 기법을 활용하여 이직준비 여부

를 판단하는 예측모형이 구축되는 실험 과정에 대해 설명한다. 4장에서는 연구 결과에 대해 설명하고, 5장에서는 근로자의 고용유지에 도움을 주는 방안들을 제시하고자 한다. 마지막으로 6장에서는 본 연구의 요약과 함께 시사점과 한계점에 대해 다룬다.

II. 문헌 연구

먼저 근로자들을 직종별로 구분하지 않고 이직에 영향을 주는 요인을 분석한 기존 연구들을 요약하면 다음과 같다. 김보인(2017)은 근로자들의 이직 의도에 조직분배공정성, 직무모호성, 상사의 배려행동, 직무갈등 순으로 영향을 준다고 하였으며, 조직공식성과 상사의 구조주도행동은 크게 영향을 주지 않는다고 언급하였다[13]. 이진복과 박원준(2017)은 상사의 이직 의도가 같은 조직의 직원 이직에 어떠한 영향을 주는가를 분석하였는데, 그 결과 상사의 이직 의도가 직원의 이직 의도에 영향을 미친다는 사실을 발견하였다[14]. 성지미와 안주엽(2016)은 청년층을 대상으로 일자리 만족도와 이직 의사의 관계를 조사하였는데, 일자리나 일의 만족도가 높을수록 이직 의사나 실제 이직 비중이 현저하게 낮아지는 것을 확인하였다[15]. 여기에서 일자리 만족도는 모든 수준에서 이직 의사에 통계적으로 유의미한 효과를 미치지만, 일의 수준이나 요구하는 기술의 수준, 전공과의 일치도는 부정적이라고 응답한 경우에만 이직 의도에 통계적으로 유의미한 효과를 미친다고 하였다. 김경범 등(2016)의 연구에서도 패널 자료를 활용하여 청년층의 이직준비와 관련된 요인들을 인과사회적 변수와 직무관련 변수로 나누어 실증 분석을 실시하였다[16]. 그 결과 연령, 성별, 최종학력, 월소득, 고용형태와 같은 인과사회적 변수와 자신의 교육수준 대비 업무의 난이도, 업무와 전공의 일치도, 자기발전 추구, 고용안정성, 일의 자율성과 권한, 직장 내 인간관계와 같은 직무관련 변수가 이직준비에 영향을 미치는 변수로 조사되었다. 김진희(2012)는 인사제도 내재화와 직무몰입은 이직 의도를 낮추고, 역할피로는 이직 의도를 높인다고 언급하였다[17].

다음으로 특정 직종에 한정하여 이직에 영향을 주는 요인을 분석한 연구는 다음과 같다. 이재명과 강신기(2015)는 금융기관 종사자들을 대상으로 이직 및 창업 의도에 영향을 미치는 요인을 연구하였는데, 승진만족, 업무만족, 보상만족, 대인관계만족 요인이 높을수록 이직 의도가 낮아지는 것이 조사되었고, 그 중 승진만족 요인이 가장 큰 영향력을 발휘하였다[8]. 이상희와 김혁수(2014)는 외식업체 종사자들의 직무만족과 이직 의도와와의 관계를 회귀분석을 통해 조사하였다[9]. 그 결과 업무만족과 보수만족이 높을수록 이직 의도가 낮아지는 것을 확인하였고, 보수만족 요인이 이직 의도와와의 관계에서 조절 효과를 주는 것을 발견하였다. 그리고 윤은자 등(2016)은 간호사들의 이직 의도에 영향을 주는 요인들을 분석하였는데, 인과사회학적 관점에서 일반 간호사가 미혼이고 3교대 근무자이며 경력이 낮을수록 이직 의도가 높다는 것을 확인하였고, 이직 의도에 간호사의 나이는 크게 중요하지 않다는 것을 알 수 있었다[10]. 김민주(2005)는 여행업 종사자들을 대상으로 이직 의도에 영향을 주는 요인이 집단간 차이가 보이는지를 분석하고자 하였다[18]. 그 결과, 여성보다는 남성이 그리고 관광통역안내를 하는 집단보다는 관리직 집단이 이직 의도가 높다는 것이 조사되었다. 이선로와 김기영(2006)은 정보기술직 종사자들을 대상으로 이직 의도의 결정요인을 분석하였다[19]. 그 결과 직무만족, 조직몰입, 추가노력은 이직 의도에 직접적으로 영향을 주며, 종사자의 권한부여 수준은 이직 의도에는 직접 영향을 주지는 않지만 매개변수 역할을 하는 직무만족과 조직몰입에 영향을 주는 것으로 조사되었다. 이행렬과 신용석(2012)은 사회복지전담 공무원을 대상으로 연구를 진행하였는데 담당 클라이언트의 수가 적절하지 않을수록 이직 의도가 높아졌고 전문적인 직무 수행으로 발생하는 스트레스가 심할수록 이직 의도가 낮아지는 것을 알 수 있었다[20].

기존 연구와 본 연구의 차이점을 비교해 보면 다음과 같다. 첫째, 분석 대상 관점에서 살펴보면, 기존 연구에서는 전체 직종을 대상으로 하거나 한 직종에 대한 영향 요인을 분석한 연구가 진행되었다면, 본 연구에서는 수집된 데이터를 활용하여 전체 직종에 대한 분석과 특

정 직종에 관한 분석을 유사한 실험 환경에서 함께 실시함으로써 산업 전반에 대한 요인을 분석함과 동시에 특정 직종에 차별화된 요인 분석을 진행하였다는 점에서 차이가 있다. 둘째, 기존 연구에서 사용된 변수들을 실험에서 활용한 부분은 비슷하지만, 기존 연구들에 비해 다양한 변수(사용변수 39개)들을 사용하여 이직 의도에 영향을 주는 요인들을 파악하였다. 셋째, 분석 방법 관점에서 살펴보면, 기존 연구에서는 주로 다중회귀 분석이나 구조 방정식과 같은 통계적 분석 기법을 활용하여 이직준비 여부에 영향을 주는 요인들의 영향력을 분석하였다면, 본 연구에서는 데이터 마이닝 기법을 활용하여 이직준비 여부에 영향을 주는 요인들의 영향력을 분석하고 동시에 이직 여부에 관한 규칙을 통하여 근로자의 고용유지를 강화할 수 있는 실용적인 방안을 제안한 점에서 차별성을 가진다.

III. 연구 방법

1. 연구 진행 과정

데이터 마이닝 기법을 활용하여 이직준비 여부에 관한 규칙을 도출하는 전반적인 과정을 요약하면 [그림 1]과

같다. 먼저 「2015년 대졸자 직업 이동경로조사」 데이터를 수집하였다. 이후 데이터 전처리(preprocessing) 과정을 통해 수집된 데이터를 데이터 마이닝에서 분석 가능한 형태로 변환한 뒤 분석에 이용될 초기 독립변수들을 선별하였다. [그림 1]의 '실험 1'에서 보는 바와 같이, 전 직종 근로자를 대상으로 하는 통합 데이터로부터 목표변수의 데이터 균형화(data balancing)를 위해 10개의 표본을 추출하였다. 이때, 통합 데이터에는 총 24개 직종에 관한 정보가 포함되어 있다. 변수 선택(feature selection) 과정을 통해서 목표변수에 영향을 주는 독립변수들을 최종 선택하였고, 의사결정나무, 베이스넷, 인공신경망 알고리즘을 활용하여 표본별로 예측모형을 구축하였다. 그 후, 구축된 예측모형들 중에서 가장 높은 예측률을 보인 예측모형을 선택하고 해당 예측모형을 통해 근로자 전반에 관한 이직준비 여부 규칙을 도출하였다. 다음으로, [그림 1]의 '실험 2'와 같이 직종별 특성을 파악하기 위해 대표 직종으로 사무, 문화예술, 건설, 정보기술 분야를 선별하였고, 각 직종별 근로자의 데이터를 추출한 뒤, '실험 1'과 유사한 방법으로 직종별 근로자의 이직준비 여부 규칙을 도출하였다. 이때 '실험 2'의 예측모형 구축에서는 '실험 1'에서 가장 좋은 성능을 보인 알고리즘을 활용하였다.

