


지역별 예측모델을 통한 사각지대 발굴 연구

이우식·윤기찬·박규범·이인수

지역별 예측모델을 통한 사각지대 발굴 연구

연구책임자 : 이 우 식 부연구위원
공동연구원 : 윤 기 찬 연구위원
박 규 범 연구원
이 인 수 연구원



머리말

사회보장정보원은 복지사각지대 고위험 대상자 선정을 위해 복지사각지대 발굴 시스템을 운영하고 있다. 복지사각지대 발굴관리시스템은 외부 데이터를 활용해 대상자를 구성한 후 예측 모형을 구동하여 발굴 대상자를 추출한 후 대상자 규모 산정과정을 통해 특정 발굴 대상자를 검증 후 지자체에 통보한 후 지자체에서 자체적인 검증을 통해 대상자에게 서비스를 제공하는 방식으로 진행된다. 하지만 과거 방식은 지역에 대한 고려 없이 전체 지역에 대한 예측만을 수행하였다. 그래서 특정 지역에 쏠림 현상이 발생되곤 하였다.

본 보고서에서는 지역적인 특성을 반영하기 위해서 지역별 특성이 고려된 복지사각지대 대상자 발굴 시스템을 활용한 연구로 3가지 관점에서 살펴보았다. 첫째, 복지사각지대 지역단위 서울, 경기, 전북, 제주의 데이터 분석을 수행하였다. 둘째, 지역별 군집을 통해 머신러닝 알고리즘 개발 및 실험을 통해서 특정 지역이 우수하다는 결과를 도출한다. 마지막으로 본 연구로부터 복지사각지대 발굴 시스템 개선을 위해 나아가야할 방향 및 정책적 제언을 하였다.

본 연구는 이우식 부연구위원의 책임 하에 윤기찬 연구위원, 박규범 연구원과 이인수 연구원이 공동으로 참여하여 수행되었다 참여 연구진외에 본 연구 관련 전문가들의 많은 조언 및 제언이 있었다. 또한

연구보고서의 처음부터 탈고에 이르기까지 조언과 비판을 아끼지 않은
본원의 사례관리정책지원부 직원 및 보건복지부 복지정보기획과 담당
공무원들께도 심심한 감사의 뜻을 전한다.

끝으로 보고서에 수록된 모든 내용은 연구진의 의견이며, 본 원의
공식 견해가 아님을 밝힌다.

2019년 12월
사회보장정보원장
임 희 택

목 차

요 약	i
제1장 서론	3
제1절 연구의 배경 및 목적	3
제2절 연구방법 및 내용	7
제2장 이론적 배경	15
제1절 국내외 선행연구	15
제2절 머신러닝 알고리즘과 XGBoost	20
제3절 모형 평가 방법	23
제3장 데이터 탐색	27
제1절 데이터 정의	27
제2절 전체 분석데이터 대상자 속성분석	30
제3절 단독변수를 가지는 대상자 속성분석	47
제4절 가구 수에 따른 대상자 속성분석	58
제5절 지역별 대상자 속성분석	81
제4장 예측모형 구현 및 분석	115
제1절 예측 모형 구현	115
제2절 지역별 예측 모형의 분포도 및 가중치 분석	120
제3절 지역별 예측 모형 간 성능 비교 분석	155
제4절 파생 변수 투입에 따른 효과 분석	172

제5장 결론	199
제1절 결론 및 제언	199
참고문헌	207

표 목 차

<표 1-1> 협의기관 역할 및 연관 데이터	9
<표 2-1> 분류행렬 (Confusion Matrix)	23
<표 3-1> 훈련데이터(1차,2차)	27
<표 3-2> 복지사각지대 코드 및 설명	28
<표 3-3> 연령구분	28
<표 3-4> 복지사각지대 선정 대상자 및 조치 현황분석 (19년 1차, 2차)	30
<표 3-5> 시도별 19년 1, 2차 복지사각지대 선정 대상자 현황	31
<표 3-6> 연령대별 복지사각지대 선정 대상자 현황	32
<표 3-7> 연계변수별 복지사각지대 선정 대상자 현황	33
<표 3-8> 연계변수 조합 개수별 복지사각지대 선정 대상자 현황	34
<표 3-9> 지역별 선정 대상자 2019년 1차 '완료결과' 현황	36
<표 3-10> 지역별 선정 대상자 2019년 2차 '완료결과' 현황	37
<표 3-11> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	38
<표 3-12> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	39
<표 3-13> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	40
<표 3-14> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	41
<표 3-15> 선정 대상자 2019년 1차의 변수별 '완료결과' 현황	43
<표 3-16> 선정 대상자 2019년 2차 변수별 '완료결과' 현황	45
<표 3-17> 지역별 선정 대상자 2019년 1차 '완료결과' 현황	47
<표 3-18> 지역별 선정 대상자 2019년 2차 '완료결과' 현황	49
<표 3-19> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	50
<표 3-20> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	51
<표 3-21> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	52
<표 3-22> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	53
<표 3-23> 선정 대상자 1개 변수 보유자의 2019년 1차 '완료결과' 현황	54

<표 3-24> 선정 대상자 2019년 2차 '완료결과' 현황	56
<표 3-25> 지역별 선정 대상자 2019년 1차 '완료결과' 현황	58
<표 3-26> 지역별 선정 대상자 2019년 2차 '완료결과' 현황	60
<표 3-27> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	61
<표 3-28> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	62
<표 3-29> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	63
<표 3-30> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	64
<표 3-31> 선정 대상자 2019년 1차 '완료결과' 현황	66
<표 3-32> 선정 대상자 2019년 2차 '완료결과' 현황	68
<표 3-33> 지역별 선정 대상자 2019년 1차 '완료결과' 현황	70
<표 3-34> 지역별 선정 대상자 2019년 2차 '완료결과' 현황	71
<표 3-35> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	72
<표 3-36> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	73
<표 3-37> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	74
<표 3-38> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	75
<표 3-39> 선정 대상자 2019년 1차 '완료결과' 현황	77
<표 3-40> 선정 대상자 2019년 2차 '완료결과' 현황	79
<표 3-41> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	81
<표 3-42> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	82
<표 3-43> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	83
<표 3-44> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	84
<표 3-45> 선정 대상자 2019년 1차 '완료결과' 현황	85
<표 3-46> 선정 대상자 2019년 2차 '완료결과' 현황	87
<표 3-47> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	89
<표 3-48> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	90
<표 3-49> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	91
<표 3-50> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	92
<표 3-51> 선정 대상자 2019년 1차 '완료결과' 현황	93
<표 3-52> 선정 대상자 2019년 2차 '완료결과' 현황	95
<표 3-53> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	97

<표 3-54> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	98
<표 3-55> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	99
<표 3-56> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	100
<표 3-57> 선정 대상자 2019년 1차 '완료결과' 현황	101
<표 3-58> 선정 대상자 2019년 2차 '완료결과' 현황	103
<표 3-59> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황	105
<표 3-60> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황	106
<표 3-61> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황	107
<표 3-62> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황	108
<표 3-63> 선정 대상자 2019년 1차 '완료결과' 현황	109
<표 3-64> 선정 대상자 2019년 2차 '완료결과' 현황	111
<표 4-1> 지역별 훈련데이터 및 테스트데이터 현황	115
<표 4-2> 대상자 서비스 지원 내역 현황	116
<표 4-3> 변수 목록	117
<표 4-4> 기본 모형 및 지역별 모형 현황	118
<표 4-5> 기본 모형 대상자 분포도와 가중치(1인가구)	121
<표 4-6> 기본 모형 대상자 분포도와 가중치(다인가구)	124
<표 4-7> 서울특별시 모형 대상자 분포도와 가중치(1인가구)	128
<표 4-8> 서울특별시 모형 대상자 분포도와 가중치(다인가구)	132
<표 4-9> 경기도 모형 대상자 분포도와 가중치(1인가구)	135
<표 4-10> 경기도 모형 대상자 분포도와 가중치(다인가구)	138
<표 4-11> 전라북도 모형 대상자 분포도와 가중치(1인가구)	141
<표 4-12> 전라북도 모형 대상자 분포도와 가중치(다인가구)	145
<표 4-13> 제주특별자치도 모형 대상자 분포도와 가중치(1인가구)	149
<표 4-14> 제주특별자치도 모형 대상자 분포도와 가중치(다인가구)	152
<표 4-15> 기본 모형과 서울특별시 모형 분류행렬 및 AUC값 비교(1인가구)	156
<표 4-16> 기본 모형과 서울특별시 모형 Lift Chart값 비교(1인가구)	157
<표 4-17> 기본 모형과 서울특별시 모형 분류행렬 및 AUC값 비교(다인가구)	158
<표 4-18> 기본 모형과 서울특별시 모형 Lift Chart값 비교(다인가구)	159
<표 4-19> 기본 모형과 경기도 모형 분류행렬 및 AUC값 비교(1인가구)	160

<표 4-20> 기본 모형과 경기도 모형 Lift Chart값 비교(1인가구)	161
<표 4-21> 기본 모형과 경기도 모형 분류행렬 및 AUC값 비교(다인가구)	162
<표 4-22> 기본 모형과 경기도 모형 Lift Chart값 비교(다인가구)	163
<표 4-23> 기본 모형과 전라북도 모형 분류행렬 및 AUC값 비교(1인가구)	164
<표 4-24> 기본 모형과 전라북도 모형 Lift Chart값 비교(1인가구)	165
<표 4-25> 기본 모형과 전라북도 모형 분류행렬 및 AUC값 비교(다인가구)	166
<표 4-26> 기본 모형과 전라북도 모형 Lift Chart값 비교(다인가구)	167
<표 4-27> 기본 모형과 제주특별자치도 모형 분류행렬 및 AUC값 비교(1인가구)	168
<표 4-28> 기본 모형과 제주특별자치도 모형 Lift Chart값 비교(1인가구)	169
<표 4-29> 기본 모형과 제주특별자치도 모형 분류행렬 및 AUC값 비교(다인가구)	170
<표 4-30> 기본 모형과 제주특별자치도 모형 Lift Chart값 비교(다인가구)	171
<표 4-31> 지역파생 모형(지역변수)	172
<표 4-32> 지역파생 모형 대상자 분포도와 가중치(1인가구)	173
<표 4-33> 지역파생 모형 대상자 분포도와 가중치(다인가구)	177
<표 4-34> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(서울시 1인가구)	181
<표 4-35> 기본 모형과 지역파생 모형 Lift Chart값 비교(서울시 1인가구)	182
<표 4-36> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교 (서울시 다인가구)	183
<표 4-37> 서울특별시 모형과 지역파생 모형 Lift Chart값 비교 (다인가구)	184
<표 4-38> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(경기도 1인가구)	185
<표 4-39> 기본 모형과 지역파생 모형 Lift Chart값 비교 (경기도 1인가구)	186
<표 4-40> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(경기도 다인가구)	187
<표 4-41> 경기도 모형과 지역파생 모형 Lift Chart값 비교 (다인가구)	188
<표 4-42> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(전북 1인가구)	189
<표 4-43> 기본 모형과 지역파생 모형 Lift Chart값 비교 (전북 1인가구)	190
<표 4-44> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(전북 다인가구)	191
<표 4-45> 기본 모형과 지역파생 모형 Lift Chart값 비교 (전북 다인가구)	192
<표 4-46> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(제주 1인가구)	193
<표 4-47> 기본 모형과 지역파생 모형 Lift Chart값 비교 (제주 1인가구)	194
<표 4-48> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(제주 다인가구)	195
<표 4-49> 기본 모형과 지역파생 모형 Lift Chart값 비교 (제주 다인가구)	196

그림 목 차

[그림 1-1] 연구흐름도	11
[그림 2-1] 빅데이터 분석 단계에서 머신러닝 (허진경, 2018)	21

요 약

- 본 연구에서는 지역별 예측모델의 효용성을 살펴보기 위해 국내외 사례 선행 연구 조사부터, 데이터 탐색과 예측모델 개발 및 평가까지 수행하였다.
- 국내외 사례 연구 파트에서 국내 사례의 경우 복지사각지대에 대한 개념이 과거에는 단순히 저소득층을 기준으로 복지사각지대 대상자에 초점이 맞춰져 복지사각지대 발굴 수행을 하는 사례에 대해 검토를 하였다.
- 머신러닝 알고리즘에 대한 부분에서는 본 연구에서 머신러닝 기법으로 활용하는 XGBoost에 대한 세부적인 설명과 모형 평가를 위한 분류행렬, ROC커브, AUC, 리프트 도표에 대한 설명에 대하여 살펴보았다.
- 3장 데이터 탐색에서는 복지사각지대 19년 1차와 2차 데이터를 기반으로 전체 대상자의 속성분석, 단독변수와 가구 수에 따른 속성분석, 지역별 대상자 속성분석까지 진행하였다.
- 예측 모형 구현 및 분석에서는 기본모형, 지역별 모형 그리고 파생 변수를 투입했을 때의 효과를 살펴보았다. 서울특별시 모형 기준으로 1인 가구의 경우 기본 모형이 더욱 더 좋으며, 다

인 가구의 경우 서울특별시 모형이 더 좋았다. 경기도 모형도 마찬가지로 1인 가구의 경우 기본 모형이 더욱 좋았으며, 다인 가구는 경기도 모형이 더 우수한 결과를 보여주었다. 전라북도 와 제주특별자치도의 경우 1인 가구 다인 가구 모두 기본 모형 이 더 우수한 결과를 보여주었다.

- 이와 같은 다양한 분석을 기반으로 본 연구에서는 다음과 같은 정책적 제언을 하였다.
- 첫째, 현재 정확도에서 더욱 더 정교한 정확도를 가지기 위해서는 데이터 입수기관과 수집되는 데이터에 대한 정합성, 수집되는 데이터가 분석대상자로 구성되는 세부적인 프로세스, 분석대상자에서 예측모형의 입력 데이터로 전환하기 위한 전처리 과정 및 최종 발굴 대상자가 되기까지에 존재하는 각 필터링 등을 재정비 하여 더욱 더 정확도를 재고해야 될 것이다.
- 둘째, 인구 수에 따라 지역을 분리하고 분리된 지역이 특정 인구 수 이상이 될 경우, 지역 모형을 도입하고, 그렇지 않을 경우 지역 파생 변수 또는 지역 융합 데이터 기반의 발굴 등으로 복지사각지대 발굴시스템 운영을 수행한다면 적중률 향상에 도움일 될 것이다.
- 셋째, 연령이 높은 노년기 층에 속하는 대상자의 경우 별도의 모형 또는 별도의 복지서비스 전달을 하기 위한 방안을 마련한다면 더욱 더 노년기층에 속하는 대상자에게 복지 서비스를 많이 주면서 복지사각지대 관리시스템에서는 타 연령층에 많은 대상자에게 복지 서비스 기회를 줄 수 있을 것이다.
- 넷째, 지자체 공무원이 처리되는 실적까지 고려된 새로운 모형

평가 방법이 추가적으로 고려되어 가상 환경에서 실험한 결과가 지자체 공무원이 처리 된 실적까지 예측이 될 수 있는 지표가 만들어져야 정확한 실험이 가능할 것으로 보인다. 예를 들어, 분류 행렬의 민감도 + 정확도 + 지자체 공무원 수 + 처리량 대비 정확도 등을 모형 평가에 넣는 것이다.

- 다섯째, 복지사각지대 발굴시스템의 정확도 제고 및 서비스 범위를 넓히기 위해서는 비정형 데이터의 확대를 수행해야 될 것이다. 특히 비정형 데이터 중 복지와 보건 사례관리용 비정형 데이터는 유용하게 활용될 것으로 판단된다.
- 여섯째, 머신러닝 알고리즘의 지도 학습을 보완하기 위한 Weakly-supervised learning 기법 활용은 인공지능을 더욱 더 강화하는데 도움을 줄 것이다.
- 마지막, 복지사각지대 발굴관리시스템은 행정데이터와 최신 머신러닝 기법을 활용한 대표적인 시스템으로써 사회보장정보원의 다양한 시스템과의 접목이 가능하다.
- 이상 7가지 정책적인 제언은 사회보장정보원의 복지사각지대 시스템의 고도화 및 차세대 사회보장정보시스템으로 발전하는데 큰 도움이 될 것으로 판단된다. 특히 본 연구는 차세대 사회보장정보시스템에서 예측모형의 다각화를 위한 기초 자료로 활용될 수 있으며, 연령별 모형, 가구유형 모형 등을 만들 때 참고자료로 활용될 수 있을 것으로 판단된다.

서 론

제1절 연구의 배경 및 목적

제2절 연구방법 및 내용

제1장 서론

제1절 연구의 배경 및 목적

과학기술의 발달은 우리 사회를 정보화 사회로 빠르게 만들었으며, 모든 사람과 사물을 인터넷 네트워크 환경에 연결되어 서로 유기적으로 정보를 주고받을 수 있는 환경으로 만들었다. 이런 환경에서 매일 다양한 종류의 미디어 정보 및 데이터가 누적되고 있고 빅데이터의 활용도는 높아지고 있다.

복지 사각지대의 개념에 대해서는 다양하게 논의되고 있으며 아직까지도 학자들마다 의견이 분분하여 정확한 대상자에 대한 정의와 범위에 대해서 서로 다른 의견을 가지고 있다. 사회보장급여법에 쓰여 있는 사각지대에 대한 정의는 사회보장에 대한 사각지대 혹은 사회보장급여에 대한 사각지대로 말하고 있는데 이것은 공공부조와 사회보험과 사회서비스로 확대된 개념으로 볼 수 있을 것 같다. 더 확장되어 사각지대에 대한 개념을 구인회·백학영(2008)는 정의하기 위해 사회보장 제도 적용에서 사회보장급여에서 제외된 집단으로 정의를 한다. 공공부조 영역에서의 복지사각지대에 대한 논의는 주로 기초생활보장제도를 중심으로 진행되어 왔다. 김은하 외(2015)는 복지사각지대 범위에 대한 것을 자격기준에 적합하지 않아 비수급 상태인 경우와 그렇지 않은 경우로 분리하였으며 수급/저격성 여부에 대해서도 고려하였다. 최근 망우동 사건등 주 소득을 상실하여 집안의 생계가 어려

운 가정에게 지원해주는 긴급복지지원법에 관심이 높아지고 있는 상황이다. 그래서 복지에 대한 욕구가 존재함에도 생활의 어려움 때문에 소외 계층으로 보고 있고, 자격기준에 배제된 대상자 등 발굴하지 못한 대상자도 제시되어지고 있다. (보건복지부, 2017) 최근에는 복지 사각지대에 대한 개념이 해외에서 Non-take-up 또는 Blind Spot로써 정의 되고 있다. 복지사각지대 (Non-take-up, Blind Spot)은 자신에게 자격이 있어 복지 서비스를 받을 수 있음에도 불구하고 복지 서비스를 받아 이익을 얻고자 하는 주장을 하지 않는 가구를 말하며, 경제이론에 비추어보면 비합리적으로 행동하는 사람을 의미한다 (Harnisch, 2019). 국내에서도 복지사각지대에 대한 대상자를 단순히 경제적으로 어려운 사각지대 대상자 말고도 빈곤을 넘어 고립, 관계 단절, 정신적 인지적인 문제가 있는 경우까지도 사각지대 대상으로 확대하여 보고 있다 (보건복지부, 2019).

사회보장정보원에서는 복지사각지대 대상자를 능동적으로 발굴하기 위해 머신러닝 기술을 활용하여 2015년 12월부터 복지사각지대 발굴 관리시스템을 개통하여 2개월에 한 번씩 정기적으로 운영해오고 있다. 복지사각지대 발굴관리시스템은 입수자료의 연계기관의 협의부터 협의 후 데이터를 받아오기 위한 컬럼의 정의와 범위 및 수집주기와 개인정보 동의 등 복잡한 과정 및 최종적으로 법적 근거로 인한 데이터 수집 등 복잡한 과정을 거친 데이터 기반의 시스템이다. 매년 복지사각지대 발굴관리시스템은 단순한 설문지를 통해 수집하는 방식에서부터 한 개의 의사결정나무 또는 로지스틱 기법과 같은 단순한 통계기법 기반으로 발전하였으며, 이후에 로지스틱, 엘라스틱, 그라디언트 부스팅 모델의 융합 등 다양한 기법으로 발전되고 있는 상황이다. 이와 같은 데이터의 연계 기관의 확대 그리고 머신러닝 알고리즘의

발전과 더불어 현재는 가구 수에 따른 군집화 등으로 다양화 되어가고 있는 상황이다.

이에 따라 복지사각지대에 대한 발굴은 현재 다양화에 맞춰 발전되고 있는 상황에서 본 연구에서는 복지사각지대의 지역별 군집 수행을 통해 전체 지역 데이터를 서울, 경기, 전북, 제주로 군집화 하여 영향도를 비교 분석하고자 한다.

전체 지역을 군집화 하여 분리하게 되면 기존 복지 사각지대 발굴 관리 시스템에서 수행하는 예측 모델의 대표 값이 지역 특성이 존재하는 예측 모델이 된다. 예를 들어 전체 지역을 서울과 경기도 지역으로 분리하여 예측 모델을 통한 예측 수행을 하고자 할 때, 서울이라고 하는 지역이 예측 분류의 제일 상단에 위치하게 되어 분리가 되게 된다. 즉, 서울이라고 하는 큰 바구니에 대상자를 묶은 후 예측을 수행한다는 이야기가 된다. 전체 지역이 동일한 특성을 가지고 있고 같은 수의 수급자를 가지고 있으면 의미가 없는 작업이지만 만약 지역마다 특성을 가지고 있고, 각 지역마다 특성이 다르다면 지역별 군집을 통해 예측 모델 분리 및 지역 변수의 범주화를 통한 예측 변수 중 예측비율이 높은 변수를 추가하는 방법을 사용한다면 예측력을 증강시킬 수 있다.

본 연구진은 이전 연구에서 머신러닝 알고리즘에 초점을 맞춰 연구를 진행하여 복지사각지대의 개념정립부터 아키텍처의 구성 등을 세부적으로 살펴보았다.

본 연구의 목적은 복지사각지대 지역별 예측모형 평가를 위해 총 5개의 장으로 구성되어 있다. 1장에서는 지역별 예측모형 연구를 위한 연구 배경 및 목적에 대해 설명한다. 2장에서는 복지사각지대 시스템

설명과 간단한 모형 평가 방법에 대해 살펴본다. 3장에서는 지역별 예측모형에 활용되는 데이터에 대한 특징을 살펴보기 위한 데이터 탐색을 수행한다. 4장에서는 지역별 예측모형 구현 및 분석을 수행하기 위해 지역별 예측모형을 구현하기 위한 설계부터 모형으로부터 나온 결과 분석까지 수행한다. 마지막 5장에서는 연구의 요약을 하며 예측 모형 개선을 위한 방안 및 정책제언을 하고자 한다.

제2절 연구방법 및 내용

본 연구의 연구방법은 다음과 같이 나뉘 볼 수 있다.

첫째, 본 연구는 지역별 예측모형 연구를 수행하기 위한 실증 데이터 기반의 연구로써 우선적으로 데이터 탐색을 수행한다. 데이터 탐색을 수행하기 위해 본 연구진은 사회보장정보원에서 수행하는 복지사각지대발굴시스템에서 추출되는 피드백데이터와 2달 또는 1달에 한 번씩 입수되는 외부연계 데이터와의 융합을 수행하여 대상자의 개별 속성 데이터를 기반으로 연구를 위한 데이터를 마련한다. 이렇게 마련된 데이터를 기반으로 본 연구에서는 19년도 1차와 2차 데이터로 분리하여 변수 개수별, 연령별, 지역별, 변수속성별로 기술통계 기반의 데이터 탐색을 수행한다. 변수의 세부 속성은 외부 연계기관으로부터 입수되는 속성변수로서 작년 연구진이 수행한 연구 중 사회보장정보원의 자체과제인 “복지사각지대 대상자 발굴을 향상을 위한 인공지능 시스템 활용 연구”을 참고하여 각 속성변수의 세부 내용을 파악한다.

둘째, 복지사각지대 지역별 예측 모형 연구의 경우 데이터 입수부터 예측모형 구동과 현장조사까지 일련의 과정이 서로 연결되어 있어 각 단계별 고려가 있어야 정확한 예측 모형이 개발 될 수 있다. 이를 위해 본 연구 책임자의 경우 복지사각지대 예측모형 구동 및 개발 경험이 있기 때문에 보건복지부와의 주기적인 회의와 복지사각지대 발굴관리시스템을 구동하기 위한 사회보장정보원의 수급자관리부, 희망복지중앙지원단과의 주기적인 정보 공유 및 인터뷰를 통하여 지역별 예측 모형의 세부적인 데이터 파악 및 전체적인 프로세스에 대한 이

해를 수행하여 지역별 예측 모형 개발을 하는데 있어 실무 담당자가 고려된 예측 모형을 개발한다. 본 연구 과제에서 법, 제도, 공무원의 전달체계 및 데이터 정합성과 최신 머신러닝 알고리즘 등 모든 것을 고려하기에는 범위가 넓기 때문에, 사전에 보건복지부와 협의 하였던 지역별 예측모형을 개발하는데 있어 기존에 사회보장정보원에서 활용하는 변수를 사용하며, 예측 모형 기법의 경우 XGboost 기법으로 한정한다. 또한 예측 모형을 시군구로 분리 할 때 수급자 비율이 가장 높은 서울특별시와 경기도 그리고 중간 비율을 가지고 있는 전라북도와 낮은 비율에 속하는 제주특별자치도를 고려하여 4개의 지역별 예측 모형을 기반으로 분석을 수행한다. 다음은 본 연구를 수행하기 위해 DB 접근 및 데이터 추출과 연구 타당성 및 과제 심의와 관련된 이해 관계자를 보여준다. 연구센터의 경우 복지사각지대 지역별 예측 모형을 연구하는 연구 수행자 역할을 수행하며, 통계센터는 차수별 복지사각지대 발굴 실적 관련 데이터에 대한 정형 데이터 추출 역할을 수행한다. 그리고 수급자관리부의 경우 복지사각지대 외부 연계기관으로 수집되는 데이터를 관리하고 분석데이터를 구성하는 역할을 수행하여 실제 예측 모형에서 활용하는 원천 데이터를 구성해 주는 역할을 한다. 마지막으로 보건복지부 복지정보기획과의 경우 본 연구를 수행하는데 있어 정책적인 결정이나 연구 수행하는데 필요한 협의를 맡게 된다. 본 연구진은 이와 같은 이해관계자와의 회의를 통해 본 연구를 협의하면서 진행하며 최종 결과 공유를 통해 본 연구 결과가 실제 사회보장정보원의 복지사각지대발굴시스템에 적용되어 정책결정까지 활용될 수 있게 한다.