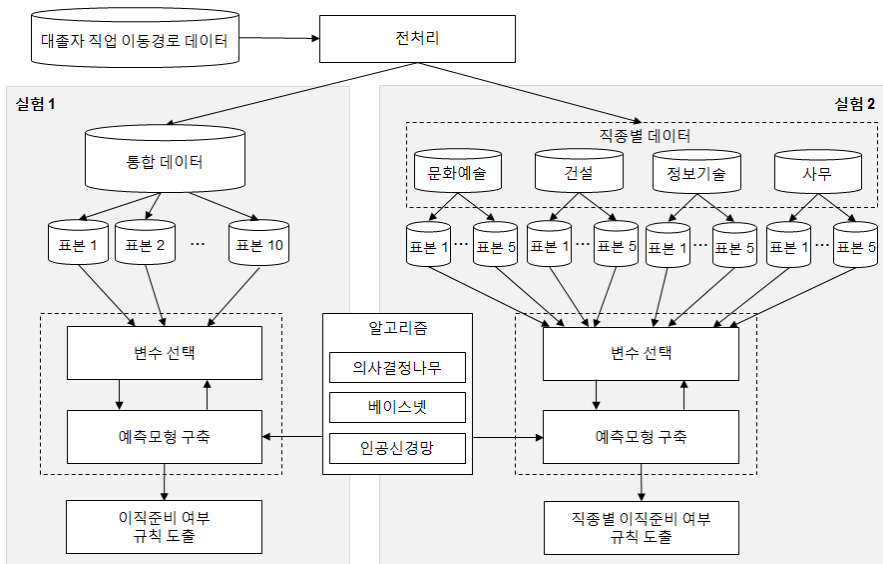


그림 1. 연구 진행 과정

2. 분석 데이터

본 연구에서는 한국고용정보원 주관으로 조사된 「2015년 대졸자 직업 이동경로조사」 데이터 18,056건을 분석 데이터로 수집하였다. 수집된 통합 데이터는 2015년 9월 1일을 기준으로 설문 대상자의 현재 경제활동 상황, 현재 일자리, 이직준비 여부 등에 관한 정보를 포함하고 있다. 그 후, 전처리 과정을 통해 10,883건의 데이터를 선별하였다. 선별된 데이터에서 1,835명이 이직을 준비하였고 나머지 9,048명은 이직을 준비하지 않았다. 전체 직종에서의 이직준비 비율은 16.86%임을 알 수 있었다. 또한 직종별 데이터의 경우, 사무(2800건), 문화예술(752건), 건설(323건), 정보기술(576건) 중 이직을 준비하는 사람은 사무(556명), 문화예술(129명), 건설(58명), 정보기술(72명)이며, 이직을 준비하지 않는 사람은 사무(2244명), 문화예술(623명), 건설(265명), 정보기술(504명)이다. 여기에서 각 직종별 이직준비 비율은 사무는 19.86%, 문화예술은 17.15%, 건설은 17.96%, 그리고 정보기술은 12.50%로 분석되었다.

3. 목표변수 및 독립변수

설문조사 대상자의 이직준비 여부에 대한 예측모형을 구축하기 위해 이직준비 여부를 나타내는 ‘Target’ 변수를 목표변수로 선택하였다. 목표변수의 값은 이직을 준비하고 있음을 의미하는 ‘준비함’과 이직을 준비하지 않고 있음을 의미하는 ‘준비안함’으로 구분된다.

독립변수 선정에 앞서 문헌연구에서 조사된 인공통계학적 변수, 직무관련 변수, 경제관련 변수, 고용유지관련 변수 등이 고려되었고, 그 중 이직준비 여부에 영향을 줄 것으로 판단된 39개의 독립변수들을 선정하였다. 그리고 이들 변수가 지닌 데이터 값들 중에서 의미가 없거나 널 값이 존재해 사용하기에 부적절한 7,173개의 데이터는 삭제 처리하였다. [표 1]은 전처리 과정을 통해 선택된 독립변수 39개와 목표변수 1개에 관한 정보를 나타내고 있다.

표 1. 실험에서 사용한 변수 정보

변수	설명	변수값
F1	성별	남, 여
F2	연령	숫자값
F3	혼인 상태	미혼, 기혼, 이혼, 사별
F4	부양 자녀 유무	예, 아니오
F5	가족의 경제적 지원 여부	예, 아니오
F6	가족에게 경제적 지원 제공 여부	예, 아니오
F7	이용 교통수단	걸어서, 자가용, 대중교통, 통근버스, 고속시외버스, 기차, 자전거, 기타
F8	통근 시간(분)	숫자값
F9	종사상 지위	상용근로자, 임시근로자, 일용근로자
F10	계약기간 지정 여부	정함, 정하지 않음
F11	근로시간 형태	전일제, 시간제
F12	정규직 여부	예, 아니오
F13	주당 정규 근로일	숫자값
F14	주당 정규 근로시간	숫자값
F15	월 평균 휴일 근로	숫자값
F16	월 평균 근로소득	숫자값
F17	임금 및 소득	불만족, 보통, 만족
F18	고용 안정성	불만족, 보통, 만족
F19	근무환경	불만족, 보통, 만족
F20	근로시간 및 근무일수	불만족, 보통, 만족
F21	개인의 발전가능성	불만족, 보통, 만족
F22	인간관계	불만족, 보통, 만족
F23	복리후생제도	불만족, 보통, 만족
F24	인사체계 및 승진제도	불만족, 보통, 만족
F25	일에 대한 사회적 평판	불만족, 보통, 만족
F26	일의 자율성과 권한	불만족, 보통, 만족
F27	일자리에 대한 사회적 평판	불만족, 보통, 만족
F28	일과 자신의 적성 일치도	불만족, 보통, 만족
F29	직무관련 교육 또는 훈련	불만족, 보통, 만족
F30	자신의 교육수준 대비 일의 수준	낮음, 알맞음, 높음
F31	자신의 수준 대비 일의 기술수준	낮음, 알맞음, 높음
F32	일과 전공 일치도	맞지 않음, 보통, 맞음
F33	전공지식 도움정도	도움안됨, 보통, 도움됨
F34	외국어 능력 요구정도	요구되지 않음, 보통, 요구됨
F35	국민연금 가입여부	가입됨, 가입안됨
F36	건강보험 가입여부	가입됨, 가입안됨
F37	고용보험 가입여부	가입됨, 가입안됨
F38	산재보험 가입여부	가입됨, 가입안됨
F39	취직 위해 구직활동 경험유무	예, 아니오
Target	현재 이직준비 여부	준비함, 준비안함

4. 데이터 균형화

데이터 마이닝 기법을 이용하여 예측모형을 구축할 때, 목표변수의 클래스 분포를 파악해야 한다. 목표변수의 관측 값이 특정 클래스에 많이 속하게 된다면 관측 대상이 많은 클래스를 중심으로 학습이 많이 이루어지게 되며, 이 경우 특정 클래스만을 잘 분류하게 되는 편향된 예측모형이 구축된다[21]. 따라서 예측모형을 구축하기 전에 목표변수 내에 존재하는 클래스들의 비율을 동등하게 맞추어 주는 데이터 균형화 작업이 필요하다. 본 연구의 목표변수인 'Target'의 값은 2개 클래스인 '준비함'과 '준비안함'으로 구성되어 있다. 전체 직종에 대한 통합 데이터의 경우 목표변수의 각 클래스별 관측 대상의 수는 '준비함'이 1,835명이고 '준비안함'이 9,048명이기 때문에 목표변수의 클래스 분포가 '준비안함' 방향으로 많이 편향되어 있는 것을 알 수 있다. 본 연구에서는 이를 해결하기 위해 10개의 표본을 추출할 때 언더샘플링(under-sampling)을 실시하여 클래스의 비율을 맞추는 작업을 진행하였다. 그 결과 표본별 3,670명(준비함: 1,835명, 준비안함: 1,835명)으로 이루어진 데이터 셋을 최종 분석 데이터 셋으로 결정하였다.

직종별 데이터의 경우에도 앞서 언급한 방법에 따라 직종별로 5개의 표본을 추출할 때 언더샘플링을 실시하였다. 그 결과, 표본별 사무는 1112명(준비함: 556명, 준비안함: 556명), 문화예술은 258명(준비함: 129명, 준비안함: 129명), 건설은 116명(준비함: 58명, 준비안함: 58명) 그리고 정보기술은 144명(준비함: 72명, 준비안함: 72명)으로 데이터를 균형화하여 최종 분석 데이터 셋을 구성하였다.

5. 변수 선택

앞서 언급한 39개의 독립변수들 중에서 예측모형을 구축하는데 아무런 도움을 주지 못하는 변수들이 존재할 수 있다. 목표변수의 예측에 큰 도움이 되지 않음에도 불구하고 예측모형이 구축될 때 이용되는 변수들이 있다면 예측모형의 정확성을 저하시킬 수 있다. 따라서 예측모형을 구축할 때 불필요한 변수들을 제거하는 작업이 필요하다[22]. 이와 같이, 여러 독립변수들 중 목표변수를 분류할 때 중요한 역할을 하는 변수들을 선택

하는 과정을 변수 선택이라고 한다. 본 연구에서는 변수를 선정하기 위한 알고리즘으로 이득비(gain ratio)를 사용하여 독립변수들의 중요도를 평가하였고, 예측모형을 구축하는데 이용될 독립변수들을 래퍼(wrapper) 방법 중 하나인 역방향 제거를 통해 선택하였다. 역방향 제거란 예측모형을 구축하는 때 회마다 가장 중요성이 떨어지는 변수를 하나씩 제거해 나가면서 예측모형을 구축하는 방식으로 중요도가 가장 높은 변수가 한 개 남을 때까지 예측모형을 반복해서 구축한다[23]. 이 과정을 통해서 변수의 개수만큼 예측모형이 구축되며, 구축된 예측모형들 중에서 정확도가 가장 높은 예측모형의 변수들을 최종 변수로 선택하게 된다.

IV. 연구 결과

1. 통합 데이터의 예측모형 구축 결과

본 연구에서는 통합 데이터로부터 이직준비 여부에 관한 예측모형을 구축하기 위해 오픈소스 데이터 마이닝 툴인 웨카(Weka) 버전 3.8을 이용하였으며, 의사결정나무, 베이즈넷, 인공신경망 알고리즘을 사용하여 예측모형을 구축하였다. [그림 1]의 '실험 1'과 같이, 통합 데이터로부터 언더샘플링 방식으로 10개의 표본들을 추출하였고, 추출된 표본들을 각각 66%의 학습데이터와 34%의 검증데이터 비율로 분할하여 실험을 진행하였다. 여기에서 학습데이터는 예측모형의 학습에 사용되며, 검증데이터는 학습데이터를 통해 구축된 예측모형의 성능을 평가할 때 사용된다. 선정된 39개의 독립변수들 중 중요도가 낮은 변수를 하나씩 제거하면서 예측모형의 예측률을 계산하였다. 그 결과 하나의 표본마다 서로 다른 3개의 알고리즘을 적용하고, 각 알고리즘마다 역방향 제거를 통한 39개의 예측모형을 구축하여 표본 10개에 대해 최종적으로 1,170개의 예측모형이 구축되었다. [표 2]는 10개의 표본에 대한 알고리즘 별 예측모형의 평균 예측률과 3개의 알고리즘이 각 표본에서 가장 높은 예측률을 기록한 횟수를 보여주고 있다. 그 결과 의사결정나무 기반 예측모형의 평균 예측률(65.737%)이 다른 예측모형들의 평균 예측률 보다 근소

한 차이로 높게 나타났고, 10개의 표본들 중에서 4개의 표본에서 최고 예측률을 보여주었다.

표 2. 통합 데이터의 예측모형 결과 비교

예측모형	평균 예측률	상위 예측률 보유 횟수
DT	65.737	4
BN	65.719	4
ANN	65.367	2

[표 3]은 10개의 표본 중에서 예측률이 가장 높은 예측모형을 지닌 표본 3개에 관한 정보를 보여주고 있다. 여기에서 변수 순서가 낮을수록 중요도가 낮은 변수이며, 1번 변수부터 마지막 변수 순으로 변수가 제거되면서 기록된 예측모형의 예측률을 보여주고 있다. 예를 들어, 표본 1과 표본 7의 경우 의사결정나무 기반의 예측모형에서 최고 예측률을 보여주었는데, 표본 1에서는 5개의 변수를 사용했을 때 66.82%, 표본 7에서는 6개의 변수를 사용했을 때 67.86%의 예측률을 기록하였다. 표