〈표 1-1〉 협의기관 역할 및 연관 데이터

협의 기관		역할	연관 데이터
사회보장 정보원	연구센터	복지사각지대 연구 수행	
	통계센터	복지사각지대 정형 통계 협의	사각지대 실적 통계 DW
	수급자관리부	복지사각지대 운영 현황	수집 데이터 분석 데이터
	희망복지중앙 지원단	지자체 협의 및 환류 데이터 결과	지자체 협의 복지 서비스
보건복지부 복지정보기획과		정책 결정 연구 수행 협의	

셋째, 본 연구에서 개발한 지역별 예측 모형 개발을 하기 위해서 기존 사회보장정보원 복지사각지대 발굴시스템에서 활용하는 예측 모형 변수와 예측 모형 기법을 참조한다. 본 연구에서 활용하는 예측 모형 변수의 경우 기존 보건복지부 수탁과제로 수행된 “사회보장정보 시스템을 활용한 복지 사각지대 발굴방안 연구”와 후속확대 연구를 참고하여 현재 사회보장정보원에서 운영 중인 예측모형에 변수를 활용한다. 따라서 본 연구의 경우 예측 모형에서 활용하는 변수는 고정되어 있다는 가정을 가지고 연구를 수행한다. 또한 머신러닝 기법으로는 분류에 우수한 성능을 가지는 부스팅 기법을 활용하는 XGboost 기법을 사용하는 것으로 한정한다. 최신 머신러닝 기법에 대한 비교는 본 연구 수행 전 사회보장정보원 연구센터 자체연구 중 “복지사각지대 대상자 발굴률 향상을 위한 인공지능 시스템 활용 연구”에서 수행한 내용으로써 본 연구에서는 지역별 예측 모형을 구성하였을 때의 영향도 및 효과에 초점을 맞춰 연구를 수행한다. 본 연구의 가설은

예측 모형을 활용하였을 때 지역별 예측 모형이 효과가 있는지 그렇지 않은지에 대한 것으로 실제 실 데이터의 실험을 통해 가설 검증을 수행한다.

마지막으로 본 연구의 결론 부분에서는 본 연구에서 수행한 데이터 탐색과 예측 모형 간 비교를 통해 나온 결과를 기반으로 복지사각지대 발굴시스템으로부터 더 높은 정확도를 가지고 많은 대상자에게 복지 서비스를 전달하기 위한 정책적인 제언을 수행하고자 한다. 본 연구에서 수행하는 정책적인 제언의 경우 시스템 기반으로 행정데이터의 활용도를 더 높여 더 많은 대상자에게 복지 서비스를 전달시키기 위한 기초 자료로 활용될 수 있을 것으로 예상된다.

연구 내용은 다음과 같이 구성될 수 있다.

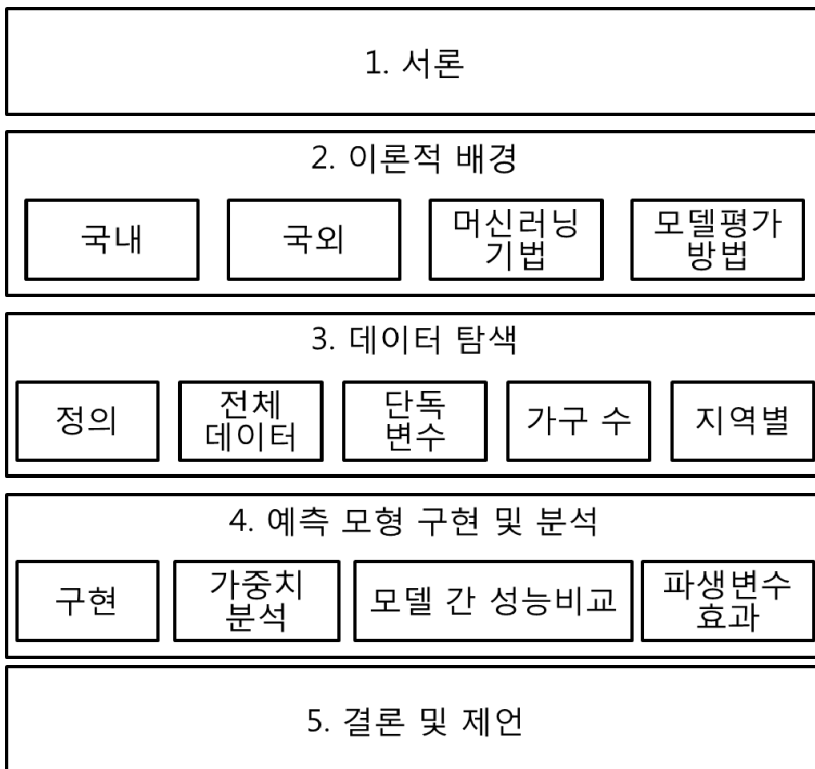
이론적 배경부분에서는 국내외 선행 연구에 대해서 살펴본다. 국내 연구는 주로 행복e음과 보건복지부 수탁과제로 수행된 복지사각지대 발굴 관련된 시스템 연구를 주로 살펴보았으며, 국외 연구의 경우 오스트리아, 독일, 스위스, 프랑스에서는 복지사각지대 발굴에 대하여 살펴보았다. 그리고 머신러닝에 대한 전반적인 개념과 본 연구에서 활용하는 XGBoost 기법에 대한 세부 설명을 하고자 한다.

데이터 탐색 부분에서는 전체 분석데이터 대상자 속성부터, 단독변수를 보유한 대상자, 가구 수에 따라서 달라지는 대상자의 분포와 지역별 대상자 속성 분포에 대하여 살펴본다.

다음으로 예측 모형 구현 및 분석에서는 서울특별시, 경기도, 전라북도, 제주특별자치도 모형을 구현하여 구현된 모형의 가중치 값을 통해 변수간의 영향을 분석한다. 그리고 모형 성능 평가 방법을 활용하여 모형 간 성능분석을 수행하고자 한다. 또한 군집으로 분리된 지

역별 모형과 지역 파생 변수를 활용한 모형 간의 효과를 분석하고자 한다.

마지막 결론에서는 본 연구를 요약하고 본 연구를 통하여 사회보장 정보원 복지사각지대발굴시스템 개선 및 정책 제언을 수행하고자 한다.



[그림 1-1] 연구흐름도

이론적 배경

제1절 국내외 선행연구

제2절 머신러닝 기법

제3절 모델평가 방법

제2장 이론적 배경

제1절 국내외 선행연구

본 절에서는 지역별 예측 모델 선행 연구를 살펴보기 위하여 복지 사각지대 발굴시스템의 국내 발굴 사례와 국외 발굴 사례를 살펴보고, 지역 정보를 머신러닝 기법을 활용한 사례를 살펴본다.

1. 복지사각지대 국내 발굴 사례

복지사각지대의 체계적 발굴을 위해 관련 기관에서 보유한 정보를 연계하거나 정보시스템 활용의 필요성을 제시하는 연구가 수행되었다(최균 외, 2012). 2014년 발표된 정부의 복지사각지대 발굴 및 지원 종합대책안(보건복지부, 2014)은 사각지대에 대한 상시 발굴체계 구축의 일환으로 각 부처 및 공공기관 등이 보유하고 있는 잠재위기가구 정보 연계 및 DB 구축을 통해 위기가구를 주기적으로 관리하는 내용을 포함하고 있다. 김은하·추병주(2014)는 사회보장정보시스템을 이용해 복지사각지대 대상자 발굴을 위해서 데이터베이스 구축하기 전 단계에서 추후 활용도 높은 데이터 탐색을 하였으며 데이터베이스 설계 방안 제시 등을 하였다. 사회보장정보시스템 상에서 수집 가능한 정보를 파악하여 사례관리의 욕구영역을 기반으로 데이터 관리 항목을 제시하고 있다. 김은하 외(2015)는 복지사각지대 발굴과 관련된 정보 연계 및 활용, 정보시스템 구축에 필요한 법·제도 개선방안을 검토하

여 복지사각지대 발굴 및 지원을 위한 정보시스템 구축방향을 제시하였다. 이 때 개인/가구/일자리/경제/심리사회적 특성 등 분류해서 정보를 파악하였으며, 이것을 바탕으로 하여 통계방법론을 활용해 복지사각지대 대상자 예측을 수행하는 모형 개발을 하였다. 김은하 외(2016)는 신규 연계정보의 도출, 지자체 사각지대 발굴업무의 효율화 방안을 후속연구에서 제시하였다. 복지사각지대 발굴시스템으로부터 최종 발굴대상자로 선정된 대상자의 인구사회학적인 연령, 지역, 변수 등에 대해 추가 분석 실시하였으며, 기존 모형 대비 수정 모형이 어느정도 성능이 개선되었는지를 비교분석 하였다. 더군다나 복지사각지대 대상자 발굴을 위해 신규 연계정보를 파악하여 13개의 연계정보 후보군을 도출하였다. 최현수 외(2016)는 1인가구의 위험요인 분석을 위해서 다양한 마이크로 데이터를 활용해 빅데이터의 2분 8척하였다. 그리고 1인가구 예측 모형과 정책 지원 방안에 대해서도 제안하였다. 최현수 외(2018)은 사회보장정보시스템을 활용한 복지사각지대 발굴 방안 확대 연구를 통하여 2019년 보건복지부는 “포용적 사회보장 정보 기반 마련을 위해 사회보장정보시스템 새로 구축한다!”라는 타이틀을 가지고 보도자료를 발표하였다. 보도 자료에는 복지사각지대 발굴 시스템 2.0 도입을 할 계획을 제시하였다. 과거 단순 단전·단수 정보, 건강보험료 체납정보 기반으로 경제적 어려움에 처한 사람을 위주로 사각지대를 발굴하였다면, 복지사각지대 발굴시스템 2.0에서는 위기 가구에 대한 정의를 다양화 하여 빈곤을 넘어 고립, 관계단절, 정신적·인지적인 문제가 있는 케이스까지 포용할 수 있는 연구를 수행하여 사회적 논의를 진행할 것을 계획하고 있다 (보건복지부, 2019). 이와 같은 연구 흐름속에 현재 보건복지부는 차세대 사회보장정보시스템에서 추구하는 위기 가구 이외 가구를 발굴하기 위한 일환으로 복지사각지대 발굴시스템의 다각화를 모색하고 있다. 본 연구는 차세대

복지사각지대 2.0 도입 전 복지사각지대 발굴시스템의 다각화의 한 축인 지역 단위 예측모형을 분리하고자 하는 첫 단추인 연구라고 할 수 있다.

2. 가난, 빈곤을 인공지능 기술을 활용한 국외 연구 사례

국내 복지사각지대와 직접적으로 연관이 높은 연구는 많이 없기 때문에 본장에서는 가난, 빈곤 등 복지사각지대에 관련된 다양한 사례를 인공지능 기술을 활용한 국외 연구 사례를 살펴보고자 한다.

개발도상국의 경제 생계에 대한 데이터는 많이 희박하기 때문에 Neal Jean 연구에서는 위성 사진을 활용하여 지역 수준의 경제 변동을 설명하는 것을 인공지능을 활용하였다(Neal Jean et. al., 2019). 이때 나이지리아, 탄자니아, 우간다, 말라위, 르완다 등 5개의 아프리카 국가의 설문조사 및 75%이상을 설명할 수 있는 고화질 이미지를 CNN 신경망 기술을 활용하여 빈곤 추적을 하고자 하였다. 결과적으로 연구에서는 제한된 데이터를 위성 이미지와 설문조사 등을 통해서 빈곤을 어느정도 추적이 가능하다는 가능성을 보여주었다. 그 밖의 Linden McBride의 연구에서는 (Linden McBride and Austin Nichols, 2015) 기존에 빈곤에 대한 테스트를 수행하는 Proxy means test (PMT)에 대한 빈곤 타겟팅 도구를 개선하기 위한 연구를 수행하였다. 이 때 기존에 PMT의 경우 내부에 있는 샘플 데이터를 고려하기 때문에 내부 데이터에 대해서는 많은 에러가 존재하지 않는다는 장점이 있다. 하지만 실제 빈곤 타겟팅에 대한 목표는 외부에 대한 샘플 예측 에러에 대해 줄이는 것이기 때문에 연구에서는 외부 데이터에 샘플 예측 에러를 줄이기 위해 머신러닝 기법 적용을 수행하였다. 이 때 연구진은 최신 머신러닝 알고리즘인 Quantile 회귀 포레스트라는 머신러닝 기법을 적용하였으며, 2~18% 정도 정확도를

개선할 수 있음을 보여주었다. 기존 위성을 활용하여 빈곤 지역에 대한 연구와 비슷하게 (Anthony Perez. et. al. 2019) 연구에서는 지역 빈곤 지표를 설명하는 자료를 수집하기 위해 Landsat 7에서 제공하는 아프리카 대륙의 무료 및 공개적으로 이용가능한 주간 위성 이미지에 대해서 CNN기법을 활용하여 지역 경제 생계를 예측하기 위한 모델 구축을 수행하였다. 위성 이미지의 5%만이 라벨과 연관될 수 있기 때문에 GAN을 활용해 semi-supervised 접근법을 활용하는 방법이 아닌 경사 패널티 기법을 활용하는 Wasserstein GAN 기법을 활용해 안정적인 훈련을 수행하는 연구를 하였다. 제안한 모델의 경우 높은 상관관계를 가지고 예측을 안정적으로 수행하지만 “oracle”에 의해 달성된 0.66 상관 계수를 극복하는데 한계가 존재하기 때문에 WGAN을 통해서 극복하기 어렵다는 것을 입증하였다. 이와 같은 다양한 머신러닝을 활용한 빈곤 예측에 대한 연구이외에도 복지사각지대 시스템을 간접적으로 독일, 스위스, 프랑스에서 사회보장혜택과 빈곤 퇴치에 목적을 두고 변수 등 연구를 수행한 것은 존재한다. 특히 빈곤격차 (Fuchs et al., 2019). (Heckman, 1976). (Hamisch, 2019). 빈곤격차, 한부모 가정, 연령이 많은 가구주 등이 정(+)에 영향을 미치는 것을 통해 본 연구에서 활용하는 변수와 영향이 비슷하다는 것을 알 수 있었다. 그 밖에도 스위스와 유럽에서 농촌지역보다는 도시지역이 사각지대 대상이 낮게 나온다는 연구와 장기요양보호를 받는 노인이 사각지대 대상으로 중요하다는 연구도 존재한다 (Hümbelin, 2019, Crettaz 2009) (Arrighi et al., 2015).

3. 지역정보를 활용한 머신러닝 연구 사례

본 절에서는 지역정보와 머신러닝을 활용한 연구 사례를 살펴보고

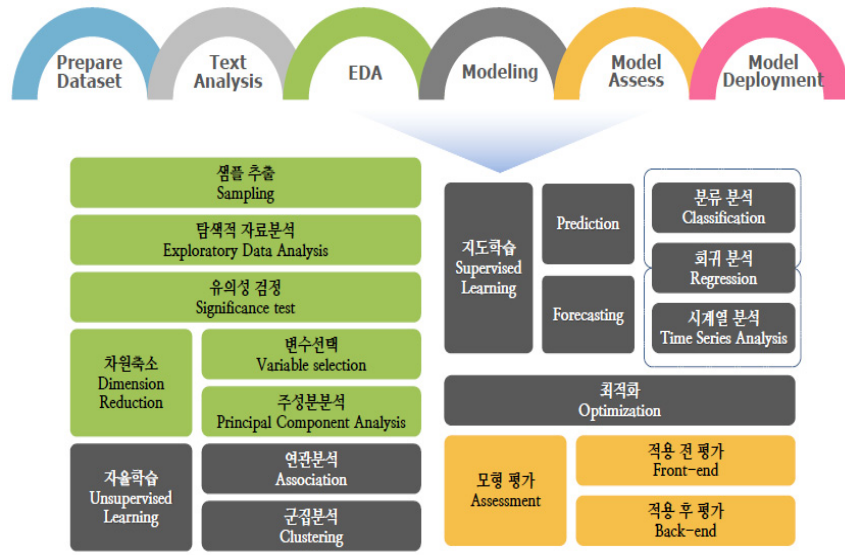
자 한다. 박병호 (1995)는 교통사고의 심각성에 대한 인식 증가로 교통사고를 예측하기 위한 지역별 모형을 제안하고 특성을 비교 분석하였다. 우선 교통사고 모형의 형태는 지역 특성에 따라 각기 다양하고 설명변수도 경제, 사회, 문화, 제도 및 기술적 요인 등에 따라 다양하기 때문에 이를 위해 연구에서는 PTM과 지수함수의 결합모형을 사용하였다. 지역별 인구, 자동차보유대수, 사망자 수 등을 변수로 활용하여 지역별 예측모형을 활용하여 교통안전 정책을 수립하는 근거자료로 제시하였다. 허선영 외(2018)은 머신러닝 기법을 활용하여 범죄발생 위험 지역을 예측 하였다. 이를 위해 연구진은 빅데이터 구축이 가능한 범죄정보와 도시지역 자료를 바탕으로 한국형 범죄발생 위험 지역 예측 모형을 개발하였다. 이 때 범죄발생 확률을 지도로 시각화함으로써 사용자 이해도록 높이도록 하였다. 실제 실험 결과 J 시의 시나리오 별로 위험지역이 다르게 나타나 시민들과 경찰관계자에게 도움이 될 수 있는 자료로 제공할 수 있게 되었다. 박일수 외(2012)은 질환성 심정지 발생의 지역별 변이에 관한 연구를 수행하기 위해 행정 구역별 건강상태와 심정지발생에 미치는 지표를 분석용 데이터 셋에 구축하였으며, 심정지 조사자료와 지역사회 건강조사자료를 지표선정을 위해 활용하였다. 그 밖에도 머신러닝을 활용한 범죄발생 위험 지역 예측 연구 (허선영, 2018), 강원도 지역의 산불 발생 예측모형 연구 (채경재, 2018), 수도권 지역의 호우피해 예측함수 연구 (최창현, 2018), 상주지역 산사태 취약성 분석 (강경희, 2018), 광역시 지역주민의 원전계속운전 수용성 결정요인 분석 (박호정, 2018), 머신러닝을 활용한 가뭄 취약 지역 분석 등의 다양한 지역관련 머신러닝 연구 (신정훈, 2018)가 존재한다. 하지만 공공데이터를 활용하여 머신러닝 기법과 지역이라는 것을 고려하여 사회 취약계층을 발굴하는 연구는 본 연구가 공공데이터를 활용한 큰 시도라고 볼 수 있다.

제2절 머신러닝 알고리즘과 XGBoost

본 절에서는 모형들에 대한 일반적인 개념, 과정, 장/단점 등에 대한 설명을 하고, 4장에서 본 절에서 설명한 모형에 데이터를 접목한 모형 구동 사례를 가지고 구체적인 설명을 하기로 한다.

1. 머신러닝 알고리즘 개념

머신러닝 알고리즘이란 현재 상황을 인지해서 종합적으로 분석하여 미래를 예측하는데 활용된다. 이를 위해 수집, 정제, 학습, 임무, 성능측정, 구축 순으로 일반적으로 처리되고 있다. 머신러닝 알고리즘은 누적되는 데이터의 양이 늘어날수록 학습능력이 뛰어나 성능 향상을 가져올 수 있다. (허진경, 2018) 머신러닝은 전통적인 통계 기법을 사용하는 방법과 기계학습을 통해 데이터를 학습하여 예측 모형을 만드는 방식이 존재한다. 전통적인 통계적인 기법의 경우 로지스틱, 의사결정나무, K-NN 기법이 대표적으로 존재하며, 기계학습 모형의 경우에는 신경망, SVM, XGBoost 등이 대표적으로 존재한다. 일반적으로 전통적 통계기법에 비해서 최신 머신러닝 알고리즘 성능이 더욱 더 좋다. 하지만 최신 머신러닝 알고리즘은 전통적인 통계적 기법에 비해 설명력이 떨어진다는 단점을 지니고 있다.



[그림 2-1] 빅데이터 분석 단계에서 머신러닝 (허진경, 2018)

전처리 단계에서는 수집되는 데이터에 대해서 예러가 많은 데이터의 경우 예러를 최대한 줄이고 예측 모형에 활용할 수 있는 형태로 변형 시키는 작업을 수행한다. 때에 따라서 전처리 단계에서는 결측값을 채우기 위해 결측치 채우기 위한 알고리즘 기법을 활용하기도 한다. 이후 변수 선택을 위한 단계에서는 다중공선성이 일어나지 않는 변수 그리고 모형에서 종속변수를 가장 잘 설명할 수 있는 독립변수를 선택하는 작업을 수행한다. 최종적으로 전처리와 변수 선택단계에서 선택된 변수를 활용하여 예측 모형을 훈련시키는 작업을 하게 되는데, 이 때 다양한 예측 모형을 모형 성능평가 방법을 통해 비교 분석하여 가장 성능이 좋은 예측 모형을 선택한다. 그리고 최종 선택된 모형은 활용되지 않은 데이터 예측에 쓰여 데이터 예측을 위해서 사용된다.

2. XGBoost (eXtreme Gradient Boosting)

XGBoost (Tianqi Chen et. al. 2016)는 앙상블 기법 중 하나인 그레디언트 부스팅 기법을 기반으로 수행되는 앙상블 모형으로써, 과거 결정 트리와 같은 단일 트리 모형에 비해 더욱 더 정확하면서도 병렬적으로 처리하기 때문에 빠른 속도로 문제를 해결하기 적합한 모형이라고 할 수 있다. Gradient 부스팅을 바탕으로 동작되는 XGboost 앙상블 모형은 다수의 트리 구조 모형을 동시에 만드는 방식이 아닌 순차적으로 진행되어 만들어지는 방식으로 구동된다. 순차적으로 진행되어 만들어지는 트리구조의 모형은 차수가 지날수록 에러률이 줄어드는 효과를 가져 온다. CART (Classification And Regression Tree) 기반의 앙상블 모형을 활용하는 XGBoost의 경우 약한 분류기 모형을 활용하여 최종적으로 점수를 계산하여 모형을 만드는 방식으로 진행된다. 이런 이유로 CART 모형 기반의 모형이 만들어지면 모형 간에 스코어를 비교하면서 최종 만들어진 모형이 적합한지 아닌지에 대해 판단할 수 있게 된다. 이 때 생성되는 다양한 분류기를 약 분류기라고 하여 (Weak Classifier)라고 부른다. 그리고 최종 생성된 분류기에 데이터를 넣게 되면 데이터는 양의 방향 또는 음의 방향으로 보여준다. 만약 만들어진 모형이 양의 방향과 음의 방향을 잘 구분할 수 있으면 좋은 모형이다.

제3절 모형 평가 방법

본 절에서는 본 연구에서 제시된 모형의 성능을 평가하기 위해서 데이터마이닝에서 활용되는 모형 평가 방법들을 살펴본다. 분류결과의 성능을 평가하는 데에는 분류행렬 (S.V. Stehman, 1997), ROC 곡선 (T. Fawcett, 2006), 리프트 도표 (S. Tuffery, 2011) 등이 활용되며 관련 지표로 정확도, 오분류율, 민감도, 특이도 등이 사용된다.

혼돈행렬이라고 하는 것은 분류를 수행하는 예측 모형을 평가하는 도구 중에 하나로 실제 값과 예측 값을 표를 통해 정리하고 정 분류율, 오분류율, 민감도, 특이도를 가지고 예측 모형이 잘 만들어졌는지를 확인하는 도구라고 할 수 있다.

<표 2-1> 분류행렬 (Confusion Matrix)

실제	예측	
	A (N)	B (Y)
A(N)	True Negatives(TN) : 올바르게 분류된 A 데이터 수	False Positives(FP) : 잘못 분류된 A 데이터 수
B(Y)	False Negatives(FN) : 잘못 분류된 B 데이터 수	True Positives(TP) : 올바르게 분류된 B 데이터 수

이 때 정분류율이라고 하는 것은 올바른 답을 올바르게 예측하고, 틀린 답을 틀리게 예측한 지표라고 할 수 있다. 분류행렬의 표에서 공식으로 표현하면 다음과 같이 표현될 수 있다.

$$(TP+TN)/(TP+TN+FN+FP).$$

다음으로 오분류율은 정분류율과는 다르게 잘못된 답을 선택되어 잘못된 답을 예측한 비율을 의미한다.

$$(FN+FP)/(TP+TN+FN+FP).$$

다음으로 민감도라고 하는 것은 실제 답인 것을 답으로 예측하는 성능평가 척도라고 할 수 있다. 반대로 특이도의 경우에는 실제 0의 값을 0으로 맞추는 비율이라 할 수 있다. 공식으로 다음과 같다.

$$\text{민감도: } TP/(TP+FN)$$

$$\text{특이도: } TN/(TN+FP)$$

데이터 탐색

제1절 데이터 정의

제2절 전체 분석데이터 대상자 속성분석

제3절 단독변수를 갖는 대상자 속성분석

제4절 가구 수에 따른 대상자 속성분석

제5절 지역별 대상자 속성분석

제3장 데이터 탐색

제1절 데이터 정의

본 장에서는 환류데이터에 속성변수를 추가한 훈련데이터 기반의 분석을 수행한다. 훈련데이터는 차수별로 중복되는 대상자는 존재하지 않으며, 훈련데이터 생성하는 규칙에 의거하여 생성된 훈련데이터이다. 본 훈련데이터에는 공공부조와 민간서비스가 중복이 될 경우 공공부조 우선순위가 높게 하여 공공부조를 받은 것으로 카운트를 매긴다.

이에 훈련데이터의 모수는 다음과 같이 구성된다. 2019년 1차와 2차 훈련데이터 구성에 대하여 분석을 수행한다. 첫 번째로 기술통계 기반의 빈도 분석을 통해 2019년 1차와 2차의 현황을 우선적으로 분석하고자 한다.

<표 3-1> 훈련데이터(1차,2차)

구분	2019년 1차	2019년 2차
총 데이터 수	56,436건	53,965건

발굴대상자의 구분코드는 다음과 같이 구성되며, 본 연구에서는 200, 300, 600코드에 대해서만 고려하여 대상자 분석을 수행한다.

비대상자는 분배는 되었지만 방문하거나 상담한 결과 서비스를 제공할 필요가 없다고 판단된 대상자라고 할 수 있다. 미처리자는 지자

체에 대상자로 분배는 되었지만 어느 조치도 취하지 않은 경우를 의미한다. 그리고 요청접수는 서비스를 받고 싶어하는 대상자가 서비스 요청을 접수 한 상황을 의미한다. 다음으로 조치중 대상자의 경우 접수는 되었지만 아직까지 조치완료가 되지 않은 상황을 의미한다. 마지막으로 조치완료의 경우 서비스를 최종적으로 대상자에게 주기위해 공공부조 또는 민간서비스 연계를 하여 최종 서비스를 대상자에게 전달된 경우를 의미한다.