표 3. 통합 데이터의 표본별 예측모형 예측률 비교

순서	표본 1				표본 7				표본 9			
	변수	DT	BN	ANN	변수	DT	BN	ANN	변수	DT	BN	ANN
1	F15	61.85	66.67	57.37	F8	64.10	66.42	58.49	F2	62.74	64.98	61.05
2	F1	61.85	66.67	57.37	F15	64.10	66.42	58.49	F8	62.74	64.98	61.05
3	F7	61.85	66.58	59.29	F1	64.10	66.42	58.49	F15	62.74	64.98	61.05
4	F35	61.85	66.42	58.89	F7	64.10	66.50	59.93	F1	62.74	64.98	61.05
5	F34	61.85	66.34	59.21	F35	63.62	66.42	58.89	F7	61.85	64.82	55.92
6	F39	61.93	66.26	59.45	F34	63.70	66.10	60.41	F35	62.82	65.14	56.25
7	F6	62.09	66.18	59.05	F39	63.22	65.94	62.41	F34	62.82	64.98	57.61
8	F20	63.30	66.18	58.81	F6	63.06	66.10	60.65	F6	62.90	64.98	57.77
9	F8	63.30	66.50	60.97	F3	63.70	66.18	58.49	F39	62.98	65.22	57.13
10	F3	63.38	66.50	59.21	F4	63.94	65.86	60.57	F4	64.58	65.06	58.09
11	F19	63.14	66.10	60.57	F5	63.70	65.94	60.49	F19	64.58	64.82	58.25
12	F26	64.26	66.42	59.13	F20	63.46	65.94	60.49	F3	65.14	64.74	55.20
13	F4	64.26	65.86	61.13	F19	63.06	65.86	61.77	F22	64.50	64.58	56.16
14	F32	64.34	65.54	60.25	F27	63.06	65.78	59.69	F5	62.74	64.74	55.28
15	F22	64.66	66.10	60.09	F32	64.10	65.70	58.89	F26	61.53	64.50	57.05
16	F33	64.18	66.18	59.93	F25	64.18	65.46	59.85	F20	60.97	64.58	59.21
17	F27	64.26	65.38	60.73	F22	64.90	66.02	62.09	F37	61.45	64.50	59.69
18	F25	64.18	65.62	61.93	F33	65.70	66.50	60.33	F27	62.33	64.50	58.33
19	F5	64.42	65.54	61.93	F37	66.50	66.10	59.37	F32	62.66	65.30	58.49
20	F14	64.42	65.46	61.29	F26	65.54	66.58	58.65	F25	62.74	65.30	59.37
21	F37	65.14	65.46	61.37	F10	66.66	66.82	61.53	F29	62.90	65.54	58.81
22	F2	65.14	65.22	60.17	F29	64.90	66.34	58.73	F33	63.70	65.54	61.45
23	F13	61.45	65.22	60.01	F2	65.14	66.18	61.05	F28	63.54	65.54	61.29
24	F38	61.21	65.22	59.21	F28	65.14	66.18	61.61	F38	63.78	65.70	65.70
25	F28	61.93	66.10	59.85	F17	64.90	66.18	61.77	F23	63.86	65.78	61.37
26	F10	64.50	66.26	62.25	F38	66.80	66.02	64.50	F10	64.02	65.22	58.01
27	F29	62.82	65.54	63.22	F23	66.98	66.34	61.85	F18	63.54	64.82	62.09
28	F23	63.38	66.42	62.41	F18	66.10	66.42	63.86	F36	63.86	65.70	62.17
29	F17	63.46	66.18	63.30	F30	65.86	66.50	65.22	F17	64.34	64.82	62.58
30	F36	63.62	65.94	62.25	F36	66.26	67.30	64.90	F24	65.38	65.06	63.86
31	F16	63.62	66.10	62.82	F24	66.02	67.46	66.10	F13	65.06	65.62	64.50
32	F31	65.46	65.46	63.86	F31	66.98	67.38	67.14	F30	65.14	65.62	65.54
33	F18	66.26	65.94	63.14	F16	65.14	65.78	66.98	F31	65.62	66.90	66.10
34	F30	66.18	64.98	62.98	F21	67.86	67.30	67.62	F21	63.70	64.90	65.62
35	F21	66.82	65.06	67.06	F13	63.54	63.30	63.54	F14	62.58	63.78	62.41
36	F24	64.82	63.70	64.58	F14	63.54	63.30	63.62	F16	62.58	63.78	63.30
37	F11	63.22	62.82	62.58	F12	63.54	63.38	63.30	F11	62.41	62.98	62.41
38	F12	63.22	63.22	63.22	F9	61.37	61.45	61.45	F12	62.41	62.41	62.41
39	F9	60.89	60.89	60.89	F11	56.16	56.16	56.16	F9	61.61	61.61	61.61

본 9에서는 베이즈넷 기반 예측모형이 7개의 변수를 사용했을 때 가장 높은 예측률인 66.90%를 기록하였다.

앞서 분석된 결과를 바탕으로, 통합 데이터에서 전체 근로자의 이직준비 여부를 가장 잘 예측하는 모형은 의사결정나무 기반 예측모형임을 알 수 있다.

2. 직종별 데이터의 예측모형 구축 결과

직종별 데이터로부터 직종별 이직준비 여부에 관한 예측모형이 구축되는 과정을 요약하면 [그림 1]의 '실험 2'와 같다. 먼저 사무, 문화예술, 건설, 정보기술 분야로 직종별 데이터를 추출한 뒤, 직종별로 5개의 표본을 언더샘플링으로 생성하였다. '실험 2'에서는 '실험 1'에서 가장 좋은 성능을 보여준 알고리즘인 의사결정나무를 활용하여 예측모형을 구축하였다. 역방향 제거를 이용하여 표본 당 39개의 예측모형을 구축하였고, 직종마다 195개의 예측모형을 만들어 최종 780개의 예측모형을 구축하였다. [표 4]는 예측모형이 직종별 표본에서 가장 높은 예측률을 기록하였을 때의 결과를 보여준다.

표 4. 직종별 표본에서의 최고 예측률

직종	표본	최고 예측률
사무	표본 1	65.60
	표본 2	67.46
	표본 3	64.02
	표본 4	65.60
	표본 5	65.87
문화예술	표본 1	60.22
	표본 2	67.04
	표본 3	67.04
	표본 4	63.63
	표본 5	68.18
건설	표본 1	79.48
	표본 2	66.66
	표본 3	76.92
	표본 4	74.35
	표본 5	76.92
정보기술	표본 1	73.46
	표본 2	77.55
	표본 3	67.34
	표본 4	65.30
	표본 5	67.34

[표 5]는 직종별 최고 예측률을 기록한 표본의 결과를 나타내고 있다. 사무 분야는 표본 2에서 20개의 변수

표 5. 직종별 최고 예측률을 기록한 표본 결과

순서	사무		문화예술		건설		정보기술	
	표본 2		표본 5		표본 1		표본 2	
	변수	DT	변수	DT	변수	DT	변수	DT
1	F2	57.14	F2	52.27	F2	61.53	F2	51.02
2	F8	57.14	F8	52.27	F8	61.53	F8	51.02
3	F13	57.14	F13	52.27	F13	61.53	F11	51.02
4	F14	57.14	F14	52.27	F15	61.53	F13	51.02
5	F15	57.14	F15	52.27	F16	61.53	F14	51.02
6	F6	57.14	F16	52.27	F39	61.53	F15	51.02
7	F34	57.14	F33	52.27	F27	61.53	F16	51.02
8	F7	57.14	F27	52.27	F10	66.66	F6	51.02
9	F1	58.73	F34	52.27	F1	66.66	F1	51.02
10	F27	58.73	F39	52.27	F31	61.53	F34	51.02
11	F32	58.73	F1	52.27	F34	53.84	F25	51.02
12	F25	58.73	F38	52.27	F21	53.84	F23	51.02
13	F39	57.67	F19	51.13	F26	53.84	F24	51.02
14	F35	60.58	F11	50.00	F6	53.84	F19	51.02
15	F37	61.64	F29	51.13	F25	53.84	F32	51.02
16	F19	61.37	F5	48.86	F19	53.84	F27	48.97
17	F22	61.90	F26	54.54	F33	71.79	F28	71.42
18	F20	61.90	F6	54.54	F35	71.79	F20	71.42
19	F26	65.87	F25	52.27	F18	61.53	F10	71.42
20	F33	67.46	F20	57.95	F32	61.53	F21	77.55
21	F4	67.19	F32	57.95	F4	48.71	F39	77.55
22	F29	67.19	F10	59.09	F3	48.71	F26	77.55
23	F38	66.93	F23	60.22	F30	43.58	F36	69.38
24	F23	66.93	F35	60.22	F24	43.58	F31	69.38
25	F3	66.66	F28	61.36	F28	41.02	F22	69.38
26	F5	66.66	F22	61.36	F23	41.02	F17	69.38
27	F28	67.19	F24	65.90	F17	38.46	F33	69.38
28	F18	64.55	F36	65.90	F29	38.46	F18	69.38
29	F31	65.60	F37	65.90	F37	41.02	F12	69.38
30	F36	66.40	F12	65.90	F7	41.02	F7	69.38
31	F30	66.40	F4	65.90	F5	41.02	F3	69.38
32	F24	65.34	F3	65.90	F38	79.48	F4	69.38
33	F21	65.60	F21	68.18	F36	79.48	F29	69.38
34	F10	64.02	F17	67.04	F12	79.48	F37	65.30
35	F17	64.02	F9	64.77	F20	79.48	F35	65.30
36	F16	62.16	F31	63.63	F22	41.02	F38	65.30
37	F11	62.16	F7	63.63	F11	41.02	F30	65.30
38	F12	62.16	F18	59.09	F9	41.02	F9	57.14
39	F9	59.25	F30	50.00	F14	41.02	F5	42.85

를 사용하였을 때 67.46%의 예측률을 보여주었다. 문화 예술 분야는 표본 5에서 7개의 변수를 사용하였을 때 68.18% 그리고 건설 분야는 표본 1에서 5개의 변수를 사용하였을 때 79.48%의 예측률을 기록하였다. 또한 정보기술 분야는 표본 2에서 18개의 변수를 사용하였을 때 77.55%의 예측률을 얻을 수 있었다.