〈표 3-2〉 복지사각지대 코드 및 설명

코드	설명
700	미지원
100	미처리
200	단순상담
300	비대상
400	요청접수
500	조치중
600	조치완료

연령은 다음과 같이 구성한다. 영·유아기는 만0~6세로 아동·청소년기는 만7~18세, 청년기는 만19~39세, 장·중년기는 만40~64세, 노년기는 만65세 이상으로 구성하였다.

〈표 3-3〉 연령구분

구분	영·유아기	아동·청소년기	청년기	장·중년기	노년기
연령대	만0~6세	만7~18세	만19~39세	만40~64세	만65세 이상

데이터의 지역 구성은 다음과 같이 서울특별시, 부산광역시, 대구광역시, 인천광역시, 광주광역시, 대전광역시, 울산광역시, 세종특별자치시, 경기도, 강원도, 충청북도, 충청남도, 전라북도, 전라남도, 경상북도, 경상남도, 제주특별자치도로 총 17개의 시군구로 구성되어 있다.

제2절 전체 분석데이터 대상자 속성분석

본 절에서는 복지사각지대 데이터의 기초 분석을 수행하기 위하여 시도별, 선정 대상자별, 연령대별, 변수의 개수별로 분석을 수행한다. 또한 지역별 특성 분석을 추가로 하기 위해서 17개의 시군구 중 서울특별시, 경기도, 전라북도, 제주특별자치도로 분리하여 지역 특성 분석을 수행한다.

<표 3-4> 복지사각지대 선정 대상자 및 조치 현황분석 (19년 1차, 2차)

(단위: 명, %)

조치결과	19년1차 2019.01	비율 (%)	19년 2차 2019.03	비율 (%)
전체	56,436	100%	53,965	100%
단순상담 -200	21,786	38.60%	24,918	46.17%
비대상 -300	5,399	9.57%	4,640	8.60%
조치완료 -600	29,251	51.83%	24,407	45.23%

복지사각지대 선정 대상자 및 조치 현황은 <표 3-4>와 같이 볼 수 있다. 19년 1차에서는 단순 상담의 경우 38.6%, 비대상은 9.57%, 조치완료의 경우 51.83%로 조사되었다. 그리고 19년 2차의 경우 단순 상담은 46.17%, 비대상은 8.6%, 그리고 조치완료의 경우 45.23%로 조사되었다. 이와 같은 결과를 통하여 전체적인 조치완료는 6% 정도 떨어지는 것을 확인할 수 있었고, 단순상담의 비율은 8% 정도 증가되는 것을 볼 수 있었다.

<표 3-5> 시도별 19년 1, 2차 복지사각지대 선정 대상자 현황

(단위: 명, %)

조치결과	19년1차	비율(%)	19년2차	비율(%)	증감수	증감비율(%)
전체	56436	100%	53965	100%		
서울	4584	8%	6282	12%	1698	4%
부산	2486	4%	3012	6%	526	1%
대구	1576	3%	2064	4%	488	1%
인천	1844	3%	2155	4%	311	1%
광주	1127	2%	1361	3%	234	1%
대전	910	2%	1255	2%	345	1%
울산	539	1%	750	1%	211	0%
세종	190	0%	174	0%	-16	0%
경기	13012	23%	11835	22%	-1177	-1%
강원	2469	4%	1840	3%	-629	-1%
충북	2421	4%	2136	4%	-285	0%
충남	3473	6%	2730	5%	-743	-1%
전북	3747	7%	3520	7%	-227	0%
전남	4759	8%	4212	8%	-547	-1%
경북	4536	8%	3572	7%	-964	-1%
경남	4984	9%	4606	9%	-378	0%
제주	922	2%	922	2%	0	0%
결측	2857	5%	1539	3%	-1318	-2%

19년 1차와 2차에서 시도별 복지대상자 선정 현황을 분석해볼 결과는 <표 3-5>와 같이 볼 수 있다. 1차를 기준으로 경기도가 가장 높은 23%의 비율을 보여주고 있으며, 다음으로 경상남도 9%, 전라남도, 경상북도, 서울특별시 순으로 대상자 비율이 8~9%로 보여주고 있다. 2차 기준으로 비율을 보면 경기도가 1차와 마찬가지로 22%로 가장 높은 비율을 보여주고 있으며, 서울특별시가 12%로 높은 비율을 보여주고 있다. 경상남도과 전라남도의 세종특별자치시, 제주시의 경우 1차와 2차의 비율은 유지되는 것을 알 수 있었다.

이와 같은 결과를 통하여 1개의 특별시와 6개의 자치시의 경우 1차에 비해 2차에서는 증가된 경향을 볼 수 있었으며, 세종특별자치시를 제외한 나머지 지역에서는 1차수에 비해 2차수에서는 감소되는 경향을 볼 수 있었다.

<표 3-6> 연령대별 복지사각지대 선정 대상자 현황

(단위: 명, %)

조치결과	19년1차	비율(%)	19년2차	비율(%)	증감수	증감비율(%)
전체	56436	100%	53965	100%		
영유	4213	7%	1722	3%	-2491	-4%
아청	2463	4%	3984	7%	1521	3%
청년	7402	13%	11435	21%	4033	8%
장중	15029	27%	17426	32%	2397	6%
노년	27329	48%	19398	36%	-7931	-12%

연령대별 복지사각지대 선정 대상자 현황은 <표 3-6>과 같이 볼 수 있다. 19년 1차에서는 노년층의 비율이 가장 높은 것을 볼 수 있으며, 2차도 마찬가지로 노년층의 비율이 가장 높은 것을 확인할 수 있다. 반면에 19년 1차에서 가장 낮은 비율을 차지하고 있는 연령대는 아동·청소년기라는 것을 알 수 있으며, 2차에서 가장 낮은 비율을 차지하고 있는 연령대의 경우는 영·유아기라는 것을 알 수 있다.

영·유아기의 경우 1차 4,213명(7%)에서 2차 1,722명(3%)로 2,491명(4%) 감소되었으며, 아동·청소년기의 경우 1차 2,463명(4%)에서 2차 3,984명(7%)로 1,521명(3%) 증가된 것을 볼 수 있다. 또한 청년기의 경우 7,402명(13%)에서 2차 11,435명(21%)로 증가된 것을 알 수 있으며, 장중년기의 경우 15,029명(27%)에서 2차 17,426명(32%)로

2,397명(6%)로 증가된 것을 알 수 있다. 마지막 노년기 층의 경우 1차 27,329명(48%)에서 19,398명(36%)로 7,931(12%)로 큰 폭으로 감소된 것을 알 수 있다.

<표 3-7> 연계변수별 복지사각지대 선정 대상자 현황

(단위: 명, %)

연계변수	19년1차	비율(%)	2차	비율(%)	증감 수	증감비율(%)
합계	36814	1	33491	1		
단전	40	0.0011	41	0.0012	1	0.0001
단수	7	0.0002	1	0.0000	-6	-0.0002
단가	102	0.0028	76	0.0023	-26	-0.0005
전기	394	0.0107	459	0.0137	65	0.0030
연금	142	0.0039	344	0.0103	202	0.0064
건보	3083	0.0837	4758	0.1421	1675	0.0583
화재	10	0.0003	11	0.0003	1	0.0001
경감	0	0.0000	0	0.0000	0	0.0000
장기	14577	0.3960	6737	0.2012	-7840	-0.1948
전세	2121	0.0576	2624	0.0783	503	0.0207
월세	11403	0.3097	11847	0.3537	444	0.0440
개별	57	0.0015	37	0.0011	-20	-0.0004
실직	106	0.0029	205	0.0061	99	0.0032
비대	515	0.0140	608	0.0182	93	0.0042
집중	133	0.0036	37	0.0011	-96	-0.0025
분유	497	0.0135	551	0.0165	54	0.0030
청각	0	0.0000	0	0.0000	0	0.0000
예방	18	0.0005	8	0.0002	-10	-0.0003
시도	9	0.0002	8	0.0002	-1	0.0000
위기	212	0.0058	420	0.0125	208	0.0068
범죄	0	0.0000	0	0.0000	0	0.0000
시설	73	0.0020	218	0.0065	145	0.0045
기초	1640	0.0445	1770	0.0529	130	0.0083
공공	122	0.0033	481	0.0144	359	0.0110
산재	16	0.0004	14	0.0004	-2	0.0000
재난	0	0.0000	0	0.0000	0	0.0000
금융	486	0.0132	1138	0.0340	652	0.0208
의료	374	0.0102	315	0.0094	-59	-0.0008
일용	462	0.0125	655	0.0196	193	0.0070
영양	215	0.0058	128	0.0038	-87	-0.0020

연계변수별 선정 대상자 현황은 <표 3-7>과 같이 볼 수 있다. 각 차수의 연계변수 비율의 경우는 각 차수의 연계변수의 전체 선정대상자를 기준으로 산출하였다. 19년 1차의 경우 피부양의무자장기요양여부 변수가 39.6%, 월세금액기준이하가구여부 변수의 경우 30.97%로 연계변수가 많이 사용되었다. 19년 2차의 경우 월세금액기준이하가구여부 변수가 35.37%로 가장 많이 사용되었으며, 피부양의무자장기요양변수가 20.12%로 두 번째로 많이 사용되었다. 건강보험료체납 변수도 14.21%로 3번째로 많이 사용되었다.

증감여부기준으로 보면 피부양장기요양여부 변수의 경우 39.6%에서 20.12%로 19.48% 감소되었다. 반면 월세금액기준이하가구의 경우 30.97%에서 35.37%로 4.4% 증가 된 것을 알 수 있으며, 전세금액기준이하가구의 경우 5.76%에서 7.83%로 2.07% 증가된 것을 볼 수 있다.

뿐만 아니라 신생아청각선별검사대상자여부, 범죄피해여부의 경우 대상자 선정 비율이 0으로써 19년 1차와 2차의 경우 영향이 없다는 것을 알 수 있다.

<표 3-8> 연계변수 조합 개수별 복지사각지대 선정 대상자 현황

(단위: 명, %)

구분	19년1차	비율 (%)	19년2차	비율 (%)	증감 수	증감 비율 (%)
전체	56436	100	53965	100		
1개	45478	80.5833	38501	71.3444	-6977	-9.2389
2개	9068	16.0678	12349	22.8834	3281	6.8156
3개	1642	2.9095	2632	4.8772	990	1.9677
4개	224	0.3969	446	0.8265	222	0.4296
5개	22	0.0390	36	0.0667	14	0.0277
6개	2	0.0035	1	0.0019	-1	-0.0017

연계변수 조합 개수별 복지사각지대 선정 대상자 현황은 <표 3-8>과 같이 볼 수 있다. 19년 1차의 경우 1개 변수를 가진 대상자의 경우 45,478명(80.58%)으로 가장 높은 대상자를 보여주었으며, 다음으로 2개 변수를 가진 대상자는 9,068명(16.07%)로 두 번째로 높은 대상자를 보여주었다. 3개 변수를 가진 대상자는 1,642명(2.91%), 4개를 가진 대상자는 224명(0.39%), 5개를 가진 대상자는 22명(0.04%), 마지막 6개 변수를 가진 대상자는 2명으로 파악되었다. 이와 비슷하게 19년 2차에서 1개 변수를 가진 대상자의 경우 38,501명(71.34%)으로 제일 높은 대상자 수를 보였다. 다음으로 2개 변수를 가진 대상자는 12,349명(22.88%)를 보였다. 3개 변수를 가진 대상자는 2,632명(4.88%), 4개 변수를 가진 대상자는 446명(0.83%), 5개 변수를 가진 대상자는 36(0.07%), 마지막으로 6개 변수를 보유한 대상자는 1명으로 추정되었다. 이런 결과를 통하여 전체적으로 1개 변수를 가진 대상자가 대부분을 차지한다는 것을 알 수 있다.

<표 3-9> 지역별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	56436	21786	38.6030	5399	9.5666	29251	51.8304
서울	4584	2399	52.3342	268	5.8464	1917	41.8194
부산	2486	1098	44.1673	236	9.4932	1152	46.3395
대구	1576	398	25.2538	50	3.1726	1128	71.5736
인천	1844	418	22.6681	130	7.0499	1296	70.2820
광주	1127	512	45.4303	152	13.4871	463	41.0825
대전	910	427	46.9231	50	5.4945	433	47.5824
울산	539	146	27.0872	2	0.3711	391	72.5417
세종	190	106	55.7895	20	10.5263	64	33.6842
경기	13012	4948	38.0264	1386	10.6517	6678	51.3219
강원	2469	1016	41.1503	772	31.2677	681	27.5820
충북	2421	1178	48.6576	187	7.7241	1056	43.6183
충남	3473	1146	32.9974	196	5.6435	2131	61.3591
전북	3747	1897	50.6272	344	9.1807	1506	40.1922
전남	4759	1493	31.3721	281	5.9046	2985	62.7233
경북	4536	1228	27.0723	424	9.3474	2884	63.5802
경남	4984	1358	27.2472	550	11.0353	3076	61.7175
제주	922	319	34.5987	47	5.0976	556	60.3037
결측	2857	1699	59.4680	304	10.6405	854	29.8915

<표 3-9>는 지역별 선정 대상자를 기준으로 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 세종특별자치시 55.79%, 전라북도 50.63%, 충청북도 48.66%, 서울특별시 44.17%, 순으로 높은 것을 볼 수 있다. 그리고 비대상의 경우 강원도 31.27%, 광주광역시 13.49%, 경상남도 11.04%, 경기도 10.65% 순으로 높은 것을 볼 수 있다. 조치완료의 경우 울산광역시 72.54%, 대구광역시 71.57%, 인천광역시 70.28%, 경상북도 63.58% 순으로 높은 것을 알 수 있다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 인천광역시 22.67%, 대구광역시 25.25%, 경상북도 27.07%, 울산광역시 27.08%가 낮은 것을 알 수 있다. 비대상 기준으로 보면 울산광역시 0.37%, 대구광역시 3.17%, 제주특별자치도 5.1%, 대전광역시 5.49%순으로 낮은 것을 알 수 있다. 마지막으로 조치완료를 낮은 순으로 보면 강원도 27.58%, 세종특별자치시 33.68%, 전라북도 40.19%, 광주광역시 41.08%순으로 낮은 것을 알 수 있다.

<표 3-10> 지역별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	53965	24918	46.1744	4640	8.5982	24407	45.2275
서울	6282	4519	71.9357	395	6.2878	1368	21.7765
부산	3012	1597	53.0212	415	13.7782	1000	33.2005
대구	2064	469	22.7229	135	6.5407	1460	70.7364
인천	2155	597	27.7030	175	8.1206	1383	64.1763
광주	1361	771	56.6495	233	17.1198	357	26.2307
대전	1255	804	64.0637	68	5.4183	383	30.5179
울산	750	190	25.3333	3	0.4000	557	74.2667
세종	174	111	63.7931	13	7.4713	50	28.7356
경기	11835	4984	42.1124	1101	9.3029	5750	48.5847
강원	1840	1010	54.8913	515	27.9891	315	17.1196
충북	2136	1236	57.8652	148	6.9288	752	35.2060
충남	2730	942	34.5055	97	3.5531	1691	61.9414
전북	3520	2109	59.9148	232	6.5909	1179	33.4943
전남	4212	1636	38.8414	219	5.1994	2357	55.9592
경북	3572	1079	30.2072	315	8.8186	2178	60.9742
경남	4606	1317	28.5931	488	10.5949	2801	60.8120
제주	922	398	43.1670	47	5.0976	477	51.7354
결측	1539	1149	74.6589	41	2.6641	349	22.6771

<표 3-10>은 지역별 선정 대상자를 기준으로 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 서울특별시 71.94%, 대전광역시 64.06%, 세종특별자치시 63.79%, 전라북도 59.91% 순으로 비율이 높은 것을 알 수 있다. 그리고 비대상의 경우 강원도 27.99%, 광주광역시 17.12%, 부산광역시 13.78%, 경상남도 10.59% 순으로 높은 것을 알 수 있다. 조치완료의 경우 울산광역시 74.27%, 대구광역시 70.74%, 인천광역시 64.18%, 충청남도 61.94% 순으로 높은 것을 알 수 있다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 대구광역시 22.72%, 울산광역시 25.33%, 인천광역시 27.70%, 경상북도 30.20% 순으로 낮은 것을 알 수 있다. 비대상 기준으로 보면 울산광역시 0.4%, 충청남도 3.55%, 제주특별자치도 5.1%, 전라남도 5.2%순으로 낮은 것을 알 수 있다. 마지막으로 조치완료를 낮은 순으로 보면 강원도 17.11%, 서울특별시 21.77%, 광주광역시 26.23%순으로 낮은 것을 알 수 있다.

<표 3-11> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	56436	21786	38.6030	5399	9.5666	29251	51.8304
영유	1384	338	24.4220	166	11.9942	880	63.5838
아청	2463	1040	42.2249	281	11.4089	1142	46.3662
청년	7402	3932	53.1208	711	9.6055	2759	37.2737
장중	15029	6421	42.7241	1395	9.2821	7213	47.9939
노년	27329	8377	30.6524	2543	9.3051	16409	60.0424
결측	2829	1678	59.3142	303	10.7105	848	29.9753

연령대별 복지사각지대 선정 대상자 2019년 1차 완료결과 현황은 <표 3-11>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층의 경우 제일 높은 단순상담 비율을 보여주었다. 반면에 노년층의 경우 30.65%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 청년층 다음으로 장중년층과 아동청소년층이 42.72%, 42.22%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아가 비대상의 비율이 가장 높으며 비슷하게 아동청소년이 11.4%로 다음으로 높은 것을 알 수 있다. 반면 노년층과 장중년층의 경우 비대상의 비율이 비슷하게 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 영유아층이 전체적으로 뽑힌 대상자의 수는 작지만 조치완료율은 63.58%로 제일 높은 것을 알 수 있다. 다음으로 노년층이 60.04%로 두 번째로 높은 비율을 보여주는 것을 알 수 있다. 청년층의 경우 37.27%로 제일 낮은 비율을 보여주었다.

<표 3-12> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	53965	24918	46.1744	4640	8.5982	24407	45.2275
영유	1721	656	38.1174	252	14.6426	813	47.2400
아청	3984	1767	44.3524	492	12.3494	1725	43.2982
청년	11435	6549	57.2715	1057	9.2436	3829	33.4849
장중	17426	8523	48.9097	1523	8.7398	7380	42.3505
노년	19398	7422	38.2617	1316	6.7842	10660	54.9541
결측	1	1	100.0000	0	0	0	0

연령대별 복지사각지대 선정 대상자 2019년 2차 완료결과 현황은 <표 3-12>와 같이 볼 수 있다. 단순상담을 기준으로 보면 2019년 1차와 마찬가지로 청년층의 경우 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아의 경우 38.11%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아가 비대상의 비율이 가장 높으며 비슷하게 아동청소년이 12.35%로 다음으로 높은 것을 알 수 있다. 반면 노년층의 비대상의 비율이 6.78%로 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년층이 54.95%로 가장 높은 조치완료율을 보여주는 것일 수 있다. 다음으로 영유아층의 경우 47.24%로 두 번째로 높은 조치완료율을 보여준다. 반면에 청년층의 경우 33.48%로 제일 낮은 조치완료율을 보여준다.

<표 3-13> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	56436	21786	38.6030	5399	9.5666	29251	51.8304
1개	45478	17687	38.8913	4626	10.1720	23165	50.9367
2개	9068	3418	37.6930	669	7.3776	4981	54.9294
3개	1642	588	35.8100	88	5.3593	966	58.8307
4개	224	81	36.1607	14	6.2500	129	57.5893
5개	22	12	54.5455	2	9.0909	8	36.3636
6개	2	0	0.0000	0	0.0000	2	100.0000

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-13>의 경우 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 6개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 1개를 가진 대상이 38.89%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 4개, 3개 그리고 6개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 5개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 6개를 가진 대상이 100%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 58.83%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 5개를 가진 대상이 36.36%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-14> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	53965	24918	46.1744	4640	8.5982	24407	45.2275
1개	38501	17762	46.1339	3449	8.9582	17290	44.9079
2개	12349	5736	46.4491	968	7.8387	5645	45.7122
3개	2632	1208	45.8967	183	6.9529	1241	47.1505
4개	446	195	43.7220	39	8.7444	212	47.5336
5개	36	16	44.4444	1	2.7778	19	52.7778
6개	1	1	100.0000	0	0.0000	0	0.0000

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-14>의 경우 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 6개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 1개를 가진 대상이 38.89%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 4개, 3개 그리고 6개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 5개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 5개를 가진 대상이 52.78%로 제일 높은 것을 알 수 있다. 다음으로 변수 4개를 가진 대상이 47.53%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 6개를 가진 대상이 0%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-15> 선정 대상자 2019년 1차의 변수별 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	71	27	38.0282	4	5.6338	40	56.3380
단수	18	10	55.5556	1	5.5556	7	38.8889
단가	209	103	49.2823	6	2.8708	100	47.8469
전기	847	397	46.8713	58	6.8477	392	46.2810
연금	339	173	51.0324	27	7.9646	139	41.0029
건보	7038	3532	50.1847	440	6.2518	3066	43.5635
화재	15	5	33.3333	0	0.0000	10	66.6667
경감	0	0		0		0	
장기	26351	9040	34.3061	2788	10.5802	14523	55.1137
전세	4401	1833	41.6496	461	10.4749	2107	47.8755
월세	20830	7887	37.8637	1644	7.8925	11299	54.2439
개별	84	22	26.1905	5	5.9524	57	67.8571
실직	190	69	36.3158	15	7.8947	106	55.7895
비대	1073	455	42.4045	105	9.7856	513	47.8099
집중	195	51	26.1538	13	6.6667	131	67.1795
분유	778	186	23.9075	97	12.4679	495	63.6247
청각	0	0		0		0	
예방	25	6	24.0000	1	4.0000	18	72.0000
시도	19	10	52.6316	0	0.0000	9	47.3684
위기	459	199	43.3551	48	10.4575	212	46.1874
범죄	0	0		0		0	
시설	127	49	38.5827	7	5.5118	71	55.9055
기초	2659	819	30.8011	208	7.8225	1632	61.3765
공공	216	80	37.0370	14	6.4815	122	56.4815
산재	28	11	39.2857	1	3.5714	16	57.1429
재난	0	0		0		0	
금융	1222	668	54.6645	71	5.8101	483	39.5254
의료	720	285	39.5833	63	8.7500	372	51.6667
일용	1232	649	52.6786	122	9.9026	461	37.4188
영양	412	105	25.4854	95	23.0583	212	51.4563

<표 3-15>는 선정 대상자 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단수여부 55.56%, 금융연체여부의 54.66%, 일용근로자여부 52.68%, 자살시도대상자여부 52.63% 순으로 단순상담의 비율이 높은 것을 알 수 있다. 비대상의 경우 영양플러스미지원가구 23.06%, 기저귀조제분유지원대상자여부 12.47%, 피부양 의무자장기요양여부 10.58%, 전세금액기준이하가구여부 10.47%, 위기학생여부 10.46% 순으로 비율이 높은 것을 알 수 있다. 조치완료의 경우 자살예방관리대상자여부 72%, 고용보험개별연장급여대상자여부 67.86%, 방문건강집중관리군여부 67.18%, 화재피해여부 66.67%, 기저귀조제분유지원대상자여부 63.62% 순으로 높게 나타났다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 기저귀조제분유지원대상자여부 23.91%, 자살예방관리대상자여부 24%, 영양플러스미지원가구 25.49%, 방문건강집중관리군여부 26.15%, 고용보험개별연장급여대상자여부 26.19% 순으로 낮은 것을 볼 수 있다. 비대상 기준으로는 화재피해여부와 자살시도대상자여부가 0%로 제일 낮으며, 단가스여부 2.87%, 산재요양종결후근로단절자여부 3.57%, 자살예방관리대상자여부 4% 순으로 낮은 것을 볼 수 있다. 조치완료의 경우 일용근로자여부 37.42%, 단수도여부 38.89%, 금융연체여부 39.53%, 국민연금체납여부 41%, 건강보험료체납여부 43.56% 순으로 낮은 것을 알 수 있다.

<표 3-16> 선정 대상자 2019년 2차 변수별 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	97	47	48.4536	10	10.3093	40	41.2371
단수	2	1	50.0000	0	0.0000	1	50.0000
단가	174	90	51.7241	8	4.5977	76	43.6782
전기	1169	614	52.5235	97	8.2977	458	39.1788
연금	797	392	49.1844	61	7.6537	344	43.1619
건보	11274	5769	51.1708	782	6.9363	4723	41.8929
화재	26	12	46.1538	3	11.5385	11	42.3077
경감	0	0		0		0	
장기	12047	4593	38.1257	736	6.1094	6718	55.7649
전세	6197	2982	48.1201	621	10.0210	2594	41.8590
월세	25659	11559	45.0485	2365	9.2170	11735	45.7344
개별	68	25	36.7647	6	8.8235	37	54.4118
실직	523	268	51.2428	50	9.5602	205	39.1969
비대	1523	755	49.5732	162	10.6369	606	39.7899
집중	68	21	30.8824	11	16.1765	36	52.9412
분유	1005	358	35.6219	104	10.3483	543	54.0299
청각	0	0		0		0	
예방	9	1	11.1111	0	0.0000	8	88.8889
시도	18	8	44.4444	2	11.1111	8	44.4444
위기	942	408	43.3121	119	12.6327	415	44.0552
범죄	0	0		0		0	
시설	416	178	42.7885	21	5.0481	217	52.1635
기초	3551	1453	40.9181	347	9.7719	1751	49.3101
공공	1004	454	45.2191	69	6.8725	481	47.9084
산재	31	13	41.9355	4	12.9032	14	45.1613
재난	0	0		0		0	
금융	3467	2170	62.5901	161	4.6438	1136	32.7661
의료	732	333	45.4918	84	11.4754	315	43.0328
일용	1879	1057	56.2533	175	9.3135	647	34.4332
영양	387	163	42.1189	97	25.0646	127	32.8165

<표 3-16>은 선정 대상자 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 금융연체 62.59%, 일용근로자여부 56.25%, 전기료 체납여부 52.52%, 단가스여부 51.72%, 고용보험실직사유대상여부 51.24% 순으로 단순상담의 비율이 높은 것을 알 수 있다. 비대상의 경우 영양플러스미지원가구여부 25.06%, 방문건강집중관리군여부 16.18%, 산재요양종결후근로단절자여부 12.9%, 위기학생여부 12.63%, 화재피해여부 11.54% 순으로 비율이 높은 것을 알 수 있다. 조치완료의 경우 자살예방관리대상자여부 88.89%, 피부양의무자장기요양여부 55.76%, 고용보험개별연장급여대상자여부 54.41%, 기저귀조제분유지원대상자여부 54.03%, 방문건강집중관리군여부 52.94% 순으로 높은 것을 알 수 있다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 피부양의무자장기요양여부 38.13%, 고용보험개별연장급여대상자여부 36.76%, 기저귀조제분유지원대상자여부 35.62%, 방문건강집중관리군여부 30.88%, 시설퇴소여부 42.79% 순으로 낮은 것을 볼 수 있다. 비대상 기준으로는 자살예방관리대상자여부, 단수도여부의 경우 0%로 나타났다. 그리고 단가스여부 4.59%, 금융연체여부 4.64%, 시설입퇴소여부 5.05% 순으로 낮게 나타났다. 조치완료의 경우 금융연체여부 32.77%, 영양플러스미지원가구여부 32.82%, 일용근로자여부 34.43%, 전기료체납여부 39.18%, 고용보험실직사유대상여부 39.19% 순으로 낮게 나타났다.