지금까지의 결과를 통해 사무 분야를 제외한 나머지 분야들의 예측모형에서 전체 직종의 예측모형 보다 높은 예측률을 기록한 것을 알 수 있다. 이를 통해 예측모형은 전체 직종의 데이터를 사용하는 것 보다는 데이터들 간의 유사도가 높은 직군별 데이터를 사용하였을 때 성능이 향상되는 것을 확인할 수 있다.

V. 고용유지 강화 방안

1. 근로자 전반에 관한 고용유지 방안

본 절에서는 ‘실험 1’에서 가장 좋은 성능을 보여준 의사결정나무 기반 예측모형의 결과를 활용하여 전체 직종에 관한 이직준비 여부 규칙을 도출하고, 근로자 전반에 관한 고용유지 강화 방안을 제시하고자 한다. 일반적으로 의사결정나무로 구축된 예측모형의 경우, 의사결정나무를 구성하는 분류규칙들이 이해하기 쉬운

형태로 만들어지기 때문에 규칙을 통한 결과 해석이 용이한 장점이 있다[23].

[표 3]의 표본 7에 제시된 의사결정나무 기반 예측모형들 중, 예측률이 가장 높은 예측모형의 분류규칙을 활용하여 근로자 전반에 관한 이직준비 여부 규칙을 도출하고자 한다. 그 결과 6가지 변수인 F11(근로시간 형태), F9(종사상 지위), F12(정규직 여부), F14(주당 정규 근로시간), F13(주당 정규 근로일), F21(개인의 발전가능성)을 사용하여 의사결정나무를 구축하였을 때 가장 높은 예측률인 67.86%를 얻을 수 있었다. [그림 2]는 의사결정나무 기반 예측모형에서 도출된 12개의 전체 규칙을 보여준다. 이직준비 여부에 관한 규칙은 7개의 ‘준비함’ 규칙과 5개의 ‘준비안함’ 규칙으로 구성된다.

[표 6]은 [그림 2]에서 보여준 12개의 규칙에 관한 세부 내용들을 설명하고 있다. 본 연구에서는 각 규칙들의 주요 특성들을 분석하고 이를 토대로 근로자 전반에 관한 고용유지 강화 방안을 제안하고자 한다.

첫 번째, 규칙 1, 2, 3을 비교하여 살펴보면, 근로자의 종사상 지위가 임시근로자(고용계약기간이 1개월 이상 1년 미만인 경우 또는 고용계약이 정해지지 않은 경우 일정한 사업의 필요에 의해 고용된 사람)이거나 일용근로자(고용계약기간이 1개월 미만이거나 매일매일 고용되어 근로의 대가로 일급 또는 일당제 급여를 받고 일

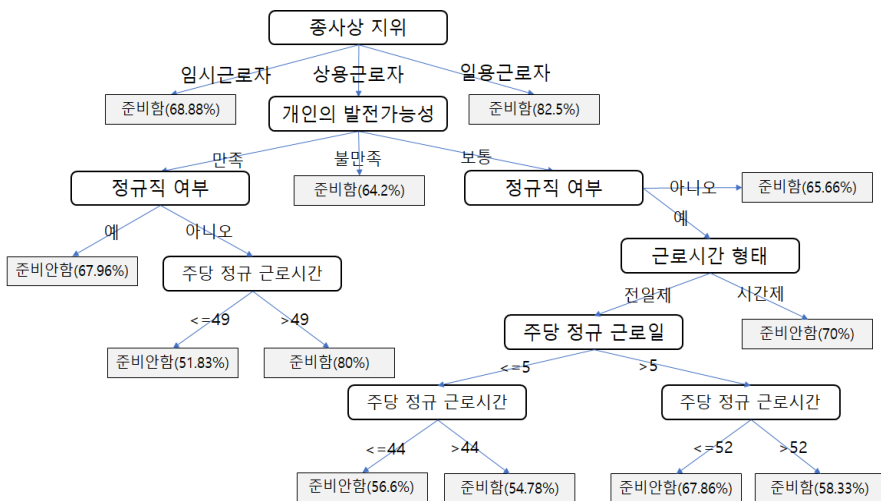


그림 2. 의사결정나무로 도출된 이직준비 여부에 관한 규칙

표 6. 근로자 전반에 관한 이직준비 여부 규칙

규칙	내용	예측 결과	예측률	예측수
1	종사상 지위=임시근로자	준비함	68.88%	813
2	종사상 지위=일용근로자	준비함	82.50%	80
3	종사상 지위=상용근로자 & 개인의 발전가능성=불만족	준비함	64.20%	500
4	종사상 지위=상용근로자 & 개인의 발전가능성=만족 & 정규직 여부=예	준비안함	67.96%	1233
5	종사상 지위=상용근로자 & 개인의 발전가능성=만족 & 정규직 여부=아니오 & 주당 정규 근로시간<=49	준비안함	51.83%	137
6	종사상 지위=상용근로자 & 개인의 발전가능성=만족 & 정규직 여부=아니오 & 주당 정규 근로시간>49	준비함	80.00%	10
7	종사상 지위=상용근로자 & 개인의 발전가능성=보통 & 정규직 여부=아니오	준비함	65.66%	99
8	종사상 지위=상용근로자 & 개인의 발전가능성=보통 & 정규직 여부=예 & 근로시간 형태=시간제	준비안함	70.00%	10
9	종사상 지위=상용근로자 & 개인의 발전가능성=보통 & 정규직 여부=예 & 근로시간 형태=전일제 & 주당 정규 근로일<=5 & 주당 정규 근로시간<=44	준비안함	56.60%	553
10	종사상 지위=상용근로자 & 개인의 발전가능성=보통 & 정규직 여부=예 & 근로시간 형태=전일제 & 주당 정규 근로일<=5 & 주당 정규 근로시간>44	준비함	54.78%	115
11	종사상 지위=상용근로자 & 개인의 발전가능성=보통 & 정규직 여부=예 & 근로시간 형태=전일제 & 주당 정규 근로일>5 & 주당 정규 근로시간<=52	준비안함	67.86%	84
12	종사상 지위=상용근로자 & 개인의 발전가능성=보통 & 정규직 여부=예 & 근로시간 형태=전일제 & 주당 정규 근로일>5 & 주당 정규 근로시간>52	준비함	58.33%	36

하는 사람)일 경우 이직을 준비하는 것으로 분석되었다. 그리고 종사상 지위가 상용근로자(고용계약기간이 1년 이상인 정규직원 또는 특별한 고용계약이 없이 기간이 정해져 있지 않아도 소정의 채용절차에 의해 입사하여 인사관리 규정을 적용받는 사람)일 경우 개인의 발전가능성에 대해 만족하지 못할 경우 이직을 준비하는 것으로 나타났다. 이를 통해, 무엇보다 현재 근로자의 종사상 지위가 이직준비 여부에 가장 큰 영향을 주는 것을 알 수 있다. 따라서, 기업에서는 근로자의 고용유지를 위해서 근로자에게 고용계약을 통한 직업의 안정성을 우선적으로 보장해 주어야 한다. 또한 직업의 안정성이 보장된 상용근로자의 경우에는 고용유지를 위해 개인의 발전가능성을 높일 수 있는 다양한 기회를 마련해 주어야 한다.