제3절 단독변수를 가지는 대상자 속성분석

<표 3-17> 지역별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상당		비대상		조치완료	
전체	42677	16022	37.5425	4324	10.1319	22331	52.3256
서울	2762	1494	54.0912	165	5.9739	1103	39.9348
부산	1633	703	43.0496	165	10.1041	765	46.8463
대구	1027	276	26.8744	38	3.7001	713	69.4255
인천	1129	255	22.5864	87	7.7059	787	69.7077
광주	751	344	45.8056	114	15.1798	293	39.0146
대전	641	311	48.5179	41	6.3963	289	45.0858
울산	377	107	28.3820	1	0.2653	269	71.3528
세종	165	94	56.9697	17	10.3030	54	32.7273
경기	10665	4126	38.6873	1233	11.5612	5306	49.7515
강원	2105	872	41.4252	671	31.8765	562	26.6983
충북	2039	985	48.3080	165	8.0922	889	43.5998
충남	2996	974	32.5100	171	5.7076	1851	61.7824
전북	3184	1631	51.2249	304	9.5477	1249	39.2274
전남	4165	1319	31.6687	242	5.8103	2604	62.5210
경북	3997	1101	27.5457	389	9.7323	2507	62.7220
경남	4304	1165	27.0678	483	11.2221	2656	61.7100
제주	737	265	35.9566	38	5.1560	434	58.8874

<표 3-17>은 지역별 선정 대상자를 기준으로 변수를 1개 가진 대상자에 한정하여 2019년 1차 완료결과 현황을 보여주고 있다. 단순상당 기준으로 제일 높은 비율을 보여주는 지역을 보면 세종특별자치시

56.96%, 서울특별시 54.09%, 전라북도 51.22%, 대전광역시 48.52%, 충청북도 48.31% 순으로 높게 나타났다. 비대상 기준으로 강원도 31.88%, 광주광역시 15.18%, 경기도 11.56%, 경상남도 11.22%, 세종특별자치시 10.3% 순으로 높게 나타나는 것을 알 수 있었다. 그리고 조치완료의 경우 울산광역시 71.35%, 인천광역시 69.71%, 대구광역시 69.43%, 경상북도 62.72%, 전라남도 62.52% 순으로 높게 나타났다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 인천광역시 22.59%, 대구광역시 26.87%, 경상남도 27.07%, 경상북도 27.55%, 울산광역시 28.38% 순으로 낮게 나타났다. 비대상 기준으로 보면 울산광역시 0.27%, 대구광역시 3.7%, 제주특별자치도 5.16%, 충청남도 5.71%, 전라남도 5.81% 순으로 낮게 나타났다. 조치완료 기준으로 보면 강원도 26.69%, 세종특별자치시 32.73%, 광주광역시 39.01%, 전라북도 39.23%, 서울특별시 39.93% 순으로 낮게 나타났다.

<표 3-18> 지역별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	37422	16953	45.3022	3419	9.1363	17050	45.5614
서울	4101	3001	73.1773	259	6.3155	841	20.5072
부산	1984	1031	51.9657	302	15.2218	651	32.8125
대구	1319	321	24.3366	95	7.2024	903	68.4610
인천	1284	376	29.2835	129	10.0467	779	60.6698
광주	863	498	57.7057	167	19.3511	198	22.9432
대전	856	561	65.5374	41	4.7897	254	29.6729
울산	499	133	26.6533	3	0.6012	363	72.7455
세종	120	81	67.5000	12	10.0000	27	22.5000
경기	8453	3647	43.1444	823	9.7362	3983	47.1194
강원	1346	736	54.6805	369	27.4146	241	17.9049
충북	1576	920	58.3756	109	6.9162	547	34.7081
충남	2047	667	32.5843	74	3.6150	1306	63.8007
전북	2625	1530	58.2857	178	6.7810	917	34.9333
전남	3342	1307	39.1083	173	5.1765	1862	55.7151
경북	2726	845	30.9978	237	8.6941	1644	60.3081
경남	3644	1021	28.0187	420	11.5258	2203	60.4555
제주	637	278	43.6421	28	4.3956	331	51.9623

<표 3-18>은 지역별 선정 대상자를 기준으로 변수를 1개 가진 대상자에 한정하여 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 서울특별시 73.18%, 세종특별자치시 67.5%, 대전광역시 65.54%, 충청북도 58.38%, 전라북도 58.29% 순으로 높은 것을 보여주었다. 비대상 기준으로 강원도 27.41%, 광주광역시 19.35%, 부산광역시 15.22%, 경

상남도 11.53%, 인천광역시 10.05% 순으로 높게 나타났다. 조치완료를 기준으로 보면 울산광역시 72.75%, 대구광역시 68.46%, 충청남도 63.8%, 인천광역시 60.67%, 경상남도 60.46% 순으로 높게 나타났다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 대구광역시 24.34%, 울산광역시 26.65%, 경상남도 28.02%, 인천광역시 29.28%, 경상북도 30.99% 순으로 낮게 나타났다. 비대상 기준으로 보면 울산광역시 0.6%, 충청남도 3.62%, 제주특별자치도 4.4%, 대전광역시 4.78%, 전라남도 5.17% 순으로 낮게 나타났다. 조치완료의 경우 강원도 17.9%, 서울특별시 20.51%, 세종특별자치시 22.5%, 광주광역시 22.94%, 대전광역시 29.67% 순으로 낮게 나타났다.

<표 3-19> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	45478	17687	38.8913	4626	10.1720	23165	50.9367
영유	3683	1892	51.3712	416	11.2951	1375	37.3337
아청	1441	641	44.4830	171	11.8668	629	43.6502
청년	5437	2945	54.1659	574	10.5573	1918	35.2768
장중	11752	5034	42.8353	1144	9.7345	5574	47.4302
노년	23165	7175	30.9735	2321	10.0194	13669	59.0071

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

연령대별 복지사각지대 선정 대상자 중 변수 1개를 보유한 대상자의 2019년 1차 완료결과 현황은 <표 3-19>와 같이 볼 수 있다. 단

순상담을 기준으로 보면 청년층의 경우 제일 높은 단순상담 비율을 보여주었다.

반면에 노년층의 경우 30.97%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 청년층 다음으로 영유아와 아동청소년 층이 51.37%, 44.48%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 아동청소년 층이 11.87%, 영유아층이 11.3%으로 비대상 비율이 높게 나타났다. 그리고 장중년층의 경우 비대상 비율이 가장 낮게 나타났다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년층이 59%, 장중년층이 47.43% 순으로 높게 나타났다. 청년층은 35.28%, 영유아층은 37.33% 순으로 낮게 나타났다.

<표 3-20> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	38501	17762	46.1339	3449	8.9582	17290	44.9079
영유	1093	448	40.9881	187	17.1089	458	41.9030
아청	2184	1012	46.3370	281	12.8663	891	40.7967
청년	7469	4365	58.4416	740	9.9076	2364	31.6508
장중	11926	5917	49.6143	1142	9.5757	4867	40.8100
노년	15829	6020	38.0315	1099	6.9430	8710	55.0256

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

연령대별 복지사각지대 선정 대상자 중 변수 1개를 보유한 대상자의 2019년 2차 완료결과 현황은 <표 3-20>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층의 경우 제일 높은 단순상담 비율을 보여주었다.

반면에 노년층의 경우 38.03%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 청년층 다음으로 장중년층과 아동청소년층이 49.61%, 46.34%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아가 비대상의 비율이 가장 높으며 비슷하게 아동청소년이 12.87%로 다음으로 높은 것을 알 수 있다. 반면 노년층과 장중년층의 경우 비대상의 비율이 비슷하게 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년층이 55.03%, 영유아가 41.9% 순으로 높게 나타났다. 그리고 청년층의 경우 31.65%로 제일 낮은 비율을 보여주었다.

<표 3-21> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	45478	17687	38.8913	4626	10.1720	23165	50.9367
1개	45478	17687	38.8913	4626	10.1720	23165	50.9367

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-21>은 변수 조합 개수별 선정대상자의 1차 완료현황을 보여주고 있다. 1개 변수만을 가진 대상자만 존재한다. 이 때 조치완료는 50.93%로 가장 높게 보여주었으며, 단순상담의 경우 38.89%로 두 번째로 높게 나타났다. 마지막으로 비대상의 경우 10.17%로 나타났다. 2~6개 변수를 가진 대상자는 표시하지 않았다. 그 이유는 1개 대상자만을 분석하기 때문에 나머지 대상의 변수는 보유하고 있지 않기 때문이다.

<표 3-22> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상당		비대상		조치완료	
전체	38501	17762	46.1339	3449	8.9582	17290	44.9079
1개	38501	17762	46.1339	3449	8.9582	17290	44.9079

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-22>는 변수 조합 개수별 선정대상자의 2차 완료현황을 보여주고 있다. 1개 변수만을 가진 대상자만 존재한다. 이 때 단순상당은 46.13%로 가장 높게 보여주었으며, 조치완료의 경우 44.91%로 두 번째로 높게 나타났다. 마지막으로 비대상의 경우 8.96%로 나타났다. 마찬가지로 2~6개 변수를 가진 대상자는 표시하지 않았다. 그 이유는 1개 대상자만을 분석하기 때문에 나머지 대상의 변수는 보유하고 있지 않기 때문이다.

〈표 3-23〉 선정 대상자 1개 변수 보유자의 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상당		비대상		조치완료	
단전	4	0	0.0000	0	0.0000	4	100.0000
단수	0	0		0		0	
단가	4	2	50.0000	0	0.0000	2	50.0000
전기	24	15	62.5000	3	12.5000	6	25.0000
연금	11	4	36.3636	0	0.0000	7	63.6364
건보	3434	1888	54.9796	242	7.0472	1304	37.9732
화재	1	1	100.0000	0	0.0000	0	0.0000
경감	0	0		0		0	
장기	22099	7668	34.6984	2532	11.4575	11899	53.8441
전세	2201	1017	46.2063	255	11.5856	929	42.2081
월세	15646	6064	38.7575	1298	8.2961	8284	52.9464
개별	63	18	28.5714	5	7.9365	40	63.4921
실직	5	2	40.0000	1	20.0000	2	40.0000
비대	331	195	58.9124	54	16.3142	82	24.7734
집중	0	0		0		0	
분유	142	40	28.1690	26	18.3099	76	53.5211
청각	0	0		0		0	
예방	0	0		0		0	
시도	0	0		0		0	
위기	46	22	47.8261	8	17.3913	16	34.7826
범죄	0	0		0		0	
시설	0	0		0		0	
기초	0	0		0		0	
공공	1	1	100.0000	0	0.0000	0	0.0000
산재	0	0		0		0	
재난	0	0		0		0	
금융	392	248	63.2653	27	6.8878	117	29.8469
의료	74	23	31.0811	20	27.0270	31	41.8919
일용	663	396	59.7285	70	10.5581	197	29.7134
영양	337	83	24.6291	85	25.2226	169	50.1484

<표 3-23>은 1개 변수 보유자의 선정 대상자 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 화재피해여부변수를 보유한 대상자와 공공임대주택체납여부 변수를 보유한 대상자가 1명씩 나왔는데 모두 단순상담을 하였다. 그리고 금융연체, 전기료체납여부, 일용근로자여부 순으로 63.27%, 62.5%, 59.73% 순으로 높게 측정되었다. 비대상의 경우 의료비용과다지출가구여부 27.03%, 영양플러스미지원가구 25.22%, 고용보험실직사유대상여부 20%, 기저귀조제분유지원대상자여부 18.3%, 위기학생여부 17.39% 순으로 높게 나왔다. 조치완료의 경우 단전여부 4건이 모두 조치완료가 되었으며, 국민연금체납여부의 경우 63.63%, 고용보험개별연장급여대상자여부 63.49%, 피부양의무자장기요양여부 53.84%, 기저귀조제분유지원대상자여부 53.52% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 단전여부의 경우 0%, 영양플러스미지원가구 24.63%, 기저귀조제분유지원대상자여부 28.17%, 고용보험개별연장급여대상자여부 28.57%, 의료비용과다지출가구여부 31.08% 순으로 낮게 나왔다. 비대상 기준으로 살펴보면 다음과 같이 볼 수 있다. 단전여부, 국민연금체납여부, 단가스여부, 화재피해가구여부, 공공임대주택여부의 경우 0%가 나왔다. 그리고 금융연체, 건강보험료체납여부, 고용보험개별연장급여대상자여부, 월세금액기준이하가구여부가 6.88%, 7.05%, 7.94%, 8.29% 순으로 낮게 나왔다. 조치완료 기준으로 살펴보면 다음과 같다. 화재피해가구, 공공임대주택은 0건이며, 고용보험비대상여부 24.77%, 전기료체납여부 25%, 일용근로자여부 29.71%, 금융연체대상자여부 29.85%, 위기학생여부 34.78% 순으로 낮게 나왔다.

〈표 3-24〉 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	2	0	0.0000	0	0.0000	2	100.0000
단수	0	0		0		0	
단가	5	3	60.0000	0	0.0000	2	40.0000
전기	79	39	49.3671	15	18.9873	25	31.6456
연금	70	29	41.4286	13	18.5714	28	40.0000
건보	3336	1916	57.4341	255	7.6439	1165	34.9221
화재	6	2	33.3333	1	16.6667	3	50.0000
경감	0	0		0		0	
장기	9278	3454	37.2279	572	6.1651	5252	56.6070
전세	3011	1539	51.1126	335	11.1259	1137	37.7615
월세	18358	8232	44.8415	1780	9.6960	8346	45.4625
개별	47	18	38.2979	5	10.6383	24	51.0638
실직	162	83	51.2346	27	16.6667	52	32.0988
비대	480	277	57.7083	75	15.6250	128	26.6667
집중	1	0	0.0000	1	100.0000	0	0.0000
분유	73	42	57.5342	16	21.9178	15	20.5479
청각	0	0		0		0	
예방	0	0		0		0	
시도	0	0		0		0	
위기	182	79	43.4066	36	19.7802	67	36.8132
범죄	0	0		0		0	
시설	15	9	60.0000	1	6.6667	5	33.3333
기초	63	25	39.6825	12	19.0476	26	41.2698
공공	234	106	45.2991	24	10.2564	104	44.4444
산재	10	2	20.0000	3	30.0000	5	50.0000
재난	0	0		0		0	
금융	1590	1097	68.9937	71	4.4654	422	26.5409
의료	275	134	48.7273	34	12.3636	107	38.9091
일용	915	548	59.8907	98	10.7104	269	29.3989
영양	309	128	41.4239	75	24.2718	106	34.3042

<표 3-24>는 1개 변수 보유자의 선정 대상자 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 금융연체대상자여부 68.99%, 단가스여부 60%, 시설입퇴소여부 60%, 일용근로자여부 59.89%, 고용보험비대상여부 57.71% 순으로 높게 나왔다. 비대상의 경우 방문건강집중관리군여부의 경우 1건이 비대상으로 100%가 나왔다. 그리고 산재요양종결후근로단절자여부 30%, 영양플러스지원가구 24.27%, 기저귀조제분유지원대상자여부 21.92%, 위기학생여부 19.78% 순으로 높게 나왔다. 조치완료의 경우 단전여부의 경우 2건으로 2건 모두 조치완료가 되었다. 그리고 피부양의무자장기요양여부 56.61%, 고용보험개별연장급여대상자여부 51.06%, 산재요양종결후근로단절자여부 50%, 화재피해여부 50% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 단전여부와 방문건강집중관리군여부의 경우 0건이며, 산재요양종결후근로단절자여부 20%, 화재피해여부 33.33%, 피부양의무자장기요양여부 37.23% 순으로 낮게 나왔다. 비대상 기준 단전여부와 단가스 여부가 0건으로 나왔으며, 금융연체 4.46%, 피부양장기요양여부 6.16%, 시설입퇴소여부 6.67% 순으로 낮게 나왔다. 마지막 조치완료의 경우 기저귀조제분유지원대상자여부 20.55%, 금융연체여부 26.54%, 고용보험비대상여부 26.66%, 일용근로자여부 29.39%, 전의료체납여부 31.65% 순으로 낮게 나왔다.

제4절 가구 수에 따른 대상자 속성분석

<표 3-25> 지역별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	18425	6597	35.8046	1112	6.0353	10716	58.1601
서울	2988	1480	49.5315	142	4.7523	1366	45.7162
부산	1705	706	41.4076	144	8.4457	855	50.1466
대구	1077	269	24.9768	28	2.5998	780	72.4234
인천	1074	238	22.1601	65	6.0521	771	71.7877
광주	721	331	45.9085	82	11.3731	308	42.7184
대전	582	276	47.4227	21	3.6082	285	48.9691
울산	375	85	22.6667	0	0.0000	290	77.3333
세종	35	18	51.4286	0	0.0000	17	48.5714
경기	2531	807	31.8846	163	6.4401	1561	61.6752
강원	607	243	40.0329	121	19.9341	243	40.0329
충북	649	264	40.6780	43	6.6256	342	52.6965
충남	847	236	27.8630	26	3.0697	585	69.0673
전북	967	448	46.3289	45	4.6536	474	49.0176
전남	1514	526	34.7424	52	3.4346	936	61.8230
경북	1134	281	24.7795	74	6.5256	779	68.6949
경남	1390	321	23.0935	99	7.1223	970	69.7842
제주	229	68	29.6943	7	3.0568	154	67.2489
결측	0	0		0		0	

<표 3-25>는 지역별 선정 대상자를 기준으로 1인가구 대상자에 한정하여 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 세종특별자치시 51.43%, 서울특별시 49.53%, 대전광역시 47.42%, 전라북도 46.33%, 광주광역시 45.91% 순으로 높게 나타났다. 비대상 기준으로 강원도 19.93%, 광주광역시 11.37%, 부산광역시 8.45%, 경상남도 7.12%, 충청북도 6.63% 순으로 높게 나타나는 것을 알 수 있었다. 그리고 조치완료의 경우 울산광역시 77.33%, 대구광역시 72.42%, 인천광역시 71.79%, 경상남도 69.78%, 충청남도 69.78% 순으로 높게 나타났다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 인천광역시 22.16%, 울산광역시 22.67%, 경상남도 23.09%, 경상북도 24.78%, 대구광역시 24.97% 순으로 낮게 나타났다. 비대상 기준으로 보면 울산광역시와 세종특별자치도는 0건이며, 대구광역시 2.6%, 제주특별자치도 3.06%, 충청남도 3.07% 순으로 낮게 나타났다. 조치완료 기준으로 보면 강원도 40.03%, 광주광역시 42.72%, 서울특별시 45.72%, 세종특별자치시 48.57%, 대전광역시 48.97% 순으로 낮게 나타났다.

<표 3-26> 지역별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	17268	7411	42.9175	933	5.4031	8924	51.6794
서울	2681	1775	66.2066	124	4.6251	782	29.1682
부산	1232	635	51.5422	119	9.6591	478	38.7987
대구	751	157	20.9055	28	3.7284	566	75.3662
인천	734	215	29.2916	31	4.2234	488	66.4850
광주	432	236	54.6296	54	12.5000	142	32.8704
대전	448	284	63.3929	11	2.4554	153	34.1518
울산	272	62	22.7941	0	0.0000	210	77.2059
세종	45	29	64.4444	4	8.8889	12	26.6667
경기	2799	1049	37.4777	177	6.3237	1573	56.1986
강원	536	317	59.1418	85	15.8582	134	25.0000
충북	677	334	49.3353	25	3.6928	318	46.9719
충남	906	234	25.8278	16	1.7660	656	72.4062
전북	1049	537	51.1916	40	3.8132	472	44.9952
전남	1781	707	39.6968	53	2.9759	1021	57.3273
경북	1159	385	33.2183	57	4.9180	717	61.8637
경남	1604	402	25.0623	102	6.3591	1100	68.5786
제주	162	53	32.7160	7	4.3210	102	62.9630
결측	0	0		0		0	

<표 3-26>은 지역별 선정 대상자를 기준으로 1인가구 대상자에 한정하여 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 서울특별시 66.21%, 세종

특별자치시 64.44%, 대전광역시 63.39%, 강원도 59.14%, 광주광역시 54.63% 순으로 높게 나타났다. 비대상 기준으로 강원도 15.86%, 광주광역시 12.5%, 부산광역시 9.66%, 세종특별자치시 8.89%, 경상남도 6.36% 순으로 높게 나타나는 것을 알 수 있었다. 그리고 조치완료의 경우 울산광역시 77.21%, 대구광역시 75.37%, 충청남도 72.41%, 경상남도 68.58%, 인천광역시 66.49% 순으로 높게 나타났다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 대구광역시 20.91%, 울산광역시 22.79%, 경상남도 25.06%, 충청남도 25.83%, 인천광역시 29.29% 순으로 낮게 나타났다. 비대상 기준으로 보면 울산광역시의 경우 0건이며, 충청남도 1.77%, 대전광역시 2.46%, 충청남도 3.07% 순으로 낮게 나타났다. 조치완료 기준으로 보면 강원도 25%, 세종특별자치시 26.67%, 서울특별시 29.17%, 광주광역시 32.87%, 대전광역시 34.15% 순으로 낮게 나타났다.

<표 3-27> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	18468	6615	35.8187	1114	6.03206	10739	58.1492
영유	0	0		0		0	
아청	5	1	20	0	0	4	80
청년	1744	1026	58.8303	120	6.88073	598	34.289
장중	6477	2721	42.0102	450	6.94766	3306	51.0422
노년	10199	2849	27.9341	542	5.31425	6808	66.7517
결측	43	18	41.8605	2	4.65116	23	53.4884

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

1인가구의 연령대별 복지사각지대 선정 대상자의 2019년 1차 완료 결과 현황은 <표 3-27>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년기층이 제일 높은 단순상담 비율을 보여주었다. 반면에 아동·청소년층의 경우 20%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 아동·청소년기층 다음으로 노년기층과 장·중년기층이 27.93%, 42.01%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 장·중년층이 비대상의 비율이 가장 높으며 비슷하게 청년기층이 6.88%로 다음으로 높은 것을 알 수 있다. 반면 아동·청소년기층의 경우 비대상의 비율이 가장 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 아동·청소년층이 80%, 노년층이 66.75% 순으로 높게 나타났다. 그리고 청년기층의 경우 34.29%로 제일 낮은 비율을 보여주었다.

<표 3-28> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	18699	8477	45.3340	972	5.1981	9250	49.4679
영유	0	0		0		0	
아청	4	1	25.0000	1	25.0000	2	50.0000
청년	2777	1898	68.3471	149	5.3655	730	26.2874
장중	4909	2418	49.2565	325	6.6205	2166	44.1230
노년	11009	4160	37.7873	497	4.5145	6352	57.6982

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동·청소년기, 청년: 청년기, 장중: 장·중년기, 노년: 노년기

1인가구의 연령대별 복지사각지대 선정 대상자의 2019년 2차 완료 결과 현황은 <표 3-28>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 아동청소년층의 경우 25%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 노년층 다음으로 장중년층이 37.79%, 49.26%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 아동청소년층이 비대상의 비율이 25%로 가장 높은 것을 알 수 있다. 반면 노년층의 경우 4.51%로 비대상의 비율이 비슷하게 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년층이 57.69%, 아동청소년층이 50% 순으로 높게 나타났다. 그리고 청년층의 경우 26.28%로 제일 낮은 비율을 보여주었다.

<표 3-29> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	18468	6615	35.81871	1114	6.032055	10739	58.14923
1개	14558	5171	35.51999	949	6.518753	8438	57.96126
2개	3378	1254	37.12256	141	4.174067	1983	58.70337
3개	488	171	35.04098	21	4.303279	296	60.65574
4개	41	17	41.46341	3	7.317073	21	51.21951
5개	3	2	66.66667	0	0	1	33.33333
6개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-29>의 경우 1인가구에서 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 5개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 4개를 가진 대상이 41.46%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 1개, 3개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 4개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 1개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 3개를 가진 대상이 60.65%로 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 58.7%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 5개를 가진 대상이 33.33%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-30> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	18699	8477	45.33398	972	5.198139	10739	57.43088
1개	14706	6535	44.43764	798	5.426357	7373	50.136
2개	3543	1725	48.68755	152	4.29015	1666	47.0223
3개	398	194	48.74372	19	4.773869	185	46.48241
4개	50	21	42	3	6	26	52
5개	2	2	100	0	0	0	0
6개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-30>의 경우 1인 가구에서 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 5개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 48.74%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 1개, 4개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 3개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 1개를 가진 대상이 50.14%로 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 47.02%로 두 번째로 높은 것을 알 수 있다. 그리고 0건을 제외하고 변수 3개를 가진 대상이 46.48%로 제일 조치완료율이 낮은 것을 알 수 있다.