두 번째, 규칙 5, 6을 살펴보면, 상용근로자 중 현재 일자리에서 개인의 발전가능성에 대해 만족하는 비정규직의 경우, 주당 정규 근로시간을 고려하여 이직준비를 결정하는 것을 알 수 있다. 즉, 비정규직이더라도 개인의 발전가능성에 대해 만족하고 주 5일 근무제 기준으로 하루 근로시간이 9.8시간을 넘지 않을 경우 이직을 준비하지 않는 경향이 나타났다. 하지만, 동일한 조

건에서 하루 근로시간이 9.8시간을 넘을 경우 이직을 준비하는 것을 알 수 있다. 따라서, 기업에서는 상용근로자 중 비정규직 근로자의 이직을 방지하기 위해서 개인의 발전가능성에 만족할 수 있는 프로그램의 개발과 적정 근로시간을 보장해 주는 방안이 마련되어야 한다.

세 번째, 규칙 9, 10, 11, 12에 의하면, 기업의 전일제 정규직의 경우 주당 정규 근로일과 주당 정규 근로시간을 고려하여 이직준비를 생각하는 것을 알 수 있다. 즉, 주당 정규 근로일이 5일 이하일 경우 일주일에 44시간, 하루에 8.8시간을 기준으로 이직준비를 결정하였으며, 주당 정규 근로일이 5일을 초과할 경우 일주일에 52시간(근무일 7일 기준 약 7.43시간 또는 6일 기준 약 8.67시간)을 넘으면 이직을 준비하는 것으로 분석되었다. 따라서 기업에서는 정규직 직원의 이직을 방지하기 위해서 적정 근로일과 근로시간을 보장해 주어야 하며, 초과 업무가 발생될 경우 이에 상응하는 적절한 보상체계가 제공되어야 한다.

네 번째, 규칙 8과 같이 개인의 발전가능성이 보통이어서도 정규직이고 시간제 근로가 가능할 경우에는 이직준비를 하지 않는다. 특정 시간에 일을 하지만 근로조건이 정규직과 동일한 대우를 받는 근로자들이 여기에

속하는데, 이들 대부분은 학업이나 출산과 같은 개인적인 상황에 의해 시간제 근로를 실시하고 있다. 따라서 기업에서는 특정 상황에 놓이게 된 정규직의 이직을 방지하기 위해서 단시간 근로제, 탄력적 근로시간제, 재택근무제 등과 같은 다양한 유연근무제도를 도입하는 방안이 고려되어야 한다.

2. 직종별 근로자의 고용유지 방안

본 절에서는 직종별 데이터 분석을 통해 특정 직종에 차별화된 이직준비 요인들을 발견하고, 각 직종별 특성을 고려한 고용유지 방안을 제안하고자 한다. 이를 위해, [표 5]에서 제시된 의사결정나무 기반 예측모형의 직종별 최고 예측률에 관한 결과를 활용하고자 하며, 직종별 근로자의 고용유지 방안은 ‘준비함’ 규칙을 중심으로 설명하고자 한다.

먼저 [표 5]의 사무 분야 결과를 통해 사무직 근로자의 이직준비 여부에 영향을 주는 20개의 요인과 10개의 이직준비 규칙을 파악하였다. 사무직에서는 F9(종사상 지위)가 가장 영향력이 큰 요인으로 나타났다. 근로자 전반에 영향을 주는 요인들 중에서는 F9(종사상 지위),

F12(정규직 여부), F11(근로시간 형태), F21(개인의 발전가능성)이 사무직 근로자의 이직준비 여부에 영향을 미치는 것으로 나타났다.

[표 7]에 제시된 사무직 근로자들의 이직준비 규칙을 정리하면 다음과 같다. 첫째, 규칙 1과 같이 사무직에서도 종사상 지위가 임시근로자이면 가장 많이 이직을 준비하는 것을 확인할 수 있다. 둘째, 규칙 2, 3, 4, 5, 9, 10을 살펴보면 상용근로자 중 임금 및 소득에 대한 만족도가 낮고 미혼이면 이직을 준비하는 것을 알 수 있다. 셋째, 규칙 6, 7, 8을 통해 임금 및 소득에 대한 만족도가 보통이더라도 인사체계 및 승진제도에 만족하지 못한다면 이직을 준비하는 것으로 나타났다. 따라서 사무직 근로자들의 고용유지를 강화하기 위해서는 임시근로자의 비율을 점진적으로 줄이면서, 상용근로자를 위한 합리적인 임금체계와 공정한 인사체계를 마련하는 노력이 필요하다.

다음으로 [표 5]의 문화예술 분야 결과를 통해 문화예술직 근로자의 이직준비 여부에 영향을 주는 7개의 요인과 9개의 이직준비 규칙을 발견하였다. 문화예술직에서는 F30(자신의 교육수준 대비 일의 수준)이 가장

표 7. 사무직 근로자들의 이직준비 규칙

규칙	내용	예측결과	예측률	예측수
1	종사상 지위=임시근로자	준비함	71.90%	242
2	종사상 지위=상용근로자 & 임금 및 소득=불만족 & 혼인상태=미혼 & 인사체계 및 승진제도=보통	준비함	70.45%	44
3	종사상 지위=상용근로자 & 임금 및 소득=불만족 & 혼인상태=미혼 & 인사체계 및 승진제도=불만족 & 개인의 발전가능성=불만족	준비함	68.29%	41
4	종사상 지위=상용근로자 & 임금 및 소득=불만족 & 혼인상태=미혼 & 인사체계 및 승진제도=만족	준비함	84.38%	32
5	종사상 지위=상용근로자 & 임금 및 소득=불만족 & 혼인상태=미혼 & 인사체계 및 승진제도=불만족 & 개인의 발전가능성=보통	준비함	58.62%	29
6	종사상 지위=상용근로자 & 임금 및 소득=보통 & 인사체계 및 승진제도=불만족 & 전공지식 도움정도=도움안됨	준비함	75.00%	28
7	종사상 지위=상용근로자 & 임금 및 소득=보통 & 인사체계 및 승진제도=불만족 & 전공지식 도움정도=도움됨	준비함	75.00%	20
8	종사상 지위=상용근로자 & 임금 및 소득=보통 & 인사체계 및 승진제도=보통 & 월 평균 근로 소득(<=283 & 자신의 교육수준 대비 일의 수준=낮음 & 전공지식 도움정도=도움안됨	준비함	56.25%	16
9	종사상 지위=상용근로자 & 임금 및 소득=불만족 & 혼인상태=미혼 & 인사체계 및 승진제도=보통 & 계약기간 지정여부=예	준비함	86.67%	15
10	종사상 지위=상용근로자 & 임금 및 소득=불만족 & 혼인상태=미혼 & 인사체계 만족도=보통 & 계약기간 지정여부=아니오 & 자신의 교육수준 대비 일의 수준=낮음 & 직무관련 교육 또는 훈련 만족도=불만족	준비함	84.62%	13