〈표 3-31〉 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	23	9	39.1304	0	0.0000	14	60.8696
단수	2	2	100.0000	0	0.0000	0	0.0000
단가	53	24	45.2830	0	0.0000	29	54.7170
전기	250	116	46.4000	11	4.4000	123	49.2000
연금	41	31	75.6098	1	2.4390	9	21.9512
건보	976	442	45.2869	34	3.4836	500	51.2295
화재	2	0	0.0000	0	0.0000	2	100.0000
경감	0	0		0		0	
장기	4974	1590	31.9662	188	3.7797	3196	64.2541
전세	646	250	38.6997	29	4.4892	367	56.8111
월세	13489	4927	36.5261	921	6.8278	7641	56.6462
개별	0	0		0		0	
실직	108	39	36.1111	7	6.4815	62	57.4074
비대	404	121	29.9505	18	4.4554	265	65.5941
집중	75	13	17.3333	2	2.6667	60	80.0000
분유	0	0		0		0	
청각	0	0		0		0	
예방	14	5	35.7143	0	0.0000	9	64.2857
시도	6	2	33.3333	0	0.0000	4	66.6667
위기	0	0		0		0	
범죄	0	0		0		0	
시설	65	25	38.4615	3	4.6154	37	56.9231
기초	1000	311	31.1000	58	5.8000	631	63.1000
공공	23	13	56.5217	0	0.0000	10	43.4783
산재	15	5	33.3333	0	0.0000	10	66.6667
재난	0	0		0		0	
금융	273	136	49.8168	15	5.4945	122	44.6886
의료	302	128	42.3841	10	3.3113	164	54.3046
일용	216	81	37.5000	9	4.1667	126	58.3333
영양	0	0		0		0	

<표 3-31>은 1인 가구 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단수여부 100%, 국민연금체납여부 75.61%, 공공임대주택체납여부 56.52%, 금융연체대상자여부 49.82%, 전기료체납여부 46.4% 순으로 높게 나왔다. 비대상의 경우 월세금액 기준이하가구여부의 경우 6.83%로 제일 높게 나왔다. 그리고 고용보험실직사유대상여부 6.48%, 기초생활긴급지원수급탈락여부 5.8%, 금융연체대상자여부 5.49%, 시설입퇴소여부 4.62% 순으로 높게 나왔다. 조치완료의 경우 화재피해여부의 경우 2건으로 2건 모두 조치완료가 되었다. 그리고 방문건강집중관리군여부 80%, 자살시도대상자여부 66.67%, 산재요양종결후근로단절자여부 66.67%, 고용보험비대상여부 65.59% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 화재피해여부의 경우 0건이며, 방문건강집중관리군여부 17.33%, 고용보험비대상여부 29.95%, 기초생활긴급지원수급탈락여부 31.1% 순으로 낮게 나왔다. 비대상 기준 화재피해여부, 자살시도대상자여부, 산재요양종결후근로단절자여부, 자살예방관리대상자여부, 단전여부, 단가스여부, 공공임대주택체납여부, 단수여부가 0건으로 나왔으며, 국민연금체납여부 2.44%, 방문건강집중관리군여부 2.67%, 의료비용과다지출가구여부 3.31% 순으로 낮게 나왔다. 마지막 조치완료의 경우 국민연금체납여부 21.95%, 공공임대주택체납여부 43.48%, 금융연체대상자여부 44.68%, 전기료체납여부 49.2%, 건강보험료체납여부 51.23% 순으로 낮게 나왔다.

〈표 3-32〉 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	23	10	43.4783	2	8.6957	11	47.8261
단수	0	0		0		0	
단가	31	22	70.9677	0	0.0000	9	29.0323
전기	217	129	59.4470	6	2.7650	82	37.7880
연금	106	64	60.3774	6	5.6604	36	33.9623
건보	1832	995	54.3122	56	3.0568	781	42.6310
화재	5	1	20.0000	0	0.0000	4	80.0000
경감	0	0		0		0	
장기	5460	1994	36.5201	145	2.6557	3321	60.8242
전세	1394	752	53.9455	54	3.8737	588	42.1808
월세	10191	4480	43.9604	740	7.2613	4971	48.7783
개별	1	0	0.0000	0	0.0000	1	100.0000
실직	102	52	50.9804	2	1.9608	48	47.0588
비대	432	179	41.4352	23	5.3241	230	53.2407
집중	29	5	17.2414	5	17.2414	19	65.5172
분유	0	0		0		0	
청각	0	0		0		0	
예방	5	1	20.0000	0	0.0000	4	80.0000
시도	3	1	33.3333	0	0.0000	2	66.6667
위기	0	0		0		0	
범죄	0	0		0		0	
시설	251	113	45.0199	11	4.3825	127	50.5976
기초	660	240	36.3636	34	5.1515	386	58.4848
공공	26	10	38.4615	0	0.0000	16	61.5385
산재	5	1	20.0000	0	0.0000	4	80.0000
재난	0	0		0		0	
금융	1840	1299	70.5978	52	2.8261	489	26.5761
의료	178	80	44.9438	14	7.8652	84	47.1910
일용	405	233	57.5309	21	5.1852	151	37.2840
영양	0	0		0		0	

<표 3-32>는 1인 가구 대상자에서 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단가스여부 70.97%, 금융연체대상자여부 70.6%, 국민연금체납여부 60.37%, 전기료체납여부 59.45%, 일용근로자여부 57.53% 순으로 높게 나왔다. 비대상의 경우 방문건강집중관리군여부의 경우 17.24%로 제일 높게 나왔다. 그리고 단전여부 8.7%, 의료비용과다지출가구여부 7.87%, 월세금액기준이하가구여부 7.26%, 국민연금체납여부 5.66% 순으로 높게 나왔다. 조치완료의 경우 방문건강집중관리군여부의 경우 65.52%로 제일 높은 조치완료를 보여주었다. 그리고 단전여부 47.83%, 의료비용과다지출가구여부 47.19%, 월세금액기준이하가구여부 48.78%, 국민연금체납여부 33.96% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 고용보험개별연장급여대상자여부의 경우 0건이며, 방문건강집중관리군여부 17.24%, 화재피해여부 20%, 자살예방관리대상자여부 20% 순으로 낮게 나왔다. 비대상 기준 고용보험개별연장급여대상자여부, 화재피해여부, 자살예방관리대상자여부, 산재요양종결후근로단절자여부, 자살시도대상자여부, 공공임대주택체납여부, 단가스여부가 0건으로 나왔으며, 고용보험실직사유대상여부 1.96%, 피부양의무자장기요양여부 2.66%, 전기료체납여부 2.77% 순으로 낮게 나왔다. 마지막 조치완료의 경우 금융연체대상자여부 26.58%, 단가스여부 29.03%, 국민연금체납여부 33.96%, 일용근로자여부 37.28%, 전기료체납여부 37.78% 순으로 낮게 나왔다.

다음은 다인가구에 대한 대상자 속성 분석을 보여준다. 다인 가구는 가구 수가 1이 아닌 1이외의 대상자를 모두 다인가구라고 본 연구에서는 부른다.

〈표 3-33〉 지역별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	35154	13490	38.3740	3983	11.3301	17681	50.2958
서울	1596	919	57.5815	126	7.8947	551	34.5238
부산	781	392	50.1921	92	11.7798	297	38.0282
대구	499	129	25.8517	22	4.4088	348	69.7395
인천	770	180	23.3766	65	8.4416	525	68.1818
광주	406	181	44.5813	70	17.2414	155	38.1773
대전	328	151	46.0366	29	8.8415	148	45.1220
울산	164	61	37.1951	2	1.2195	101	61.5854
세종	155	88	56.7742	20	12.9032	47	30.3226
경기	10481	4141	39.5096	1223	11.6687	5117	48.8217
강원	1862	773	41.5145	651	34.9624	438	23.5231
충북	1772	914	51.5801	144	8.1264	714	40.2935
충남	2626	910	34.6535	170	6.4737	1546	58.8728
전북	2780	1449	52.1223	299	10.7554	1032	37.1223
전남	3245	967	29.7997	229	7.0570	2049	63.1433
경북	3402	947	27.8366	350	10.2881	2105	61.8754
경남	3594	1037	28.8536	451	12.5487	2106	58.5977
제주	693	251	36.2193	40	5.7720	402	58.0087

〈표 3-33〉은 지역별 선정 대상자를 기준으로 다인 가구 대상자에 한정하여 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 서울특별시 57.58%, 세종특별자치도 56.77%, 전라북도 52.12%, 충청북도 51.58%, 부산광역시 50.19% 순으로 높게 나타났다. 비대상 기준으로 강원도 34.96%, 광주광역시 17.24%, 세종특별자치시 12.9%, 경상남도 12.5%, 부산광역시 11.78% 순으로 높게 나타나는 것을 알 수 있었다. 그리고 조치완료의 경우 대구광역시 69.74%, 인천광역시 68.18%, 전라남도 63.14%, 경상북도 61.87%, 울산광역시 61.58% 순으로 높게 나타났다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 인천광역시 23.37%, 대구광역시

25.87%, 경상북도 27.84%, 경상남도 28.85%, 전라남도 29.8% 순으로 낮게 나타났다. 비대상 기준으로 보면 울산광역시 1.22%, 대구광역시 4.4%, 제주특별자치도 5.77%, 충청남도 6.47% 순으로 낮게 나타났다. 조치완료 기준으로 보면 강원도 23.52%, 세종특별자치시 30.32%, 서울특별시 34.52%, 전라북도 37.12%, 부산광역시 38.03% 순으로 낮게 나타났다.

<표 3-34> 지역별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	35158	16358	46.5271	3666	10.4272	15134	43.0457
서울	3601	2744	76.2011	271	7.5257	586	16.2733
부산	1780	962	54.0449	296	16.6292	522	29.3258
대구	1313	312	23.7624	107	8.1493	894	68.0883
인천	1421	382	26.8825	144	10.1337	895	62.9838
광주	929	535	57.5888	179	19.2680	215	23.1432
대전	807	520	64.4362	57	7.0632	230	28.5006
울산	478	128	26.7782	3	0.6276	347	72.5941
세종	129	82	63.5659	9	6.9767	38	29.4574
경기	9036	3935	43.5480	924	10.2258	4177	46.2262
강원	1304	693	53.1442	430	32.9755	181	13.8804
충북	1459	902	61.8232	123	8.4304	434	29.7464
충남	1824	708	38.8158	81	4.4408	1035	56.7434
전북	2471	1572	63.6180	192	7.7701	707	28.6119
전남	2431	929	38.2147	166	6.8285	1336	54.9568
경북	2413	694	28.7609	258	10.6921	1461	60.5470
경남	3002	915	30.4797	386	12.8581	1701	56.6622
제주	760	345	45.3947	40	5.2632	375	49.3421

<표 3-34>는 지역별 선정 대상자를 기준으로 다인 가구 대상자에 한정하여 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 제일 높은 비율을 보여주는 지역을 보면 서울특별시 76.2%, 대전광역시 64.43%, 전라북도 63.62%, 세종특별자치시 63.57%, 충청북도 61.82% 순으로 높게 나타났다. 비대상 기준으로 강원도 32.97%,

광주광역시 19.27%, 부산광역시 16.63%, 경상남도 12.86%, 경상북도 10.69% 순으로 높게 나타나는 것을 알 수 있었다. 그리고 조치완료의 경우 울산광역시 72.59%, 대구광역시 68.09%, 인천광역시 62.98%, 경상북도 60.55%, 충청남도 56.74% 순으로 높게 나타났다.

반대로 낮은 순으로 각 지역별 선정 대상자를 분석하면 다음과 같이 볼 수 있다. 단순상담 기준으로 대구광역시 23.76%, 울산광역시 26.78%, 인천광역시 26.88%, 경상북도 28.76%, 경상남도 30.48% 순으로 낮게 나타났다. 비대상 기준으로 보면 울산광역시 0.63%, 충청남도 4.44%, 제주특별자치도 5.26%, 전라남도 6.82% 순으로 낮게 나타났다. 조치완료 기준으로 보면 강원도 13.88%, 서울특별시 16.27%, 광주광역시 23.14%, 대전광역시 28.5%, 전라북도 28.61% 순으로 낮게 나타났다.

<표 3-35> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	37968	15171	39.9573	4285	11.2858	18512	48.7568
영유	1384	338	24.4220	166	11.9942	880	63.5838
아청	2458	1039	42.2701	281	11.4321	1138	46.2978
청년	5658	2906	51.3609	591	10.4454	2161	38.1937
장중	8552	3700	43.2647	945	11.0500	3907	45.6852
노년	17130	5528	32.2709	2001	11.6813	9601	56.0479
결측	2786	1660	59.5836	301	10.8040	825	29.6123

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

다인 가구의 연령대별 복지사각지대 선정 대상자의 2019년 1차 완료결과 현황은 <표 3-35>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아층의 경우 24.42%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 영유아층 다음으로 노년층과 아동청소년층이 32.27%, 42.27%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아층이 비대상의 비율이 가장 높으며 비슷하게 노년기층이 11.68%로 다음으로 높은 것을 알 수 있다. 반면 장·중년기층과 청년기층의 경우 비대상의 비율이 비슷하게 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 영유아기층이 63.58%, 노년기층이 56.05% 순으로 높게 나타났다. 그리고 청년기층의 경우 38.19%로 제일 낮은 비율을 보여주었다.

<표 3-36> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	35266	16441	46.6200	3668	10.4010	15157	42.9791
영유	1721	656	38.1174	252	14.6426	813	47.2400
아청	3980	1766	44.3719	491	12.3367	1723	43.2915
청년	8658	4651	53.7191	908	10.4874	3099	35.7935
장·중	12517	6105	48.7737	1198	9.5710	5214	41.6553
노년	8389	3262	38.8843	819	9.7628	4308	51.3530
결측	1	1	100	0	0	0	0

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장·중: 장·중년기, 노년: 노년기

다인 가구의 연령대별 복지사각지대 선정 대상자의 2019년 2차 완료결과 현황은 <표 3-36>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년기층이 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아층의 경우 38.12%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 영유아층 다음으로 노년층과 아동청소년층이 38.88%, 44.37%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아층이 비대상의 비율이 가장 높으며 비슷하게 아동청소년층이 12.34%로 다음으로 높은 것을 알 수 있다. 반면 노년기층과 장중년층의 경우 비대상의 비율이 비슷하게 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년기층이 51.35%, 영유아층이 47.24% 순으로 높게 나타났다. 그리고 청년기층의 경우 35.79%로 제일 낮은 비율을 보여주었다.

<표 3-37> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순 상담		비대상		조치 완료	
전체	37968	15171	39.9573	4285	11.2858	18512	48.7568
1개	30920	12516	40.4787	3677	11.8920	14727	47.6294
2개	5690	2164	38.0316	528	9.2794	2998	52.6889
3개	1154	417	36.1352	67	5.8059	670	58.0589
4개	183	64	34.9727	11	6.0109	108	59.0164
5개	19	10	52.6316	2	10.5263	7	36.8421
6개	2	0	0	0	0	2	100

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-37>의 경우 다인 가구에서 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 5개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 1개를 가진 대상이 40.47%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 3개, 4개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 5개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 6개를 가진 대상이 2건 모두 처리가 되어 100%로 제일 높다. 그리고 변수 4개를 가진 대상이 59.02%로 두 번째로 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 59.02%로 세 번째로 높은 것을 알 수 있다. 그리고 변수 5개를 가진 대상이 36.84%로 제일 조치완료를 이 낮은 것을 알 수 있다.

<표 3-38> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	35266	16441	46.6200	3668	10.4010	15157	42.9791
1개	23795	11227	47.1822	2651	11.1410	9917	41.6768
2개	8806	4011	45.5485	816	9.2664	3979	45.1851
3개	2234	1014	45.3894	164	7.3411	1056	47.2695
4개	396	174	43.9394	36	9.0909	186	46.9697
5개	34	14	41.1765	1	2.9412	19	55.8824
6개	1	1	100	0	0	0	0

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-38>의 경우 다인 가구에서 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 6개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 1개를 가진 대상이 47.18%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 3개, 4개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 2개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 5개를 가진 대상이 55.88%로 제일 높다. 그리고 변수 3개를 가진 대상이 47.27%로 두 번째로 높은 것을 알 수 있다. 다음으로 변수 4개를 가진 대상이 46.97%로 세 번째로 높은 것을 알 수 있다. 그리고 변수 6개를 가진 대상이 0%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-39> 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	48	18	37.5000	4	8.3333	26	54.1667
단수	16	8	50.0000	1	6.2500	7	43.7500
단가	156	79	50.6410	6	3.8462	71	45.5128
전기	597	281	47.0687	47	7.8727	269	45.0586
연금	298	142	47.6510	26	8.7248	130	43.6242
건보	6062	3090	50.9733	406	6.6975	2566	42.3293
화재	13	5	38.4615	0	0.0000	8	61.5385
경감	0	0		0		0	
장기	21377	7450	34.8505	2600	12.1626	11327	52.9869
전세	3755	1583	42.1571	432	11.5047	1740	46.3382
월세	7341	2960	40.3215	723	9.8488	3658	49.8297
개별	84	22	26.1905	5	5.9524	57	67.8571
실직	82	30	36.5854	8	9.7561	44	53.6585
비대	669	334	49.9253	87	13.0045	248	37.0703
집중	120	38	31.6667	11	9.1667	71	59.1667
분유	778	186	23.9075	97	12.4679	495	63.6247
청각	0	0		0		0	
예방	11	1	9.0909	1	9.0909	9	81.8182
시도	13	8	61.5385	0	0.0000	5	38.4615
위기	459	199	43.3551	48	10.4575	212	46.1874
범죄	0	0		0		0	
시설	62	24	38.7097	4	6.4516	34	54.8387
기초	1659	508	30.6209	150	9.0416	1001	60.3376
공공	193	67	34.7150	14	7.2539	112	58.0311
산재	13	6	46.1538	1	7.6923	6	46.1538
재난	0	0		0		0	
금융	949	532	56.0590	56	5.9009	361	38.0400
의료	418	157	37.5598	53	12.6794	208	49.7608
일용	1016	568	55.9055	113	11.1220	335	32.9724
영양	412	105	25.4854	95	23.0583	212	51.4563

<표 3-39>은 다인 가구 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 자살시도대상자여부 61.54%, 금융연체대상자여부 56.06%, 일용근로자여부 55.91%, 건강보험료체납여부 50.97%, 단가스여부 50.64% 순으로 높게 나왔다. 비대상의 경우 영양플러스미지원가구의 경우 23.06%로 제일 높게 나왔다. 그리고 고용보험비대상여부 13%, 의료비용과다지출가구여부 12.68%, 기저귀조제분유지원대상자여부 12.46%, 피부양 의무자장기요양여부 12.16% 순으로 높게 나왔다. 조치완료의 경우 자살예방관리대상자여부 81.82%, 고용보험개별연장급여대상자여부 67.86%, 기저귀조제분유지원대상자여부 63.62% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 자살예방관리대상자여부의 경우 9.09%, 기저귀조제분유지원대상자여부 23.91%, 영양플러스미지원가구 25.49%, 고용보험개별연장급여대상자여부 26.19% 순으로 낮게 나왔다. 비대상 기준 화재피해여부, 자살시도대상자여부가 0건으로 나왔으며, 단가스여부 3.85%, 금융연체대상자여부 5.9%, 고용보험개별연장급여대상자여부 5.95% 순으로 낮게 나왔다. 마지막 조치완료의 경우 일용근로자여부 32.97%, 고용보험비대상여부 37.07%, 금융연체대상자여부 38.04%, 자살시도대상자여부 38.46%, 건강보험료체납여부 42.33% 순으로 낮게 나왔다.

<표 3-40> 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	74	37	50.0000	8	10.8108	29	39.1892
단수	2	1	50.0000	0	0.0000	1	50.0000
단가	143	68	47.5524	8	5.5944	67	46.8531
전기	952	485	50.9454	91	9.5588	376	39.4958
연금	691	328	47.4674	55	7.9595	308	44.5731
건보	9442	4774	50.5613	726	7.6890	3942	41.7496
화재	21	11	52.3810	3	14.2857	7	33.3333
경감	0	0		0		0	
장기	6587	2599	39.4565	591	8.9722	3397	51.5713
전세	4803	2230	46.4293	567	11.8051	2006	41.7656
월세	15468	7079	45.7655	1625	10.5056	6764	43.7290
개별	67	25	37.3134	6	8.9552	36	53.7313
실직	421	216	51.3064	48	11.4014	157	37.2922
비대	1091	576	52.7956	139	12.7406	376	34.4638
집중	39	16	41.0256	6	15.3846	17	43.5897
분유	1005	358	35.6219	104	10.3483	543	54.0299
청각	0	0		0		0	
예방	4	0	0.0000	0	0.0000	4	100.0000
시도	15	7	46.6667	2	13.3333	6	40.0000
위기	942	408	43.3121	119	12.6327	415	44.0552
범죄	0	0		0		0	
시설	165	65	39.3939	10	6.0606	90	54.5455
기초	2891	1213	41.9578	313	10.8267	1365	47.2155
공공	978	444	45.3988	69	7.0552	465	47.5460
산재	26	12	46.1538	4	15.3846	10	38.4615
재난	0	0		0		0	
금융	1627	871	53.5341	109	6.6994	647	39.7664
의료	554	253	45.6679	70	12.6354	231	41.6968
일용	1474	824	55.9023	154	10.4478	496	33.6499
영양	387	163	42.1189	97	25.0646	127	32.8165

<표 3-40>은 다인 가구 대상자에서 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 일용근로자여부 55.9%, 금융연체대상자여부 53.53%, 고용보험비대상여부 52.38%, 화재피해여부 52.38% 순으로 높게 나왔다. 비대상의 경우 영양플러스미지원가구의 경우 25.06%로 제일 높게 나왔다. 그리고 산재요양종결후근로단절자여부 15.38%, 방문건강집중관리군여부 15.38%, 화재피해여부 14.28%, 자살시도대상자여부 13.33% 순으로 높게 나왔다. 조치완료의 경우 자살예방관리대상자여부 100%, 시설입퇴소여부 54.55%, 기저귀조제분유지원대상자여부 54.03% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준 자살예방관리대상자여부의 경우 0%, 기저귀조제분유지원대상자여부 35.62%, 고용보험개별연장급여대상자여부 37.31% 순으로 낮게 나왔다. 비대상 기준 자살예방관리대상자여부, 단수도여부가 0건으로 나왔으며, 단가스여부 5.59%, 시설입퇴소여부 6.06%, 금융연체대상자여부 6.7% 순으로 낮게 나왔다. 마지막 조치완료의 경우 영양플러스미지원가구 32.81%, 화재피해여부 33.33%, 일용근로자여부 33.65%, 고용보험비대상여부 34.46%, 고용보험실직사유대상자여부 37.29% 순으로 낮게 나왔다.

제5절 지역별 대상자 속성분석

지역별 대상자의 속성분석을 위해 본 절에서는 서울특별시, 경기도, 전라북도, 제주특별자치도 대상자의 속성을 분석하고자 한다.

<서울특별시>

우선적으로 서울특별시에서는 대상자를 기준으로 속성 분석을 수행한다.

<표 3-41> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	4584	2399	52.3342	268	5.8464	1917	41.8194
영유	95	32	33.6842	6	6.3158	57	60.0000
아청	176	115	65.3409	26	14.7727	35	19.8864
청년	799	592	74.0926	56	7.0088	151	18.8986
장중	1747	1040	59.5306	103	5.8958	604	34.5736
노년	1767	620	35.0877	77	4.3577	1070	60.5546

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

서울특별시의 연령대별 복지사각지대 선정 대상자의 2019년 1차 완료결과 현황은 <표 3-41>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아기층의 경우 33.68%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 영유아기층 다음으로 노년층과 장중년층이 35.09%, 59.53%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 아동청소년층이 비대상의 비율이 가장 높으며

비슷하게 청년층이 7%로 다음으로 높은 것을 알 수 있다. 반면 노년기층의 경우 비대상의 비율이 비슷하게 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년기층이 60.55%, 영유아기층이 60% 순으로 높게 나타났다. 그리고 청년기층의 경우 18.9%로 제일 낮은 비율을 보여주었다.

<표 3-42> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	6282	4519	71.9357	395	6.2878	1368	21.7765
영유	161	99	61.4907	24	14.9068	38	23.6025
아청	268	200	74.6269	23	8.5821	45	16.7910
청년	1421	1197	84.2365	101	7.1077	123	8.6559
장중	2055	1549	75.3771	135	6.5693	371	18.0535
노년	2377	1474	62.0109	112	4.7118	791	33.2772

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

서울특별시의 연령대별 복지사각지대 선정 대상자의 2019년 2차 완료결과 현황은 <표 3-42>와 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아기층의 경우 61.49%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 영유아기층 다음으로 노년층과 아동청소년층이 62.01%, 74.63%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아층이 비대상의 비율이 가장 높으며 비슷하게 아동청소년층이 8.58%로 다음으로 높은 것을 알 수

있다. 반면 노년기층의 경우 4.71%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년기층이 33.27%, 영유아기층이 23.6% 순으로 높게 나타났다. 그리고 청년기층의 경우 8.66%로 제일 낮은 비율을 보여주었다.

<표 3-43> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	4584	2399	52.3342	268	5.8464	1917	41.8194
1개	2762	1494	54.0912	165	5.9739	1103	39.9348
2개	1483	734	49.4943	85	5.7316	664	44.7741
3개	295	149	50.5085	15	5.0847	131	44.4068
4개	40	19	47.5000	3	7.5000	18	45.0000
5개	4	3	75.0000	0	0.0000	1	25.0000
6개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-43>의 경우 서울특별시 데이터에서 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 5개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 4개를 가진 대상이 54.09%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 3개, 4개, 2개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 4개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 1개를 가진 대상이

두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 4개를 가진 대상이 45%로 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 44.77%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 6개를 가진 대상이 25%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-44> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	6282	4519	71.9357	395	6.2878	1368	21.7765
1개	4101	3001	73.1773	259	6.3155	841	20.5072
2개	1710	1190	69.5906	109	6.3743	411	24.0351
3개	385	269	69.8701	18	4.6753	98	25.4545
4개	79	55	69.6203	9	11.3924	15	18.9873
5개	7	4	57.1429	0	0.0000	3	42.8571
6개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-44>의 경우 서울특별시 데이터에서 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 1개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 69.87%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 4개, 2개, 5개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 4개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 2개를 가진 대상이

두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 5개를 가진 대상이 42.86%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 25.45%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 4개를 가진 대상이 18.98%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-45> 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상당		비대상		조치완료	
단전	13	8	61.5385	0	0.0000	5	38.4615
단수	2	0	0.0000	0	0.0000	2	100.0000
단가	87	56	64.3678	2	2.2989	29	33.3333
전기	177	118	66.6667	10	5.6497	49	27.6836
연금	43	30	69.7674	4	9.3023	9	20.9302
건보	808	563	69.6782	36	4.4554	209	25.8663
화재	1	0	0.0000	0	0.0000	1	100.0000
경감	0	0		0		0	
장기	701	294	41.9401	26	3.7090	381	54.3509
전세	697	344	49.3544	60	8.6083	293	42.0373
월세	3136	1533	48.8839	166	5.2934	1437	45.8227
개별	16	6	37.5000	0	0.0000	10	62.5000
실직	36	18	50.0000	2	5.5556	16	44.4444
비대	166	76	45.7831	8	4.8193	82	49.3976
집중	41	11	26.8293	4	9.7561	26	63.4146
분유	95	35	36.8421	9	9.4737	51	53.6842
청각	0	0		0		0	
예방	1	1	100.0000	0	0.0000	0	0.0000
시도	5	3	60.0000	0	0.0000	2	40.0000
위기	41	24	58.5366	5	12.1951	12	29.2683
범죄	0	0		0		0	
시설	23	12	52.1739	1	4.3478	10	43.4783
기초	386	178	46.1140	35	9.0674	173	44.8187
공공	22	11	50.0000	0	0.0000	11	50.0000

구분		단순상담		비대상		조치완료	
산재	2	0	0.0000	0	0.0000	2	100.0000
재난	0	0		0		0	
금융	99	66	66.6667	4	4.0404	29	29.2929
의료	115	65	56.5217	13	11.3043	37	32.1739
일용	60	37	61.6667	4	6.6667	19	31.6667
영양	20	11	55.0000	3	15.0000	6	30.0000

<표 3-45>은 서울특별시 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 자살예방관리대상자여부 100%, 국민연금체납여부 69.77%, 건강보험료체납여부 69.78%, 전기료체납여부 66.67%, 금융연체대상자여부 66.67% 순으로 높게 나왔다. 비대상의 경우 영양플러스미지원가구의 경우 15%로 제일 높게 나왔다. 그리고 위기학생여부 12.2%, 의료비용과다지출가구여부 11.3%, 방문건강집중관리군여부 9.76%, 기저귀조제분유지원대상자여부 9.47% 순으로 높게 나왔다. 조치완료의 경우 단수여부, 화재피해여부, 산재요양종결후근로단절자여부 100%로 제일 높은 조치완료를 보여주었다. 그리고 방문건강집중관리군여부 63.41%, 고용보험개별연장급여대상자여부 62.5%, 피부양의무자장기요양여부 54.35%, 기저귀조제분유지원대상자여부 53.68% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 방문건강집중관리군여부 26.83%, 기저귀조제분유지원대상자여부 36.84%, 고용보험개별연장급여대상자여부 37.5% 순으로 낮게 나왔다. 비대상 기준 단수여부, 화재피해여부, 산재요양종결후근로단절자여부, 고용보험개별연장급여대상자여부, 공공임대주택여부, 자살시도대상자여부, 단전여부, 자살예방관리대상자여부가 0건으로 나왔으며, 단가스여부 2.3%, 피부양의무자장기요양여부 3.71%, 금융연체대상자

여부 4.04% 순으로 낮게 나왔다. 마지막 조치완료의 경우 자살예방관리대상자여부가 0건으로 나왔으며, 국민연금체납여부 20.93%, 건강보험료체납여부 25.87%, 전기료체납여부 27.68%, 위기학생여부 29.27%, 금융연체대상자여부 29.29% 순으로 낮게 나왔다.