표 8. 문화예술직 근로자들의 이직준비 규칙

규칙	내용	예측결과	예측률	예측수
1	자신의 교육수준 대비 일의 수준=낮음	준비함	78.18%	55
2	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=불만족	준비함	80.95%	21
3	자신의 교육수준 대비 일의 수준=높음 & 종사상 지위=임시근로자	준비함	80.00%	5
4	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=만족 & 이용 교통수단=도보	준비함	100%	4
5	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=보통 & 자신의 수준 대비 일의 기술 수준=낮음 & 임금 및 소득=불만족	준비함	100%	3
6	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=보통 & 자신의 수준 대비 일의 기술 수준=낮음 & 임금 및 소득=만족	준비함	100%	2
7	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=보통 & 자신의 수준 대비 일의 기술 수준=보통 & 임금 및 소득=보통 & 종사상 지위=임시근로자	준비함	62.50%	8
8	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=보통 & 자신의 수준 대비 일의 기술 수준=보통 & 임금 및 소득=불만족 & 종사상 지위=상용근로자	준비함	75.00%	12
9	자신의 교육수준 대비 일의 수준=보통 & 고용 안정성=보통 & 자신의 수준 대비 일의 기술 수준=보통 & 임금 및 소득=불만족 & 종사상 지위=일용근로자	준비함	50.00%	2

영향력이 큰 요인으로 분석되었다. 근로자 전반에 영향을 주는 요인들 중 F9(종사상 지위)와 F21(개인의 발전 가능성)가 문화예술직 근로자의 이직준비 여부에 영향을 주는 것으로 확인되었다.

[표 8]에 나타난 문화예술직 근로자들의 이직준비 규칙을 살펴보면 다음과 같다. 첫째, 규칙 1을 살펴보면 문화예술직에서는 학력과 같이 자신의 교육수준보다 하고 있는 일의 수준이 낮을 경우 가장 많이 이직을 준비하는 것을 알 수 있다. 둘째, 규칙 2를 통해 자신의 교육수준 대비 일의 수준이 보통이더라도 고용 상태가 불안정하다면 이직을 준비하는 것으로 나타났다. 셋째, 규칙 8과 같이, 임금에 대해 불만족할 경우 상용근로자라도 이직준비를 하는 것을 알 수 있다. 즉, 문화예술직 근로자들의 고용유지를 위해서 업무의 수준에 맞는 교육수준을 지닌 근로자를 고용하는 것이 중요하며, 추가로 근로자의 고용 안정을 보장하고 임금에 대한 만족도

를 높일 수 있는 방안이 필요하다.

건설직 근로자의 이직준비에 영향을 주는 요인은 [표 5]의 건설 분야 결과를 통해 파악하였으며, 그 결과 5개의 영향 요인과 4개의 이직준비 규칙을 발견하였다. 건설직에서 가장 영향력이 큰 요인은 F14(주당 정규 근로 시간)이며, F9(종사상 지위), F11(근로시간 형태), F22(인간관계), F20(근로시간 및 근로일수)의 순으로 영향을 준다. 근로자 전반에 영향을 주는 요인들 중 F14(주당 정규 근로시간), F9(종사상 지위), F11(근로시간 형태)이 건설직 근로자의 이직준비 여부에 영향을 주는 것을 알 수 있었다.

[표 9]에서 제시된 건설직 근로자들의 이직준비 규칙을 요약하면 다음과 같다. 첫째, 규칙 1, 2를 통해 건설직 근로자가 임시근로자이거나 일용근로자일 경우 이직을 준비하고 있음을 알 수 있다. 둘째, 규칙 3, 4를 살펴보면 직업 안정성이 보장된 상용근로자라도 근무시

표 9. 건설직 근로자들의 이직준비 규칙

규칙	내용	예측결과	예측률	예측수
1	종사상 지위=임시근로자	준비함	64.29%	14
2	종사상 지위=일용근로자	준비함	100%	4
3	종사상 지위=상용근로자 & 근무시간 및 근무일수=보통	준비함	58.33%	24
4	종사상 지위=상용근로자 & 근무시간 및 근무일수=불만족	준비함	65.52%	29

표 10. 정보기술직 근로자들의 이직준비 규칙

규칙	내용	예측결과	예측률	예측수
1	종사상 지위=임시근로자	준비함	90.90%	11
2	종사상 지위=상용근로자 & 직무관련 교육 또는 훈련=불만족	준비함	76.23%	26
3	종사상 지위=상용근로자 & 직무관련 교육 또는 훈련=보통 & 혼인 상태=미혼	준비함	100%	2
4	종사상 지위=상용근로자 & 직무관련 교육 또는 훈련=만족 & 일의 자율성과 권한=불만족	준비함	66.67%	3
5	종사상 지위=상용근로자 & 직무관련 교육 또는 훈련=만족 & 일의 자율성과 권한=보통	준비함	81.82%	11
6	종사상 지위=상용근로자 & 직무관련 교육 또는 훈련=만족 & 일의 자율성과 권한=만족 & 혼인상태=미혼 & 임금 및 소득=만족	준비함	80.00%	5

간이나 근무일수에 만족하지 못한다면 이직을 준비하는 것으로 밝혀졌다. 따라서 건설직 근로자들의 고용유지를 위해서는 고용 안정과 적정 근무시간을 보장하는 제도가 마련되어야 한다.

마지막으로, [표 5]의 정보기술 분야 결과를 통해 정보기술직 근로자의 이직준비 여부에 영향을 주는 18개의 요인과 6개의 이직준비 규칙을 파악하였다. 정보기술직에서는 F5(가족의 경제적 지원 여부)가 가장 영향력이 큰 요인으로 나타났으며, 근로자 전반에 영향을 주는 요인들 중에서는 F9(종사상 지위)와 F12(정규직 유무)가 정보기술직 근로자의 이직준비 여부에 영향을 주는 것으로 밝혀졌다.

[표 10]에 나타난 정보기술직 근로자들의 이직준비 규칙을 살펴보면 다음과 같다. 첫째, 규칙 1을 보면 정보기술직에서도 임시근로자일 경우 이직을 준비하고 있음을 알 수 있다. 둘째, 규칙 2, 3을 살펴보면 상용근로자이더라도 회사에서 제공하는 직무관련 교육에 대해 만족하지 못할 경우 이직을 준비하는 것으로 나타났다. 셋째, 규칙 4, 5를 통해 상용근로자이면서 회사의 교육 프로그램에 만족을 하더라도 일의 자율성과 권한 부분에서 만족하지 못하면 이직을 준비하는 것을 알 수 있다. 넷째, 규칙 6에서는 앞서 언급한 여러 요인들에 대해 전반적으로 만족하고 있더라도 근로자가 미혼이라면 이직을 준비하는 경향이 있음을 알 수 있다. 따라서 정보기술직 근로자들의 고용유지를 위해서 다양한 직무관련 교육을 실시하고 업무 진행에 있어서 자율성과 권한을 부여하는 방안이 마련되어야 한다.

VI. 결론

본 연구에서는 기업의 지속적인 인력 유지를 위해, 전체 직종과 특정 직종으로 구분된 근로자를 대상으로 이직준비 여부에 영향을 미치는 요인들을 각각 분석하고, 근로자들의 조직이탈 방지와 고용유지에 도움을 주는 실용적인 방안들을 제안하고자 하였다. 이를 위해 데이터 마이닝 기법을 활용하여 전체 직종과 특정 직종에 속한 근로자의 이직준비 여부를 예측하는 모형을 구축하였고, 그 결과를 활용하여 이직준비 여부에 관한 규칙과 고용유지 강화 방안을 제안하였다.