<표 3-46> 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	12	11	91.6667	0	0.0000	1	8.3333
단수	0	0		0		0	
단가	58	46	79.3103	0	0.0000	12	20.6897
전기	290	222	76.5517	17	5.8621	51	17.5862
연금	86	68	79.0698	6	6.9767	12	13.9535
건보	1634	1293	79.1310	88	5.3856	253	15.4835
화재	1	1	100.0000	0	0.0000	0	0.0000
경감	0	0		0		0	
장기	440	269	61.1364	18	4.0909	153	34.7727
전세	1245	882	70.8434	90	7.2289	273	21.9277
월세	3545	2408	67.9267	205	5.7828	932	26.2906
개별	6	3	50.0000	0	0.0000	3	50.0000
실직	57	43	75.4386	3	5.2632	11	19.2982
비대	222	153	68.9189	14	6.3063	55	24.7748
집중	16	5	31.2500	4	25.0000	7	43.7500
분류	71	33	46.4789	11	15.4930	27	38.0282
청각	0	0		0		0	
예방	0	0		0		0	
시도	2	1	50.0000	0	0.0000	1	50.0000
위기	43	29	67.4419	7	16.2791	7	16.2791
범죄	0	0		0		0	
시설	44	25	56.8182	2	4.5455	17	38.6364
기초	438	275	62.7854	50	11.4155	113	25.7991
공공	60	43	71.6667	2	3.3333	15	25.0000
산재	2	2	100.0000	0	0.0000	0	0.0000
재난	0	0		0		0	
금융	306	256	83.6601	18	5.8824	32	10.4575
의료	103	82	79.6117	6	5.8252	15	14.5631
일용	278	228	82.0144	15	5.3957	35	12.5899
영양	68	50	73.5294	11	16.1765	7	10.2941

<표 3-46>은 서울특별시 대상자에서 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단전여부 91.67%, 일용근로자여부 82.01%, 의료비용과다지출가구여부 79.61%, 단가스여부 79.31%, 건보료체납여부 79.13% 순으로 높게 나왔다. 비대상의 경우 방문건강집중관리군여부 25%, 위기학생여부 16.28%, 영양플러스미지원가구 16.18%, 기저귀조제분유지원대상자 15.5%, 기초생활긴급지원수급탈락여부 11.41%순으로 높게 나왔다. 조치완료의 경우 고용보험개별연장급여대상자여부와 자살시도대상자여부 50%, 방문건강집중관리군여부 43.75%, 시설입퇴소여부 38.64%, 피부양 의무자 장기요양여부 34.77% 순으로 높게 나왔다

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단수도여부, 본인부담경감여부, 신생아청각선별검사대상자여부, 자살예방관리대상자여부, 재난피해자여부의 경우 0건의 나왔으며, 방문건강집중관리군여부 31.25%, 기저귀조제분유지원대상자여부 46.48% 순으로 낮게 나왔다. 비대상기준 단전여부, 단가스여부, 화재피해여부, 고용보험개별연장급여대상자여부, 범죄피해여부 등이 0건으로 나왔으며, 공공임대주택 체납여부 3.33% 피부양 의무자 장기요양여부 4.09%순으로 낮게 나왔다. 마지막 조치완료의 경우 화재피해여부와 산재요양종결후근로단절자여부가 0건으로 나왔으며, 단전여부 8.33%, 영양플러스미지원가구 10.29%, 금융연체대상자여부 10.46%, 일용근로자여부 12.59%순으로 낮게 나왔다.

<경기도>

다음은 경기도에 거주하는 대상자를 기준으로 속성분석을 수행하고자 한다.

<표 3-47> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	13012	4948	38.0264	1386	10.6517	6678	51.3219
영유	448	116	25.8929	58	12.9464	274	61.1607
아청	652	281	43.0982	55	8.4356	316	48.4663
청년	2163	1126	52.0573	197	9.1077	840	38.8350
장중	3634	1504	41.3869	366	10.0715	1764	48.5416
노년	6115	1921	31.4146	710	11.6108	3484	56.9747

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

경기도의 연령대별 복지사각지대 선정 대상자의 2019년 1차 완료 결과 현황은 <표 3-47>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아기층의 경우 25.89%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 영유아기층 다음으로 아동청소년층과 청년층이 43.1%, 52.06%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아기층이 비대상의 비율이 가장 높으며 비슷하게 노년층이 11.61%로 다음으로 높은 것을 알 수 있다. 반면 아동청소년층의 경우 8.44%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 영유아기층이 61.16%, 노년층이 56.97% 순으로 높게 나타났다. 그리고 청년기층의 경우 38.84%로 제일 낮은 비율을 보여주었다.

<표 3-48> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	11835	4984	42.1124	1101	9.3029	5750	48.5847
영유	449	163	36.3029	59	13.1403	227	50.5568
아청	959	384	40.0417	133	13.8686	442	46.0897
청년	2641	1327	50.2461	266	10.0719	1048	39.6819
장중	4407	1924	43.6578	392	8.8949	2091	47.4472
노년	3379	1186	35.0991	251	7.4282	1942	57.4726

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

경기도의 연령대별 복지사각지대 선정 대상자의 2019년 2차 완료 결과 현황은 <표 3-48>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 노년기층의 경우 35.1%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 노년기층 다음으로 장중년기층과 청년층이 43.66%, 50.25%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 아동청소년기층이 비대상의 비율이 가장 높으며 비슷하게 영유아기층이 13.14%로 다음으로 높은 것을 알 수 있다. 반면 노년기층의 경우 7.43%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년기층이 57.47%, 영유아기층이 50.56% 순으로 높게 나타났다. 그리고 청년기층의 경우 39.68%로 제일 낮은 비율을 보여주었다.

<표 3-49> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	13012	4948	38.0264	1386	10.6517	6678	51.3219
1개	10665	4126	38.6873	1233	11.5612	5306	49.7515
2개	1930	679	35.1813	129	6.6839	1122	58.1347
3개	372	126	33.8710	22	5.9140	224	60.2151
4개	41	16	39.0244	1	2.4390	24	58.5366
5개	3	1	33.3333	1	33.3333	1	33.3333
6개	1	0	0.0000	0	0.0000	1	100.0000

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-49>의 경우 경기도 데이터에서 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 4개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 1개를 가진 대상이 38.68%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 2개, 3개, 5개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 5개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 1개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 6개를 가진 대상이 100%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 60.22%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 5개를 가진 대상이 33.33%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-50> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	11835	4984	42.1124	1101	9.3029	5750	48.5847
1개	8453	3647	43.1444	823	9.7362	3983	47.1194
2개	2638	1055	39.9924	223	8.4534	1360	51.5542
3개	633	242	38.2306	45	7.1090	346	54.6603
4개	106	39	36.7925	10	9.4340	57	53.7736
5개	5	1	20.0000	0	0.0000	4	80.0000
6개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-50>의 경우 경기도 데이터에서 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 1개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 39.99%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 3개, 4개, 5개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 4개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 5개를 가진 대상이 80%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 54.66%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 1개를 가진 대상이 47.12%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-51> 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	13	6	46.1538	3	23.0769	4	30.7692
단수	0	0		0		0	
단가	15	10	66.6667	1	6.6667	4	26.6667
전기	177	73	41.2429	17	9.6045	87	49.1525
연금	100	47	47.0000	10	10.0000	43	43.0000
건보	2294	1097	47.8204	132	5.7541	1065	46.4255
화재	0	0		0		0	
경감	0	0		0		0	
장기	7045	2508	35.5997	890	12.6331	3647	51.7672
전세	874	297	33.9817	77	8.8101	500	57.2082
월세	3576	1264	35.3468	290	8.1096	2022	56.5436
개별	18	4	22.2222	0	0.0000	14	77.7778
실직	49	14	28.5714	6	12.2449	29	59.1837
비대	316	146	46.2025	41	12.9747	129	40.8228
집중	20	3	15.0000	0	0.0000	17	85.0000
분유	176	50	28.4091	28	15.9091	98	55.6818
청각	0	0		0		0	
예방	7	1	14.2857	1	14.2857	5	71.4286
시도	0	0		0		0	
위기	67	29	43.2836	5	7.4627	33	49.2537
범죄	0	0		0		0	
시설	32	12	37.5000	1	3.1250	19	59.3750
기초	508	137	26.9685	27	5.3150	344	67.7165
공공	64	19	29.6875	2	3.1250	43	67.1875
산재	7	3	42.8571	0	0.0000	4	57.1429
재난	0	0		0		0	
금융	210	107	50.9524	12	5.7143	91	43.3333
의료	100	36	36.0000	9	9.0000	55	55.0000
일용	137	62	45.2555	13	9.4891	62	45.2555
영양	21	6	28.5714	1	4.7619	14	66.6667

<표 3-51>은 경기도 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단가스여부 66.67%, 금융연체대상자여부 50.92%, 건강보험료체납여부 47.82%, 국민연금체납여부 47%, 고용보험비대상여부 46.2% 순으로 높게 나왔다. 비대상의 경우 단전여부의 경우 23.08%로 제일 높게 나왔다. 그리고 기저귀조제분유지원대상자여부 15.91%, 자살예방관리대상자여부 14.29%, 고용보험비대상여부 12.97%, 피부양 의무자장기요양여부 12.63% 순으로 높게 나왔다. 조치완료의 경우 방문건강집중관리군여부 85%, 고용보험개별연장급여대상자여부 77.78%, 자살예방관리대상자여부 71.43%, 기초생활긴급지원수급탈락여부 67.72% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 자살예방관리대상자여부 14.29%, 방문건강집중관리군여부 15%, 고용보험개별연장급여대상자여부 22.22% 순으로 낮게 나왔다. 비대상 기준 방문건강집중관리군여부, 고용보험개별연장급여대상자여부, 산재요양종결후근로단절자여부가 0건으로 나왔으며, 공공임대주택체납여부 3.13%, 시설입퇴소여부 3.13%, 영양플러스미지원가구여부 4.76% 순으로 낮게 나왔다. 마지막 조치완료의 경우 단가스여부가 26.67%로 나왔으며, 단전여부 30.77%, 고용보험비대상여부 40.82%, 국민연금체납여부 43%, 금융연체대상자여부 43.33%, 일용근로자여부 45.26% 순으로 낮게 나왔다.

<표 3-52> 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	17	9	52.9412	1	5.8824	7	41.1765
단수	0	0		0		0	
단가	25	4	16.0000	6	24.0000	15	60.0000
전기	277	132	47.6534	18	6.4982	127	45.8484
연금	211	95	45.0237	16	7.5829	100	47.3934
건보	2846	1245	43.7456	232	8.1518	1369	48.1026
화재	12	5	41.6667	1	8.3333	6	50.0000
경감	0	0		0		0	
장기	1934	729	37.6939	196	10.1344	1009	52.1717
전세	1616	684	42.3267	158	9.7772	774	47.8960
월세	5744	2266	39.4499	485	8.4436	2993	52.1065
개별	17	7	41.1765	1	5.8824	9	52.9412
실직	196	99	50.5102	23	11.7347	74	37.7551
비대	382	199	52.0942	38	9.9476	145	37.9581
집중	22	5	22.7273	4	18.1818	13	59.0909
분유	204	80	39.2157	21	10.2941	103	50.4902
청각	0	0		0		0	
예방	1	0	0.0000	0	0.0000	1	100.0000
시도	7	3	42.8571	1	14.2857	3	42.8571
위기	169	53	31.3609	23	13.6095	93	55.0296
범죄	0	0		0		0	
시설	52	21	40.3846	3	5.7692	28	53.8462
기초	824	298	36.1650	85	10.3155	441	53.5194
공공	378	152	40.2116	22	5.8201	204	53.9683
산재	4	2	50.0000	1	25.0000	1	25.0000
재난	0	0		0		0	
금융	468	235	50.2137	30	6.4103	203	43.3761
의료	134	45	33.5821	16	11.9403	73	54.4776
일용	453	244	53.8631	41	9.0508	168	37.0861
영양	84	32	38.0952	22	26.1905	30	35.7143

<표 3-52>는 경기도 대상자에서 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 일용근로자여부 53.86%, 단전여부 52.94%, 고용보험비대상여부 52.09%, 고용보험실직사유대상여부 50.51%, 금융연체대상자여부 50.21% 순으로 높게 나왔다. 비대상의 경우 영양플러스미지원가구의 경우 26.19%로 제일 높게 나왔다. 그리고 산재요양종결후근로단절자여부 25%, 단가스여부 24%, 방문건강집중관리군여부 18.18%, 자살시도대상자여부 14.29% 순으로 높게 나왔다. 조치완료의 경우 자살예방관리대상자여부 100%, 단가스여부 60%, 방문건강집중관리군여부 59.09%, 위기학생여부 55.03% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 단가스여부 16%, 방문건강집중관리군여부 22.73%, 위기학생여부 31.36% 순으로 낮게 나왔다. 비대상 기준 자살예방관리대상자여부가 0건으로 나왔으며, 시설입퇴소여부 5.77%, 공공임대주택체납여부 5.82%, 고용보험개별연장급여대상자여부 5.88% 순으로 낮게 나왔다. 마지막 조치완료의 경우 산재요양종결후근로단절자여부 25%로 나왔으며, 영양플러스미지원가구 35.71%, 일용근로자여부 37.09%, 고용보험실직사유대상여부 37.75%, 고용보험비대상여부 37.96%, 단전여부 41.18% 순으로 낮게 나왔다.

<전라북도>

다음은 전라북도에 거주하는 대상자를 기준으로 대상자 분석을 수행하고자 한다.

<표 3-53> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	3747	1897	50.6272	344	9.1807	1506	40.1922
영유	86	29	33.7209	10	11.6279	47	54.6512
아청	163	106	65.0307	15	9.2025	42	25.7669
청년	409	278	67.9707	46	11.2469	85	20.7824
장중	860	514	59.7674	76	8.8372	270	31.3953
노년	2229	970	43.5173	197	8.8380	1062	47.6447

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

전라북도의 연령대별 복지사각지대 선정 대상자의 2019년 1차 완료결과 현황은 <표 3-53>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 영유아기층의 경우 33.72%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 영유아기층 다음으로 노년층과 장중년층이 43.52%, 59.77%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 영유아기층이 비대상의 비율이 가장 높으며 비슷하게 청년층이 11.25%로 다음으로 높은 것을 알 수 있다. 반면 장중년층의 경우 8.84%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 영유아기층이 54.65%, 노년층이 47.64% 순으로 높게 나타났다. 그리고 청년기층의 경우 20.78%로 제일 낮은 비율을 보여주었다.

<표 3-54> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	3520	2109	59.9148	232	6.5909	1179	33.4943
영유	116	69	59.4828	12	10.3448	35	30.1724
아청	321	197	61.3707	45	14.0187	79	24.6106
청년	603	452	74.9585	50	8.2919	101	16.7496
장중	1026	681	66.3743	70	6.8226	275	26.8031
노년	1454	710	48.8308	55	3.7827	689	47.3865

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

전라북도의 연령대별 복지사각지대 선정 대상자의 2019년 2차 완료결과 현황은 <표 3-54>와 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 노년기층의 경우 48.83%으로 단순상담 비율이 제일 낮은 것을 알 수 있다. 노년기층 다음으로 영유아기층과 아동청소년층이 59.48%, 61.37%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 아동청소년기층이 비대상의 비율이 가장 높으며 비슷하게 영유아기층이 10.34%로 다음으로 높은 것을 알 수 있다. 반면 노년기층의 경우 3.78%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년기층이 47.39%, 영유아기층이 30.17% 순으로 높게 나타났다. 그리고 청년기층의 경우 16.75%로 제일 낮은 비율을 보여주었다.

〈표 3-55〉 변수 조합 개수별 선정 대상자 2019년 1차 ‘완료결과’ 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	3747	1897	50.6272	344	9.1807	1506	40.1922
1개	3184	1631	51.2249	304	9.5477	1249	39.2274
2개	468	225	48.0769	36	7.6923	207	44.2308
3개	79	36	45.5696	3	3.7975	40	50.6329
4개	14	5	35.7143	1	7.1429	8	57.1429
5개	1	0	0.0000	0	0.0000	1	100.0000
6개	1	0	0.0000	0	0.0000	1	100.0000

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계변수 조합 개수 0이 존재할 수 있음

〈표 3-55〉의 경우 전라북도 데이터에서 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 1개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 48.08%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 3개, 4개, 5개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 1개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 2개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 5개와 6개를 가진 대상이 1건 밖에 안 나왔지만 모두 조치완료가 되어 100%로 제일 높은 것을 알 수 있다. 다음으로 변수 4개를 가진 대상이 57.14%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 50.63%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 2개를 가진 대상이 44.23%로 제일 조치완료율이 낮은 것을 알 수 있다.

〈표 3-56〉 변수 조합 개수별 선정 대상자 2019년 2차 ‘완료결과’ 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	3520	2109	59.9148	232	6.5909	1179	33.4943
1개	2625	1530	58.2857	178	6.7810	917	34.9333
2개	755	491	65.0331	43	5.6954	221	29.2715
3개	118	75	63.5593	7	5.9322	36	30.5085
4개	19	10	52.6316	4	21.0526	5	26.3158
5개	3	3	100.0000	0	0.0000	0	0.0000
6개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계변수 조합 개수 0이 존재할 수 있음

〈표 3-56〉의 경우 전라북도 데이터에서 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 5개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 65.03%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 3개, 1개, 4개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 4개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 1개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 1개를 가진 대상이 34.93%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 30.51%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 2개를 가진 대상이 29.27%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-57> 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	1	0	0.0000	0	0.0000	1	100.0000
단수	10	9	90.0000	1	10.0000	0	0.0000
단가	7	3	42.8571	0	0.0000	4	57.1429
전기	33	21	63.6364	1	3.0303	11	33.3333
연금	13	8	61.5385	0	0.0000	5	38.4615
건보	306	209	68.3007	16	5.2288	81	26.4706
화재	0	0		0		0	
경감	0	0		0		0	
장기	2309	1064	46.0806	216	9.3547	1029	44.5647
전세	184	98	53.2609	28	15.2174	58	31.5217
월세	1103	598	54.2158	90	8.1596	415	37.6247
개별	7	2	28.5714	1	14.2857	4	57.1429
실직	7	5	71.4286	0	0.0000	2	28.5714
비대	36	23	63.8889	5	13.8889	8	22.2222
집중	12	4	33.3333	1	8.3333	7	58.3333
분유	40	10	25.0000	10	25.0000	20	50.0000
청각	0	0		0		0	
예방	0	0		0		0	
시도	0	0		0		0	
위기	31	24	77.4194	1	3.2258	6	19.3548
범죄	0	0		0		0	
시설	11	2	18.1818	1	9.0909	8	72.7273
기초	194	76	39.1753	12	6.1856	106	54.6392
공공	10	5	50.0000	0	0.0000	5	50.0000
산재	1	0	0.0000	0	0.0000	1	100.0000
재난	0	0		0		0	
금융	41	21	51.2195	4	9.7561	16	39.0244
의료	41	10	24.3902	1	2.4390	30	73.1707
일용	27	17	62.9630	1	3.7037	9	33.3333
영양	0	0		0		0	

<표 3-57>은 전라북도 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단가스여부 90%, 위기학생여부 77.42%, 고용보험실직사유대상여부 71.43%, 건강보험료체납여부 68.3%, 고용보험비대상여부 63.89% 순으로 높게 나왔다. 비대상의 경우 저귀조제분유지원대상자여부의 경우 25%로 제일 높게 나왔다. 그리고 전세금액기준이하가구여부 15.22%, 고용보험개별연장급여대상자여부 14.29%, 고용보험비대상여부 13.89%, 단수도여부 10% 순으로 높게 나왔다. 조치완료의 경우 단전여부 100%, 산재요양종결후근로단절자여부 100%, 의료비용과다지출가구여부 73.17%, 시설입퇴소여부 72.73% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 시설입퇴소여부 18.18%, 의료비용과다지출가구여부 24.39%, 저귀조제분유지원대상자여부 25%, 고용보험개별연장급여대상자여부 28.57% 순으로 낮게 나왔다. 비대상 기준 단전여부, 산재요양종결후근로단절자여부, 단가스여부, 공공임대주택체납여부, 국민연금체납여부, 고용보험실직사유대상여부가 0건으로 나왔으며, 의료비용과다지출가구여부 2.44%, 전기료체납여부 3.03%, 위기학생여부 3.23% 순으로 낮게 나왔다. 마지막 조치완료의 경우 단수도여부의 경우 0건이 나왔으며, 위기학생여부 19.35%, 고용보험비대상여부 22.22%, 건강보험료체납여부 26.47%, 고용보험실직사유대상여부 28.57%, 전세금액기준이하여부 31.52%, 전기료체납여부 33.33% 순으로 낮게 나왔다.

<표 3-58> 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	4	2	50.0000	2	50.0000	0	0.0000
단수	0	0		0		0	
단가	2	2	100.0000	0	0.0000	0	0.0000
전기	51	34	66.6667	5	9.8039	12	23.5294
연금	32	16	50.0000	3	9.3750	13	40.6250
건보	519	371	71.4836	24	4.6243	124	23.8921
화재	0	0		0		0	
경감	0	0		0		0	
장기	1291	635	49.1867	42	3.2533	614	47.5600
전세	249	174	69.8795	22	8.8353	53	21.2851
월세	1568	1025	65.3699	106	6.7602	437	27.8699
개별	4	2	50.0000	0	0.0000	2	50.0000
실직	25	13	52.0000	2	8.0000	10	40.0000
비대	50	34	68.0000	3	6.0000	13	26.0000
집중	2	0	0.0000	0	0.0000	2	100.0000
분유	70	34	48.5714	13	18.5714	23	32.8571
청각	0	0		0		0	
예방	0	0		0		0	
시도	0	0		0		0	
위기	90	53	58.8889	22	24.4444	15	16.6667
범죄	0	0		0		0	
시설	40	25	62.5000	0	0.0000	15	37.5000
기초	265	156	58.8679	15	5.6604	94	35.4717
공공	53	41	77.3585	2	3.7736	10	18.8679
산재	1	0	0.0000	1	100.0000	0	0.0000
재난	0	0		0		0	
금융	80	57	71.2500	4	5.0000	19	23.7500
의료	50	37	74.0000	3	6.0000	10	20.0000
일용	91	59	64.8352	14	15.3846	18	19.7802
영양	43	22	51.1628	18	41.8605	3	6.9767

<표 3-58>은 전라북도 대상자에서 2019년 2차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단가스여부 100%, 공공임대주택체납여부 77.36%, 의료비용과다지출가구여부 74%, 건강보험료체납여부 71.48%, 금융연체대상자여부 71.25% 순으로 높게 나왔다. 비대상의 경우 산재요양종결후근로단절자여부의 경우 100%로 제일 높게 나왔다. 그리고 단전여부 50%, 영양플러스미지원가구 51.16%, 위기학생여부 58.89%, 기저귀조제분유지원대상자여부 48.57% 순으로 높게 나왔다. 조치완료의 경우 방문건강집중관리군여부 100%, 고용보험개별연장급여대상자여부 50%, 피부양의무자장기요양여부 47.56%, 국민연금체납여부 40.62% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 기저귀조제분유지원대상자여부 48.57%, 피부양의무자장기요양여부 49.19%, 고용보험개별연장급여대상자여부 50% 순으로 낮게 나왔다. 비대상 기준 방문건강집중관리군여부, 고용보험개별연장급여대상자여부, 시설입퇴소여부, 단가스여부가 0건으로 나왔으며, 피부양의무자장기요양여부 3.25%, 공공임대주택 3.77%, 건강보험료체납여부 4.62% 순으로 낮게 나왔다. 마지막 조치완료의 경우 단가스여부, 단전여부, 산재요양종결후근로단절자여부가 0건으로 나왔다. 그리고 영양플러스미지원가구 6.98%, 위기학생여부 16.67%, 공공임대주택체납여부 18.87%, 일용근로자여부 19.78%, 의료비용과다지출가구여부 20%, 전세금액기준이하가구여부 21.29% 순으로 낮게 나왔다.

<제주특별자치도>

다음은 제주특별자치도 대상자에 대한 속성분석을 수행하고자 한다.

<표 3-59> 연령대별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	922	319	34.5987	47	5.0976	556	60.3037
영유	28	7	25.0000	0	0.0000	21	75.0000
아청	60	22	36.6667	4	6.6667	34	56.6667
청년	153	86	56.2092	6	3.9216	61	39.8693
장중	271	117	43.1734	16	5.9041	138	50.9225
노년	410	87	21.2195	21	5.1220	302	73.6585

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

제주특별자치시 연령대별 복지사각지대 선정 대상자의 2019년 1차 완료결과 현황은 <표 3-59>와 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 노년 기층의 경우 21.22%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 노년기기층 다음으로 영유아층과 아동청소년층이 25%, 36.67%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 아동청소년기층이 비대상의 비율이 가장 높으며 비슷하게 장중년기층이 5.9%로 다음으로 높은 것을 알 수 있다. 반면 영유아기층의 경우 0%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 영유아기층이 75%, 노년층이 73.66% 순으로 높게 나타났다. 그리고 청년기층의 경우 39.87%로 제일 낮은 비율을 보여주었다.

<표 3-60> 연령대별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	922	398	43.1670	47	5.0976	477	51.7354
영유	34	17	50.0000	2	5.8824	15	44.1176
아청	136	71	52.2059	5	3.6765	60	44.1176
청년	215	111	51.6279	10	4.6512	94	43.7209
장중	352	149	42.3295	11	3.1250	192	54.5455
노년	185	50	27.0270	19	10.2703	116	62.7027

* 영·유아기: 만0~6세, 아동·청소년기: 만7~18세, 청년기: 만19~39세, 장·중년기: 만40~64세

* 영유: 영유아기, 아청: 아동청소년기, 청년: 청년기, 장중: 장중년기, 노년: 노년기

제주특별자치시 연령대별 복지사각지대 선정 대상자의 2019년 2차 완료결과 현황은 <표 3-60>과 같이 볼 수 있다. 단순상담을 기준으로 보면 청년층이 제일 높은 단순상담 비율을 보여주었다. 반면에 노년기층의 경우 27.03%로 단순상담 비율이 제일 낮은 것을 알 수 있다. 노년기기층 다음으로 장중년기층과 영유아기층이 42.33%, 50%로 높은 단순상담 비율을 보여주고 있다. 비대상을 기준으로 보면 다음과 같이 볼 수 있다. 노년기층이 비대상의 비율이 가장 높으며 비슷하게 영유아기층이 5.88%로 다음으로 높은 것을 알 수 있다. 반면 장중년기층의 경우 3.13%로 제일 낮은 비율을 보여준다. 조치완료의 경우는 다음과 같이 볼 수 있다. 노년기층이 62.7%, 장중년기층이 54.55% 순으로 높게 나타났다. 그리고 청년기층의 경우 43.72%로 제일 낮은 비율을 보여주었다.

<표 3-61> 변수 조합 개수별 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	922	319	34.5987	47	5.0976	556	60.3037
1개	737	265	35.9566	38	5.1560	434	58.8874
2개	150	49	32.6667	7	4.6667	94	62.6667
3개	33	5	15.1515	2	6.0606	26	78.7879
4개	1	0	0.0000	0	0.0000	1	100.0000
5개	1	0	0.0000	0	0.0000	1	100.0000

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계 변수 조합 개수 0이 존재할 수 있음

<표 3-61>의 경우 제주특별자치시 데이터에서 변수 조합 개수별 선정 대상자의 2019년 1차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 1개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 32.67%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 3개, 4개, 5개 순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 3개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 1개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 5개를 가진 대상이 1건 밖에 안 나왔지만 모두 조치완료가 되어 100%로 제일 높은 것을 알 수 있다. 다음으로 변수 3개를 가진 대상이 78.79%로 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 62.67%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 1개를 가진 대상이 58.89%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-62> 변수 조합 개수별 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
전체	922	398	43.1670	47	5.0976	477	51.7354
1개	637	278	43.6421	28	4.3956	331	51.9623
2개	235	106	45.1064	11	4.6809	118	50.2128
3개	47	12	25.5319	7	14.8936	28	59.5745
4개	3	2	66.6667	1	33.3333	0	0.0000
5개	0	0		0		0	

* 28종의 연계변수 이외에 복지사각지대 발굴대상자 선정하는 변수가 존재하기 때문에 연계변수 조합 개수 0이 존재할 수 있음

<표 3-62>의 경우 제주특별자치시 데이터에서 변수 조합 개수별 선정 대상자의 2019년 2차 완료결과 현황을 보여주고 있다. 단순상담 기준으로 보면 변수를 4개 가진 대상이 비율이 제일 높은 것을 알 수 있다. 다음으로 변수 2개를 가진 대상이 45.11%로 두 번째로 높은 것을 알 수 있다. 다음 순서대로 1개, 3개순으로 단순 상담의 비율이 높은 것을 알 수 있다.

비대상 기준으로 보면 다음과 같이 볼 수 있다. 변수 4개를 가진 대상이 비대상일 확률이 가장 높으며, 다음으로 3개를 가진 대상이 두 번째로 비대상일 확률이 높게 나타났다. 조치완료 기준으로 보면 변수 3개를 가진 대상이 59.57%로 제일 높은 것을 알 수 있다. 다음으로 변수 1개를 가진 대상이 51.96%로 두 번째로 높은 것을 알 수 있다. 그리고 변수 4개를 가진 대상이 0%로 제일 조치완료율이 낮은 것을 알 수 있다.

<표 3-63> 선정 대상자 2019년 1차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	1	1	100.0000	0	0.0000	0	0.0000
단수	0	0		0		0	
단가	0	0		0		0	
전기	5	1	20.0000	0	0.0000	4	80.0000
연금	8	2	25.0000	1	12.5000	5	62.5000
건보	124	55	44.3548	10	8.0645	59	47.5806
화재	0	0		0		0	
경감	0	0		0		0	
장기	340	95	27.9412	18	5.2941	227	66.7647
전세	17	4	23.5294	0	0.0000	13	76.4706
월세	497	169	34.0040	22	4.4266	306	61.5694
개별	2	1	50.0000	0	0.0000	1	50.0000
실직	4	3	75.0000	0	0.0000	1	25.0000
비대	26	15	57.6923	0	0.0000	11	42.3077
집중	3	0	0.0000	0	0.0000	3	100.0000
분유	30	2	6.6667	0	0.0000	28	93.3333
청각	0	0		0		0	
예방	0	0		0		0	
시도	1	0	0.0000	0	0.0000	1	100.0000
위기	16	6	37.5000	2	12.5000	8	50.0000
범죄	0	0		0		0	
시설	1	0	0.0000	0	0.0000	1	100.0000
기초	33	9	27.2727	2	6.0606	22	66.6667
공공	4	0	0.0000	1	25.0000	3	75.0000
산재	0	0		0		0	
재난	0	0		0		0	
금융	14	8	57.1429	1	7.1429	5	35.7143
의료	12	4	33.3333	0	0.0000	8	66.6667
일용	7	3	42.8571	1	14.2857	3	42.8571
영양	0	0		0		0	

<표 3-63>은 제주특별자치시 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 단전여부 100%, 고용보험실직사유 대상여부 75%, 고용보험비대상여부 57.69%, 금융연체대상자여부 57.14%, 고용보험개별연장급여대상자여부 50% 순으로 높게 나왔다. 비대상의 경우 공공임대주택체납여부의 경우 25%로 제일 높게 나왔다. 그리고 일용근로자여부 14.29%, 위기학생여부 12.5%, 국민연금체납여부 12.5%, 건강보험료체납여부 8.06% 순으로 높게 나왔다. 조치완료의 경우 방문건강집중관리군여부, 자살시도대상자여부, 시설입퇴소여부의 경우 100%, 기저귀조제분유지원대상자여부 93.33%, 전기료체납여부 80%, 전세금액기준이하가구여부 76.47% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 방문건강집중관리군여부, 자살시도대상자여부, 시설입퇴소여부, 공공임대주택여부의 경우 0건으로 나왔다. 그리고 기저귀조제분유지원대상자여부 6.66%, 전기료체납여부 20%, 전세금액기준이하여부 23.53% 순으로 낮게 나왔다. 비대상 기준 방문건강집중관리군여부, 자살시도대상자여부, 기저귀조제분유지원대상자여부, 전기료체납여부, 전세금액기준이하가구여부, 의료비용과다지출가구여부, 고용보험개별 연장급여대상자여부, 고용보험비대상여부, 고용보험실직실직사유대상여부, 단전여부가 0건으로 나왔으며, 월세금액기준이하여부 4.43%, 피부양장기요양여부 5.29%, 기초생활긴급지원수급탈락여부 6.06% 순으로 낮게 나왔다. 마지막 조치완료의 경우 단전여부가 0건으로 나왔다. 그리고 고용보험실직사유대상여부 25%, 금융연체대상자여부 35.71%, 고용보험비대상여부 42.31%, 일용근로자여부 42.86%, 건강보험료체납여부 47.58% 순으로 낮게 나왔다.

<표 3-64> 선정 대상자 2019년 2차 '완료결과' 현황

(단위: 명, %)

구분		단순상담		비대상		조치완료	
단전	0	0		0		0	
단수	0	0		0		0	
단가	0	0		0		0	
전기	7	0	0.0000	0	0.0000	7	100.0000
연금	11	3	27.2727	0	0.0000	8	72.7273
건보	231	101	43.7229	12	5.1948	118	51.0823
화재	0	0		0		0	
경감	0	0		0		0	
장기	101	31	30.6931	13	12.8713	57	56.4356
전세	26	7	26.9231	1	3.8462	18	69.2308
월세	636	263	41.3522	32	5.0314	341	53.6164
개별	0	0		0		0	
실직	8	2	25.0000	1	12.5000	5	62.5000
비대	25	13	52.0000	1	4.0000	11	44.0000
집중	3	1	33.3333	0	0.0000	2	66.6667
분유	34	12	35.2941	5	14.7059	17	50.0000
청각	0	0		0		0	
예방	0	0		0		0	
시도	1	1	100.0000	0	0.0000	0	0.0000
위기	43	31	72.0930	1	2.3256	11	25.5814
범죄	0	0		0		0	
시설	2	1	50.0000	0	0.0000	1	50.0000
기초	56	20	35.7143	5	8.9286	31	55.3571
공공	7	5	71.4286	0	0.0000	2	28.5714
산재	1	1	100.0000	0	0.0000	0	0.0000
재난	0	0		0		0	
금융	31	17	54.8387	2	6.4516	12	38.7097
의료	4	3	75.0000	0	0.0000	1	25.0000
일용	32	22	68.7500	2	6.2500	8	25.0000
영양	1	0	0.0000	0	0.0000	1	100.0000

<표 3-64>는 제주특별자치시 대상자에서 2019년 1차 ‘완료결과’ 현황을 보여준다. 단순상담의 경우 자살시도대상자여부, 산재요양종결후근로단절자여부 100%, 의료비용과다지출가구여부 75%, 위기학생여부 72.09%, 공공임대주택체납여부 71.43%, 일용근로자여부 68.75% 순으로 높게 나왔다. 비대상의 경우 기저귀조제분유지원대상자여부가 14.71%로 제일 높게 나왔다. 그리고 피부양의무자장기요양여부 12.87%, 고용보험실직사유대상여부 12.5%, 기초생활긴급지원수급탈락여부 8.93%, 금융연체대상자여부 6.45% 순으로 높게 나왔다. 조치완료의 경우 전기료체납여부, 국민연금체납여부의 경우 100%로 나왔다. 다음으로 국민연금체납여부 72.73%, 전세금액기준이하가구여부 69.23%, 방문건강집중관리군여부 66.67% 순으로 높게 나왔다.

반대로 낮은 순으로 보면 다음과 같이 볼 수 있다. 단순상담 기준으로 전기료체납여부, 영양플러스미지원가구여부의 경우 0건으로 나왔다. 그리고 고용보험실직실직사유대상여부 25%, 전세금액기준이하가구여부 26.93%, 국민연금체납여부 27.27% 순으로 낮게 나왔다. 비대상 기준 전기료체납여부, 영양플러스미지원가구여부, 국민연금체납여부, 방문건강집중관리군여부, 시설입퇴소여부, 공공임대주택체납여부, 의료비용과다지출가구여부, 자살시도대상자여부, 산재요양종결후근로단절자여부가 0건으로 나왔으며, 위기학생여부 2.33%, 전세금액기준이하가구여부 3.85%, 고용보험비대상여부 4% 순으로 낮게 나왔다. 마지막 조치완료의 경우 자살시도대상자여부, 산재요양종결후근로단절자여부가 0건으로 나왔다. 그리고 의료비용과다지출가구여부 25%, 일용근로자여부 25%, 위기학생여부 25.58%, 공공임대주택체납여부 28.57%, 금융연체대상자여부 38.71%, 고용보험비대상여부 44% 순으로 낮게 나왔다.

예측모형 구현 및 분석

제1절 예측 모형 구현

제2절 지역별 예측 모델의 대상자 분포도 및
가중치 분석

제3절 지역별 예측 모델 간 성능 분석

제4절 파생변수 투입에 따른 효과 분석

제4장 예측모형 구현 및 분석

제1절 예측 모형 구현

본 장에서는 지역별 예측 모형을 구현하고 각 지역별 모형에서 가상 시뮬레이션을 통해 전체 분석 대상자에서 특정 대상을 추출했을 때의 전체 대상자의 분포도와 각 지역별 모형이 가지고 있는 가중치를 분석한다. 이 때 본 연구에서는 예측 모형을 앙상블 모형인 XGBoost 만을 고려한다. XGboost 모형은 2장 선행연구에서 설명한 부스팅 모형으로써 설명력은 많이 떨어지지만 예측력은 좋은 모형으로써 분리를 위해서 많이 사용되는 모형이다.

지역별 예측모형 구현을 위해서 본 연구진은 <표 4-1>과 같이 데이터를 분리하였다. 데이터는 훈련데이터 80%와 테스트데이터 20%로 분리하여 테스트를 진행하였다.

<표 4-1> 지역별 훈련데이터 및 테스트데이터 현황

구분		훈련데이터(80%)	테스트데이터(20%)
순번	시군구	대상자 수	대상자 수
1	기본	445,550	111,387
2	서울특별시	57,464	14,232
3	경기도	101,581	25,351
4	전라북도	28,751	7,200
5	제주특별자치도	8,942	2,240

<표 4-2>는 훈련데이터와 테스트 데이터에서 종속변수인 대상자의 서비스 지원 내역을 표시하는 것을 보여준다. 대상자의 경우 공공부조를 받은 대상자를 의미하며, 비대상자의 경우 공공부조를 받지 못한 대상자를 의미한다.

<표 4-2> 대상자 서비스 지원 내역 현황

구분			훈련 데이터 80%		테스트 데이터 20%	
순번	시군구	대상자 수	Target (1)	Non-Target (0)	Target (1)	Non-Target (0)
	기본	144,474	301,076	35,803	35,803	75,584
1	서울특별시	57,464	16,973	40,491	4,216	10,016
9	경기도	101,581	32,721	68,860	8,003	17,348
13	전라북도	28,751	11,200	17,551	2,722	4,478
17	제주특별자치도	8,942	2,631	6,311	661	1,579

이 때 본 연구에서는 기본 데이터를 활용한 모형, 기본 데이터에 시군구 코드 값을 파생변수로 생성하여 만든 모형과 서울특별시, 경기도, 전라북도, 제주특별자치도로 군집화 하여 생성한 모형끼리 비교 분석을 수행한다. 모형 생성할 때에는 결정트리 기반의 부스팅 기법을 활용하는 XGBoost 기반의 분석을 수행한다.

<표 4-3>은 기본적으로 모형에서 활용하는 변수 목록을 보여준다.

복지사각지대에서 사용하는 기본적인 변수 목록은 다음과 같이 총 30개의 변수로 구성되어 있다. 각 변수는 변경 없이 사용되거나 각 변수가 가지고 있는 속성변수를 변형해서 사용되거나 파생변수를 만들어 사용된다.

〈표 4-3〉 변수 목록

변수명	컬럼명
V1	단전여부
V2	단수도여부
V3	단가스여부
V4	전기료체납여부
V5	국민연금체납여부
V6	건강보험료체납여부
V7	화재피해여부
V8	본인부담경감대상자여부
V9	피부양의무자장기요양여부
V10	전세금액기준이하가구여부
V11	월세금액기준이하가구여부
V12	고용보험개별연장급여대상여부
V13	고용보험실직사유대상여부
V14	고용보험비대상여부
V15	방문건강집중관리군여부
V16	기저귀조제분유지원대상자여부
V17	신생아난청확진자여부
V18	자살예방관리대상자여부
V19	자살시도대상자여부
V20	위기학생여부
V21	범죄피해여부
V22	시설입퇴소여부
V23	기초생활긴급지원수급탈락여부
V24	공공임대주택체납자여부
V25	산재요양종결후근로단절자여부
V26	재난피해자여부
V27	금융연체대상자여부
V28	의료비용과다지출가구여부

파생변수의 생성은 시군구 17개의 시군구를 하나의 코드화 시켜 1~17까지의 범주형 변수화 시켜 독립변수를 하나의 변수로 활용하여 수행하였다.

이에 따라 본 연구에서 고려하고 있는 모형은 다음과 같은 모형이 구성될 수 있다. 기본 모형 같은 경우 사회보장정보원에 연계되고 있는 기본적인 연계변수를 활용한 모형이다. 즉 기본 모형은 과거에 진행된 예측 모형 방법이라고 할 수 있다. 그리고 지역 파생 모형의 경우 기본적으로 연계되는 연계변수에 지역 파생변수가 추가 된 모형이다. 서울특별시, 경기도, 전라북도, 제주특별자치도 모형 같은 경우에는 기본 모형에서 지역 데이터로 분리하여 군집화하여 새로운 모형을 만든 것을 의미한다. 이와 같은 모형은 또한 1인가구와 그렇지 않은 가구로 분리하여 분리된 모형으로써 구현하였다. 그 이유는 현재 사회보장정보원의 복지사각지대발굴관리 시스템이 모형 다각화가 진행 중에 있으며, 1인/다인가구로 분리되어 수행되고 있기 때문이다.

<표 4-4> 기본 모형 및 지역별 모형 현황

모형	설명
기본 모형	사각지대 기본 연계변수
지역 파생 모형	사각지대 기본 연계변수+지역파생
서울특별시 모형	서울특별시 데이터 군집 사각지대 기본 연계변수
경기도 모형	경기도 데이터 군집 사각지대 기본 연계변수
전라북도 모형	전라북도 데이터 군집 사각지대 기본 연계변수
제주특별자치도 모형	제주특별자치도 데이터 군집 사각지대 기본 연계변수

본 장에서는 모형에 대한 평가는 대상자 분포도, 가중치, 분류행렬, AUC, Lift Chart를 통해서 모형에 대한 세부 평가를 수행하여 최적의 모형 도출을 수행하고자한다.

1. 분포도: 본 연구에서 분포도라고 하는 것은 훈련된 예측 모형에 19년도 2차수를 기준으로 하여 약 400만명이 넘는 대상자를 넣고 1인가구의 경우 약 2만명, 다인가구의 경우 약 4만명을 추출하여 연계변수별 속하는 대상자의 수를 본 것을 의미함.
2. 가중치: 가중치라고 하는 것은 XGBoost 모형에서 산출된 Gain의 값을 의미하는 것으로써, Gain은 각 변수가 가지는 상대적인 중요도라고 할 수 있음.
3. 분류 행렬 (Confusion Matrix): 모형 평가 방법중 한가지 방법으로 정확도, 민감도, 특이도, 정밀도를 산출하여 모형간의 비교를 수행할 때 쓰이는 도표임.
4. AUC(Area Under Curve): ROC커브의 넓이의 값으로써 1에 가까울수록 더욱 더 좋은 모형을 의미함.
5. Lift chart: 사후 확률에 의해 모형을 평가하는 방법으로 특정한 클래스가 상대적으로 드물고 다른 클래스보다 훨씬 더 관심이 있는 경우에 활용하는 방법으로써 상위와 하위 간 격차가 더 클수록 좋은 모형이라 할 수 있음.

2절에서는 기본 모형과 지역별 모형인 서울특별시 모형 경기도 모형 전라북도 모형과 제주특별자치도 모형에 대해 가중치를 살펴본다. 그리고 3절에서는 모형 평가를 통해 기본 모형 대신 지역별 모형에 대한 효과를 분석한다. 4절에서는 지역 파생 모형에 대한 내용에 대해 살펴보고 5절에서는 최종적으로 비교 분석을 수행한다.

제2절 지역별 예측 모델의 분포도 및 가중치 분석

2절에서는 기본 모형과 지역별 모형인 서울특별시, 경기도, 전라북도, 제주특별자치도 모형의 대상자 분포도와 가중치를 분석을 수행한다. 대상자 분포도의 경우 약 400만으로부터 위험 점수 상위 1인가구의 경우 2만, 2인이상 가구의 경우 4만명을 추출하여 어떤 대상자가 어느 변수에 많이 물리는지에 대한 분석을 수행하고자 한다.

우선적으로 기본 모형에 대한 1인가구, 다인가구에 대한 대상자 분포도와 가중치를 살펴본다. <표 4-5>는 1인가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 월세금액기준이하가구여부(H11) 변수로써 18,172명으로 가장 많은 분포를 보여주었으며, 다음으로 피부양 의무자장기요양여부 (H9)변수가 2,398명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 월세금액기준이하가구여부 (H11)가 15,241명 그리고 전세금액기준이하가구여부 (H10) 46명으로 단독 변수 보유 대상자를 보였다. 나머지 변수의 경우 단독 변수는 없으며 타 변수와 중복되는 변수로 발생하는 것을 알 수 있었다. 중복 변수 기준으로는 총 변수와 동일하게 월세금액기준이하가구여부 (H11)이 2931명, 피부양의무자장기요양여부 (H9)이 2,398명으로 높은 것을 알 수 있었다. 가중치 기준으로 보면 피부양의무자장기요양여부 (H9) 0.238, 월세전환용액변수 (H11_R) 0.125, 기초생활긴급지원수급 탈락여부 (H23) 0.107 순으로 가중치가 높은 것을 알 수 있다. 이런 결과를 통해 월세전환총액변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포되며, 피부양의무자장기요양여부 가중치가 높기 때문에 월세 다음으로 분포되며 피부양의무자장기요양여부 변수의 경우 단독변수의 수가 없기 때문에 월세 또는 전세 변수와 같이 발생하는 것을 알 수 있다.

<표 4-5> 기본 모형 대상자 분포도와 가중치(1인가구)

기본 모형		기본 모형 1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	30	0	30	0
H1_C		0	0	0	0
H1_D		0	0	0	0
H1_G	체납전체금액	0	0	0	0.000599
H1_H	월평균사용량	0	0	0	0.003292
H2	단수도여부	0	0	0	0
H2_D		0	0	0	0
H2_E		0	0	0	0
H2_I		0	0	0	0
H2_F		0	0	0	0
H3	단가스여부	13	0	13	0.001171
H3_C		0	0	0	0
H3_D		0	0	0	0
H3_E		0	0	0	0
H3_F		0	0	0	0
H4	전기료체납 여부	117	0	117	0.009193
H4_A		0	0	0	0
H4_C	체납개월수	0	0	0	0.01264
H4_D		0	0	0	0
H4_E	월평균사용량	0	0	0	0.021747
H5	국민연금체납 여부	48	0	48	0
H5_A		0	0	0	0.004469
H5_D	체납개월수	0	0	0	0.007842
H5_E	체납보험료	0	0	0	0.000812
H5_A	월 보험료	0	0	0	0.004469
H6	건강보험료 체납여부	1,209	0	1,209	0
H6_C	체납개월수	0	0	0	0.016246
H6_D	체납보험료	0	0	0	0.044748
H6_E	월 보험료	0	0	0	0.044774

기본 모형		기본 모형 1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H8_new	의료비과다	0	0	0	0
H9	피부양 의무자 장기요양여부	2,398	0	2,398	0.238417
H9_B	월보험료	0	0	0	0
H10	전세금액기준 이하가구여부	661	46	615	0.062572
H10_B	세대원수	0	0	0	0.030974
H10_A	임차보증금액	0	0	0	0.063884
H11	월세금액기준 이하가구여부	18,172	15,241	2,931	0.003881
H11_C	세대원수	0	0	0	0.078991
H11_R	월세전환총액 변수	0	0	0	0.125113
H13	고용보험실직 사유대상여부	206	0	206	0.002395
H15	방문건강집중 관리군여부	34	0	34	0.033584
H22	시설입퇴소여 부	135	0	135	0.00968
H23	기초생활긴급 지원수급탈락 여부	1,033	0	1,033	0.107491
SS1	소득인정액정 보	0	0	0	0
SS1_A	중위소득비율	0	0	0	0
region_B	지역(중)	0	0	0	0.017931
region_C	지역(소)	0	0	0	0.021933
H18	자살예방관리 대상자여부	34	0	34	0.002624
H19	자살 시도대상 자여부	6	0	6	0
H20	위기학생여부	0	0	0	0
H16	기저귀조제분 유지원대상자 여부	0	0	0	0
H21	범죄피해여부	0	0	0	0

기본 모형		기본 모형 1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H7	화재피해여부	4	0	4	0
H14	고용보험비대 상여부	878	0	878	0.0055
H24	공공임대주택 채납자여부	20	0	20	0.002678
H24_C		0	0	0	0.006052
H25	산재요양종결 후근로단절자 여부	5	0	5	0.001461
H25_D	평균임금	0	0	0	0
H27	금융연체대상 자여부	201	0	201	0.000162
H27_A	현재연체금액	0	0	0	0.01423
H27_B	연체등록건수	0	0	0	0
H27_C	연체발생기관 수	0	0	0	0
H27_M	대부연체금액 포함여부	0	0	0	0
H28	의료비용과다 지출가구여부	194	0	194	0
H12	고용보험개별 연장급여대상 여부	0	0	0	0
H17	신생아난청확 진자여부	0	0	0	0
H24_C	임대료채납금 액	0	0	0	0.006052
H26	재난피해자여 부	0	0	0	0
H28_A	본인부담금총 액	0	0	0	0.002916
AREA	지역 파생변수	0	0	0	0
합계		0	15,287	10,111	1.010521

<표 4-6>은 다인 가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 1인 가구와 마찬가지로 월세금액기준이하가구여부(H11) 변수로써 18,042명으로 가장 많은 분포를 보여주었으며, 다음으로 피부양 의무자장기요양여부 (H9)변수가 15,183명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 월세금액기준이하가구여부 (H11)가 12,449명으로 제일 대상자가 많이 분포된 변수는 1인 가구와 동일하지만, 두 번째로 높은 변수는 피부양의무자장기요양여부 (H9)변수가 11,787명으로 1인 가구와 다른 양상을 보였다. 중복 변수 기준으로는 1인 가구와 동일하게 월세금액기준이하가구여부(H11)이 5,593명으로 제일 높으며, 1인 가구는 다르게 건강보험료체납여부 (H6)가 5,096명으로 두 번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액변수 (H11_R) 0.1715, 피부양의무자장기요양여부 (H9) 0.1058, 건강보험료 체납보험료 (H6_D) 0.0844 순으로 가중치가 높은 것을 알 수 있었다. 이런 결과를 통해 1인가구와 마찬가지로 월세전환총액변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포가 많이 된다는 사실을 알 수 있었다.