주요 내용을 요약하면 다음과 같다. 전체 직종에 속한 근로자들의 경우 ‘근로시간 형태’, ‘종사상 지위’, ‘정규직 여부’, ‘주당 정규 근로시간’, ‘주당 정규 근로일’, ‘개인의 발전가능성’ 요인이 이직준비 여부에 영향을 주는 것으로 분석되었다. 주요 요인을 통해 도출된 12가지의 이직준비 여부 규칙을 활용하여 직업의 안정성 보장, 개인의 발전가능성 증진, 적정 근로일과 근로시간의 보장, 그리고 다양한 근무제도 도입의 관점에서 근로자의 고용유지 방안을 제시하였다.

직종별 근로자의 경우에서도 전체 직종에서의 분석결과와 비슷하게 종사상의 지위가 이직준비에 많은 영향을 주는 것으로 확인되었다. 그러나 일부 요인에 있어서는 직종별 차이를 보여주었다. 사무 분야의 경우 ‘인사체계 및 승진제도’와 ‘임금 및 소득’, 문화예술 분야의 경우 ‘자신의 교육수준 대비 일의 수준’과 ‘자신의 수준 대비 일의 기술수준’, 건설 분야는 ‘근무시간 및 근무

일수'와 '인간관계', 그리고 정보기술 분야의 경우 '직무 관련 교육 또는 훈련'과 '일의 자율성과 권한'이 직종별로 구분되는 영향 요인으로 분석되었다.

따라서 향후 본 연구에서 제안한 분석 방법과 고용유지 강화 방안들을 기업들이 활용할 경우 기업 근로자들의 이직률을 낮추는데 도움을 줄 수 있을 것으로 판단된다.

데이터 마이닝 분석의 특성상 분석 결과는 수집된 데이터에 영향을 받게 된다. 본 연구에서는 사전에 설문 조사된 데이터를 바탕으로 분석을 진행하였기 때문에 실시간 변동하는 노동시장의 환경 요인과 그로 인해 발생하는 다양한 외적 요소들을 분석에 반영하지 못한 점과 특정 직종의 표본 수가 상대적으로 적은 점이 연구의 한계점으로 인식된다.

참 고 문 헌

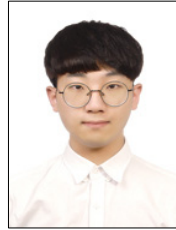
- [1] <http://kosis.kr/index/index.do>, 2018.4.24.
- [2] W. H. Mobley, *Employee Turnover, Causes, Consequences, and Control*. Reading, Addison Wesley, 1982.
- [3] A. C. Bluedorn, "The Theories of Turnover: Causes, Effects and Meaning," *Research in the Sociology of Organizations*, Vol.1, pp.75-128, 1982.
- [4] 박인규, "호텔 종사원의 직무배태성이 이직 의도에 미치는 영향에 관한 연구-임금만족을 조절변수로," *대한관광경영학회*, 제25권, 제3호, pp.61-79, 2010.
- [5] R. M. Steers and R. T. Mowday, "Employee turnover and post-decision accommodation processes," In: L. L. Cummings and B. M. Staw (Eds.). *Research in organizational behavior*, JAI Press, pp.235-281, 1981.
- [6] C. R. Forrest, L. L. Cummings, and A. C. Johnson, "Organizational Participation: A Critique and Model," *Academy of Management Review*, Vol.2, No.4, pp.586-601, 1977.
- [7] J. R. Hackman and G. R. Oldham, "Development of Job Diagnostic Survey," *Journal of Applied Psychology*, Vol.60, No.2, pp.159-170, 1975.
- [8] 이재명, 강신기, "금융기관 종사자의 이직 및 창업 의도에 영향을 미치는 요인 연구," *벤처창업연구*, 제10권, 제5호, pp.151-163, 2015.
- [9] 이상희, 김혁수, "외식업체 종사원의 이직경험에 따른 직무만족과 이직의도 영향관계," *호텔경영학연구*, 제23권, 제3호, pp.245-258, 2014.
- [10] 윤은자, 권영미, 전미순, 안정화, "간호사 이직 의도 영향요인," *한국콘텐츠학회논문지*, 제16권, 제1호, pp.94-106, 2016.
- [11] 이우경, 최수일, "IT업계 종사자들의 이직의도 결정요인," *한국콘텐츠학회논문지*, 제11권, 제5호, pp.369-383, 2011.
- [12] 지민경, "직무만족 및 전문직업성과 이직 의도와 의 관계 분석-치과위생사를 중심으로," *한국엔터테인먼트산업학회논문지*, 제9권, 제4호, pp.343-350, 2015.
- [13] 김보인, "근로자 이직 의도에 영향을 미치는 요인에 대한 연구-근로자 구성특성 별 비교," *인적자원관리연구*, 제24권, 제5호, pp.31-47, 2017.
- [14] 이진복, 박원준, "상사의 이직의도가 부하의 이직의도에 미치는 영향: 이직의도, LMX, POS의 상호작용 효과를 중심으로," *기업경영연구*, 제24권, 제3호, pp.117-136, 2017.
- [15] 성지미, 안주엽, "일자리 만족도와 이직의사 및 이직-청년층을 중심으로," *산업노동연구*, 제22권, 제2호, pp.135-179, 2016.
- [16] 김경범, 이주현, 노진원, 권영대, "청년 취업자의 이직 준비 관련 요인: 청년패널 2007-2003 자료 분석," *한국콘텐츠학회논문지*, 제16권, 제12호, pp.480-491, 2016.
- [17] 김진희, "인사제도의 내재화 직무몰입 역할피로 및 이직의도와의 관계에 대한 연구," *노동정책연구*, 제12권, 제3호, pp.119-45, 2012.
- [18] 김민주, "여행업 종사자의 개인특성에 따른 이직

의도 및 영향요인에 대한 지각 차이 분석,” 여가 관광연구, 제8권, pp.91-105, 2005.

- [19] 이선로, 김기영, “국내 IT 요원의 이직의도 결정 요인에 관한 실증 연구: 리더십과 임파워먼트 중심으로,” 정보화정책, 제13권, 제4호, pp.85-107, 2006.
- [20] 이행렬, 신용석, “사회복지전담공무원의 이직의도에 영향을 미치는 요인,” GRI 연구논총, 제14권, 제1호, pp.141-172, 2012.
- [21] N. V. Chawla, “Data Mining for Imbalanced Datasets: An Overview,” In: O. Maimon and L. Rokach (Eds.), Data Mining and Knowledge Discovery Handbook, Springer, 2005.
- [22] M. Dash and H. Liu, “Feature Selection for Classification,” Intelligent Data Analysis, Vol.1, No.3, pp.131-156, 1997.
- [23] I. H. Witten and E. Frank, *Data mining: Practical Machine Learning Tools and Techniques*, Morgan Kaufmann Publishers, 2005.

김 태 운(Taeun Kim)

준회원



- 2013년 3월 ~ 현재 : 경상대학교 경영정보학과 학사과정

<관심분야> : 데이터마이닝, 빅데이터분석 등

유 동 희(Donghee Yoo)

정회원



- 2002년 2월 : 고려대학교 MIS (경영학사)
- 2009년 2월 : 고려대학교 경영학과 MS/IS전공 (경영학박사)
- 2009년 6월 ~ 2013년 5월 : 육군사관학교 전자정보학과 조교수

• 2015년 3월 ~ 현재 : 경상대학교 경영정보학과 부교수, 경영경제연구소
<관심분야> : 데이터마이닝, 빅데이터분석, 딥러닝, 인공지능, 시맨틱웹 등

저 자 소 개

안 민 욱(Minuk An)

준회원



- 2013년 3월 ~ 현재 : 경상대학교 경영정보학과 학사과정

<관심분야> : 데이터마이닝, 빅데이터분석 등