<표 4-6> 기본 모형 대상자 분포도와 가중치(다인가구)

기본 모형		기본 모형 다인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	62	1	61	0.000175
H1_C		0	0	0	0
H1_D		0	0	0	0
H1_G	체납전체금액	0	0	0	0.005975
H1_H	월평균사용량	0	0	0	0.00895

기본 모형		기본 모형 다인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H2	단수도여부	1	0	1	0.000109
H2_D		0	0	0	0
H2_E		0	0	0	0
H2_I		0	0	0	0
H2_F		0	0	0	0
H3	단가스여부	69	0	69	0.001241
H3_C		0	0	0	0
H3_D		0	0	0	0
H3_E		0	0	0	0
H3_F		0	0	0	0
H4	전기료체납여부	735	2	733	0.008335
H4_A		0	0	0	0
H4_C	체납개월수	0	0	0	0.007794
H4_D		0	0	0	0
H4_E	월평균사용량	0	0	0	0.024603
H5	국민연금체납여부	313	0	313	0.004576
H5_A		0	0	0	0.022779
H5_D	체납개월수	0	0	0	0.002426
H5_E	체납보험료	0	0	0	0.020213
H5_A	월보험료	0	0	0	0.022779
H6	건강보험료체납여부	5,285	189	5,096	0.003978
H6_C	체납개월수	0	0	0	0.035153
H6_D	체납보험료	0	0	0	0.084443
H6_E	월보험료	0	0	0	0.064275
H8_new	의료비과다	0	0	0	0
H9	피부양 의무자 장기요양여부	15,183	11,787	3,396	0.105833
H9_B	월보험료	0	0	0	0
H10	전세금액기준 이하가구여부	6,649	3,074	3,575	0.013787

기본 모형		기본 모형 다인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H10_B	세대원수	0	0	0	0.032475
H10_A	임차보증금액	0	0	0	0.077535
H11	월세금액기준 이하가구여부	18,042	12,449	5,593	0.07286
H11_C	세대원수	0	0	0	0.040892
H11_R	월세전환총액 변수	0	0	0	0.171452
H13	고용보험실직 사유대상여부	245	0	245	0.007062
H15	방문건강집중 관리군여부	109	0	109	0.002854
H22	시설입퇴소 여부	270	0	270	0.003602
H23	기초생활긴급 지원수급탈락 여부	3,199	0	3,199	0.034712
SS1	소득인정액 정보	0	0	0	0
SS1_A	중위소득비율	0	0	0	0
region_B	지역(중)	0	0	0	0.020205
region_C	지역(소)	0	0	0	0.017242
H18	자살예방관리 대상자여부	19	0	19	0.001076
H19	자살시도대상 자여부	18	0	18	0.000155
H20	위기학생여부	1,589	0	1,589	0.020232
H16	기저귀조제분 유지원대상자 여부	1,837	0	1,837	0.021672
H21	범죄피해여부	0	0	0	0
H7	화재피해여부	11	0	11	6.78E-05
H14	고용보험비대 상여부	545	0	545	0.003778

기본 모형		기본 모형 다인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H24	공공임대주택 채납자여부	231	0	231	0.000909
H24_C		0	0	0	0.005427
H25	산재요양종결 후근로단절자 여부	12	0	12	0.005168
H25_D	평균임금	0	0	0	0
H27	금융연체 대상자여부	821	0	821	0.014912
H27_A	현재연체금액	0	0	0	0.025294
H27_B	연체등록건수	0	0	0	0
H27_C	연체발생 기관수	0	0	0	0
H27_M	대부연체금액 포함여부	0	0	0	0
H28	의료비용과다 지출가구여부	282	0	282	0.000426
H12	고용보험개별 연장급여대상 여부	3	0	3	0
H17	신생아난청확 진자여부	0	0	0	0
H24_C	임대료채납 금액	0	0	0	0.005427
H26	재난피해자 여부	0	0	0	0.000336
H28_A	본인부담금 총액	0	0	0	0.005008
AREA	지역 파생변수	0	0	0	0
합계		55,530	27,502	28,028	1.028206

다음으로 서울특별시 모형에 대한 1인가구, 다인가구에 대한 대상자 분포도와 가중치를 살펴본다. <표 4-7>은 1인가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 월세금액기준이하가구여부(H11) 변수로써 17,029명으로 가장 많은 분포를 보여주었으며, 다음으로 건강보험료체납여부 (H6)변수가 2,923명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 월세금액기준이하가구여부 (H11)가 14,923명 그리고 건강보험료체납여부 (H6) 변수가 1,937명으로 단독 변수 보유 대상자를 보였다. 중복 변수 기준으로는 총 변수와 동일하게 월세금액기준이하가구여부(H11)이 2,106명, 건강보험료체납여부 (H6)가 986명으로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세 세대원수 (H11_C)가 0.1064, 전세금액기준이하가구여부 (H10) 0.1062, 기초생활긴급지원수급탈락여부 (H23) 0.0985 순으로 가중치가 높은 것을 알 수 있다. 이런 결과를 통해 월세 세대원수변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포되며, 그렇기 때문에 전세금액기준이하가구여부의 가중치가 2순위로 0.1062를 가져도 많은 대상자가 월세금액기준이하가구여부로부터 나오는 것이다.

<표 4-7> 서울특별시 모형 대상자 분포도와 가중치(1인가구)

서울특별시 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	32	-	32	0
H1_G	체납전체금액	-	-	-	0
H1_H	월평균사용량	-	-	-	0.0024369
H2	단수도여부	-	-	-	0

서울특별시 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H3	단가스여부	154	-	154	0.0035478
H4	전기료체납여부	817	219	598	0
H4_C	체납개월수	-	-	-	0.016387
H4_E	월평균사용량	-	-	-	0.0693104
H5	국민연금체납 여부	438	-	438	0
H5_A		-	-	-	0
H5_D	체납개월수	-	-	-	0
H5_E	체납보험료	-	-	-	0.0012939
H5_A	월보험료	-	-	-	0
H6	건강보험료체납 여부	2,923	1,937	986	0.0302959
H6_C	체납개월수	-	-	-	0.0450568
H6_D	체납보험료	-	-	-	0.0871986
H6_E	월보험료	-	-	-	0.0359661
H8_new	의료비과다	-	-	-	0
H9	피부양 의무자 장기요양여부	452	-	452	0.0916802
H10	전세금액기준 이하가구여부	460	69	391	0.106199
H10_B	세대원수	-	-	-	0.0069194
H10_A	임차보증금액	-	-	-	0.0551404
H11	월세금액기준 이하가구여부	17,029	14,923	2,106	0.0866522
H11_C	세대원수	-	-	-	0.1063521

서울특별시 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H11_R	월세전환총액 변수	-	-	-	0
H13	고용보험실직 사유대상여부	175	-	175	0
H15	방문건강집중 관리군여부	14	-	14	0.0126283
H22	시설입퇴소여부	32	-	32	0.0613132
H23	기초생활긴급지 원수급탈락여부	111	-	111	0.0984988
H18	자살예방관리 대상자여부	-	-	-	0
H19	자살시도대상자 여부	2	-	2	0
H20	위기학생여부	-	-	-	0
H16	기저귀조제분유 지원대상자여부	-	-	-	0
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	1	-	1	0
H14	고용보험비대상 여부	530	-	530	0
H24	공공임대주택 채납자여부	19	-	19	0.0009628
H24_C		-	-	-	0.0119053
H25	산재요양종결 후 근로단절자여부	12	-	12	0
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	169	-	169	0.0101948

서울특별시 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H27_A	현재연체금액	-	-	-	0.050126
H28	의료비용과다 지출가구여부	19	-	19	0.0008285
H12	고용보험개별연 장급여대상여부	-	-	-	0
H17	신생아난청 확진자여부	-	-	-	0
H24_C	임대료체납금액	-	-	-	0.0119053
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0.0091054
합계		-	17,148	6,241	1.011905

<표 4-8>은 서울시 모형 다인 가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 1인 가구와 마찬가지로 월세금액기준이하가구여부(H11) 변수로써 21,253명으로 가장 많은 분포를 보여주었으며, 다음으로 전세금액기준이하가구여부 (H10)변수가 11,010명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로도 월세금액기준이하가구여부(H11)가 16,450명으로 제일 대상자가 많았고, 두 번째로 높은 변수도 전세금액기준이하가구여부 (H10) 변수가 5,598명으로 나타났다. 중복 변수 기준으로는 월세금액기준이하가구여부(H11)이 4,803명으로 제일 높으며, 건강보험료체납여부(H6)가 3,893명으로 두 번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액변수 (H11_R) 0.1645, 전세금액기준이하가구 임차보증금액(H10_A) 0.10171, 건강보험료 체납보험료 (H6_D) 0.09714, 건강보험

료 월보험료 (H6_E) 0.08391 순으로 가중치가 높은 것을 알 수 있었다. 이런 결과를 통해 1인가구와 마찬가지로 다인 가구에서는 월세전 환총액변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포가 많이 된다는 사실을 알 수 있었다.

<표 4-8> 서울특별시 모형 대상자 분포도와 가중치(다인가구)

서울특별시 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	40	1	39	0.000424
H1_G	체납전체금액	-	-	-	0.0057358
H1_H	월평균사용량	-	-	-	0.0071542
H2	단수도여부	-	-	-	0
H3	단가스여부	136	-	136	0.0038545
H4	전기료체납여부	1,608	-	1,608	0.0070085
H4_C	체납개월수	-	-	-	0.011481
H4_E	월평균사용량	-	-	-	0.0424761
H5	국민연금체납 여부	291	-	291	0.0099263
H5_A		-	-	-	0.0134449
H5_D	체납개월수	-	-	-	0.0100339
H5_E	체납보험료	-	-	-	0.0108362
H5_A	월보험료	-	-	-	0.0134449
H6	건강보험료체납 여부	4,036	143	3,893	0.0079238
H6_C	체납개월수	-	-	-	0.0588934
H6_D	체납보험료	-	-	-	0.0971473
H6_E	월보험료	-	-	-	0.0839128
H8_new	의료비과다	-	-	-	0
H9	피부양 의무자 장기요양여부	7,658	6,358	1,300	0.0599277
H10	전세금액기준 이하가구여부	11,010	7,674	3,336	0.0087347
H10_B	세대원수	-	-	-	0.0776052

서울특별시 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H10_A	임차보증금액	-	-	-	0.1017065
H11	월세금액기준 이하가구여부	21,253	16,450	4,803	0.0208877
H11_C	세대원수	-	-	-	0.0470184
H11_R	월세전환총액 변수	-	-	-	0.1645003
H13	고용보험실직 사유대상여부	209	-	209	0.0037413
H15	방문건강집중 관리군여부	26	-	26	0.0031707
H22	시설입퇴소여부	32	-	32	0.0024799
H23	기초생활긴급지 원수급탈락여부	889	-	889	0.0260293
H18	자살예방관리 대상자여부	-	-	-	0
H19	자살시도대상자 여부	15	-	15	0.0001388
H20	위기학생여부	265	-	265	0.0253912
H16	기저귀조제분유 지원대상자여부	701	-	701	0.0199427
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	4	-	4	0.0004473
H14	고용보험비대상 여부	672	-	672	0.0047228
H24	공공임대주택 채납자여부	103	-	103	3.432E-05
H24_C		-	-	-	0.0017844
H25	산재요양종결후 근로단절자여부	7	-	7	0.002896
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	1,371	503	868	0.0055195
H27_A	현재연체금액	-	-	-	0.0433724
H28	의료비용과다 지출가구여부	176	96	80	0.0011433

서울특별시 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H12	고용보험개별연 장급여대상여부	3	-	3	2.706E-05
H17	신생아난청 확진자여부	-	-	-	0
H24_C	임대료체납금액	-	-	-	0.0017844
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0.0085257
합계		-	31,225	19,280	1.015229

다음으로 경기도 모형에 대한 1인가구, 다인가구에 대한 대상자 분포도와 가중치를 살펴본다. <표 4-9>는 1인가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 월세금액기준이하가구여부(H11) 변수로써 16,291명으로 가장 많은 분포를 보여주었으며, 다음으로 전세금액기준이하가구여부 (H10)변수가 3,320명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 월세금액기준이하가구여부 (H11)가 14,355명 그리고 전세금액기준이하가구여부 (H10)변수가 2,346명으로 단독 변수 보유 대상자를 보였다. 중복 변수 기준으로는 총 변수와 동일하게 월세금액기준이하가구여부(H11)이 1,936명, 건강보험료체납여부 (H6)가 1,529명으로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액 (H11_R) 변수가 0.2586, 전세 임차보증금액 (H10_A) 0.1061, 기초생활긴급지원수급탈락여부 (H23) 0.0769 순으로 가중치가 높은 것을 알 수 있다. 이런 결과를 통해 월세 월세 전환총액 변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포되는 것을 알 수 있다.

<표 4-9> 경기도 모형 대상자 분포도와 가중치(1인가구)

경기도 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	32	2	30	0.001036
H1_C		-	-	-	0
H1_D		-	-	-	0
H1_G	체납전체금액	-	-	-	0.0036897
H1_H	월평균사용량	-	-	-	0.0021326
H2	단수도여부	-	-	-	0
H2_D		-	-	-	0
H2_E		-	-	-	0
H2_I		-	-	-	0
H2_F		-	-	-	0
H3	단가스여부	25	-	25	0.0011635
H3_C		-	-	-	0
H3_D		-	-	-	0
H3_E		-	-	-	0
H3_F		-	-	-	0
H4	전기료체납여부	85	3	82	0.0152177
H4_A		-	-	-	0
H4_C	체납개월수	-	-	-	0.0090321
H4_D		-	-	-	0
H4_E	월평균사용량	-	-	-	0.029281
H5	국민연금체납여부	125	-	125	0.0019558
H5_A		-	-	-	0.0007751
H5_D	체납개월수	-	-	-	0.0016876
H5_E	체납보험료	-	-	-	0.0007866
H5_A	월보험료	-	-	-	0.0007751
H6	건강보험료체납여부	1,545	16	1,529	0.0035762
H6_C	체납개월수	-	-	-	0.037956
H6_D	체납보험료	-	-	-	0.0636948
H6_E	월보험료	-	-	-	0.0586665
H8_new	의료비과다	-	-	-	0

경기도 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H9	피부양 의무자 장기요양여부	796	-	796	0.068819
H9_B	월보험료	-	-	-	0
H10	전세금액기준 이하가구여부	3,320	2,346	974	0.0160436
H10_B	세대원수	-	-	-	0.0595387
H10_A	임차보증금액	-	-	-	0.1060918
H11	월세금액기준 이하가구여부	16,291	14,355	1,936	0.0286525
H11_C	세대원수	-	-	-	0.031358
H11_R	월세전환총액 변수	-	-	-	0.258652
H13	고용보험실직 사유대상여부	64	-	64	0.0082527
H15	방문건강집중 관리군여부	29	-	29	0.0506238
H22	시설입퇴소여부	61	-	61	0.0068561
H23	기초생활긴급지 원수급탈락여부	392	-	392	0.0768978
SS1	소득인정액정보	-	-	-	0
SS1_A	중위소득비율	-	-	-	0
region_B	지역(중)	-	-	-	0.0146996
region_C	지역(소)	-	-	-	0.0002912
H18	자살예방관리 대상자여부	6	-	6	0.0003323
H19	자살 시도대상자 여부	4	-	4	0
H20	위기학생여부	-	-	-	0
H16	기저귀조제분유 지원대상자여부	-	-	-	0
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	3	-	3	0
H14	고용보험비대상 여부	634	-	634	0.0078478
H24	공공임대주택채 납자여부	25	-	25	0.0027647

경기도 모형		1인가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H24_C		-	-	-	0.0071625
H25	산재요양종결후 근로단절자여부	15	-	15	0.0013465
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	180	-	180	0.0019399
H27_A	현재연체금액	-	-	-	0.0176807
H27_B	연체등록건수	-	-	-	0
H27_C	연체발생기관수	-	-	-	0
H27_M	대부연체금액포 함여부	-	-	-	0
H28	의료비용과다 지출가구여부	111	-	111	0.0011492
H12	고용보험개별연 장급여대상여부	-	-	-	0
H17	신생아난청 확진자여부	-	-	-	0
H24_C	임대료체납금액	-	-	-	0.0071625
H26	재난피해자여부	-	-	-	0.000384
H28_A	본인부담금총액	-	-	-	0.0019642
		23,743	16,722	7,021	1.007938

<표 4-10>은 다인 가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 1인 가구와 마찬가지로 월세금액기준이하가구여부(H11) 변수로써 23,670명으로 가장 많은 분포를 보여주었으며, 다음으로 전세금액기준이하가구여부 (H10)변수가 9,183명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 월세금액기준이하가구여부 (H11)가 19,004명으로 제일 대상자가 많이 분포된 변수는 1인 가구와 동일하지만, 두 번째로 높은 변수는 피부양의무자장기요양여부 (H9)변수가 5,598명으로 1인 가구와 다른 양상을 보였다. 중복 변수 기준으

로는 건강보험료체납여부(H6)가 5,208명으로 제일 높으며, 월세금액 기준이하가구여부(H11)이 4,666명으로 두 번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액변수 (H11_R) 0.2219, 건강보험료 체납보험료 (H6_D) 0.09589, 건강보험료 월보험료 (H6_E) 순으로 가중치가 높은 것을 알 수 있었다. 이런 결과를 통해 1인가구와 마찬가지로 월세전환총액변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포가 많이 된다는 사실을 알 수 있었다.

<표 4-10> 경기도 모형 대상자 분포도와 가중치(다인가구)

경기도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	116	88	28	0.0001968
H1_C		-	-	-	0
H1_D		-	-	-	0
H1_G	체납전체금액	-	-	-	0.0097153
H1_H	월평균사용량	-	-	-	0.0082471
H2	단수도여부	-	-	-	0
H2_D		-	-	-	0
H2_E		-	-	-	0
H2_I		-	-	-	0
H2_F		-	-	-	0
H3	단가스여부	17	-	17	0.0008811
H3_C		-	-	-	0
H3_D		-	-	-	0
H3_E		-	-	-	0
H3_F		-	-	-	0
H4	전기료체납여부	677	-	677	0.0071839
H4_A		-	-	-	0
H4_C	체납개월수	-	-	-	0.0068719
H4_D		-	-	-	0
H4_E	월평균사용량	-	-	-	0.0289067

경기도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H5	국민연금체납 여부	191	-	191	0.0090998
H5_A		-	-	-	0.0275832
H5_D	체납개월수	-	-	-	0.0013436
H5_E	체납보험료	-	-	-	0.0118016
H5_A	월보험료	-	-	-	0.0275832
H6	건강보험료체납 여부	5,863	655	5,208	0.0045566
H6_C	체납개월수	-	-	-	0.0560464
H6_D	체납보험료	-	-	-	0.0958909
H6_E	월보험료	-	-	-	0.0801998
H8_new	의료비과다	-	-	-	0
H9	피부양 의무자 장기요양여부	6,920	5,598	1,322	0.0779058
H9_B	월보험료	-	-	-	0
H10	전세금액기준 이하가구여부	9,183	5,340	3,843	0.0360442
H10_B	세대원수	-	-	-	0.0224828
H10_A	임차보증금액	-	-	-	0.0706893
H11	월세금액기준 이하가구여부	23,670	19,004	4,666	0.0230758
H11_C	세대원수	-	-	-	0.0636293
H11_R	월세전환총액 변수	-	-	-	0.2218592
H13	고용보험실직 사유대상여부	250	-	250	0.0044252
H15	방문건강집중 관리군여부	12	-	12	0.0015574
H22	시설입퇴소여부	178	-	178	0.0024626
H23	기초생활긴급지 원수급탈락여부	1,502	-	1,502	0.0222146
SS1	소득인정액정보	-	-	-	0
SS1_A	중위소득비율	-	-	-	0
region_B	지역(중)	-	-	-	0.0038664

경기도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
region_C	지역(소)	-	-	-	0.0017644
H18	자살예방관리대상자여부	1	-	1	0
H19	자살시도대상자여부	6	-	6	0
H20	위기학생여부	365	-	365	0.0188295
H16	기저귀조제분유 지원대상자여부	599	-	599	0.0174304
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	3	-	3	0
H14	고용보험비대상 여부	530	-	530	0.0020752
H24	공공임대주택채 납자여부	235	-	235	0.0023667
H24_C		-	-	-	0.0055548
H25	산재요양종결후 근로단절자여부	13	-	13	0.0061164
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	503	-	503	0.0142888
H27_A	현재연체금액	-	-	-	0.0290281
H27_B	연체등록건수	-	-	-	0
H27_C	연체발생기관수	-	-	-	0
H27_M	대부연체금액포 함여부	-	-	-	0
H28	의료비용과다지 출가구여부	120	-	120	0.0006912
H12	고용보험개별연 장급여대상여부	1	-	1	0
H17	신생아난청확진 자여부	-	-	-	0
H24_C	임대료채납금액	-	-	-	0.0055548
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0.0031174
		50,955	30,685	20,270	1.033138

다음으로 전라북도 모형에 대한 1인가구, 다인가구에 대한 대상자 분포도와 가중치를 살펴본다. <표 4-11>은 1인가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 월세금액기준이하가구여부(H11) 변수로써 4,637명으로 가장 많은 분포를 보여주었으며, 다음으로 전세금액기준이하가구여부 (H10)변수가 4,531명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 전세금액기준이하가구여부 (H10)가 4,171명 그리고 월세금액기준이하가구여부 (H11)변수가 3,872명으로 단독 변수 보유 대상자를 보였다. 중복 변수 기준으로는 총 변수와 동일하게 월세금액기준이하가구여부(H11)이 765명, 기초생활기초지원수급탈락여부(H23) 변수가 608명으로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액 (H11_R) 변수가 0.2296, 피부양의무자장기요양여부 (H9) 0.1033, 전기료체납여부 월평균사용량 (H4_E) 0.1029 순으로 가중치가 높은 것을 알 수 있다. 이런 결과를 통해 월세 월세전환총액 변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포되는 것을 알 수 있다.

<표 4-11> 전라북도 모형 대상자 분포도와 가중치(1인가구)

전라북도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	94	72	22	0
H1_C		-	-	-	0
H1_D		-	-	-	0
H1_G	체납전체금액	-	-	-	0
H1_H	월평균사용량	-	-	-	0
H2	단수도여부	-	-	-	0
H2_D		-	-	-	0

전략복도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H2_E		-	-	-	0
H2_I		-	-	-	0
H2_F		-	-	-	0
H3	단가스여부	108	67	41	0
H3_C		-	-	-	0
H3_D		-	-	-	0
H3_E		-	-	-	0
H3_F		-	-	-	0
H4	전기료체납여부	113	42	71	0.0144477
H4_A		-	-	-	0
H4_C	체납개월수	-	-	-	0.0055163
H4_D		-	-	-	0
H4_E	월평균사용량	-	-	-	0.1029495
H5	국민연금체납 여부	62	-	62	0.0040722
H5_A		-	-	-	0.0026473
H5_D	체납개월수	-	-	-	0.0094845
H5_E	체납보험료	-	-	-	0
H5_A	월보험료	-	-	-	0.0026473
H6	건강보험료체납 여부	3,191	2,720	471	0.003655
H6_C	체납개월수	-	-	-	0.0590857
H6_D	체납보험료	-	-	-	0.0470195
H6_E	월보험료	-	-	-	0.075461
H8_new	의료비과다	-	-	-	0
H9	피부양 의무자 장기요양여부	3,516	3,145	371	0.1033461
H9_B	월보험료	-	-	-	0
H10	전세금액기준 이하가구여부	4,531	4,171	360	0
H10_B	세대원수	-	-	-	0.0017539
H10_A	임차보증금액	-	-	-	0.0810761
H11	월세금액기준 이하가구여부	4,637	3,872	765	0.0044884

전략복도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H11_C	세대원수	-	-	-	0.0632164
H11_R	월세전환총액 변수	-	-	-	0.2296383
H13	고용보험실직 사유대상여부	507	434	73	0
H15	방문건강집중 관리군여부	7	-	7	0.0513533
H22	시설입퇴소여부	502	329	173	0.0096492
H23	기초생활긴급지 원수급탈락여부	2,334	1,726	608	0.0679351
SS1	소득인정액정보	-	-	-	0
SS1_A	중위소득비율	-	-	-	0
region_B	지역(중)	-	-	-	0.0181089
region_C	지역(소)	-	-	-	0.005238
H18	자살예방관리 대상자여부	1	1	-	0
H19	자살시도대상자 여부	3	3	-	0
H20	위기학생여부	20	18	2	0
H16	기저귀조제분유 지원대상자여부	2	1	1	0
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	19	16	3	0
H14	고용보험비대상 여부	202	-	202	0.0062984
H24	공공임대주택 체납자여부	15	-	15	0.0115136
H24_C		-	-	-	0.005805
H25	산재요양종결후 근로단절자여부	57	47	10	0
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	132	-	132	0.0048745
H27_A	현재연체금액	-	-	-	0.011366
H27_B	연체등록건수	-	-	-	0

전략복도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H27_C	연체발생기관수	-	-	-	0
H27_M	대부연체금액 포함여부	-	-	-	0
H28	의료비용과다 지출가구여부	240	171	69	0
H12	고용보험개별연 장급여대상여부	-	-	-	0
H17	신생아난청 확진자여부	-	-	-	0
H24_C	임대료체납금액	-	-	-	0.005805
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0
합계		20,293	16,835	3,458	1.008452

<표 4-11>는 다인 가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 1인 가구와 마찬가지로 월세금액기준이하가구여부(H11) 변수로써 16,356명으로 가장 많은 분포를 보여주었으며, 다음으로 전세금액기준이하가구여부 (H10)변수가 10,411명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 1인 가구와는 반대로 월세금액기준이하가구여부(H11)가 13,363명으로 제일 대상자가 많이 분포된 변수가 제일 높으며, 두 번째로 높은 변수는 전세금액기준이하가구여부 (H10)변수가 8,960명으로 1인 가구와 다른 양상을 보였다. 중복 변수 기준으로는 월세금액기준이하가구여부(H11)이 2,993명이 제일 높으며, 건강보험료체납여부 (H6) 2,184명으로 두 번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액변수 (H11_R) 0.2578, 건강보험료 체납보험료 (H6_D) 0.1042, 피부양 의무자장기요양여부 (H9) 순으로 가중치가 높은 것을 알 수 있었다. 이런

결과를 통해 1인가구와 마찬가지로 월세전환총액변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포가 많이 된다는 사실을 알 수 있었다.

<표 4-12> 전라북도 모형 대상자 분포도와 가중치(다인가구)

전라북도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	28	1	27	0
H1_C		-	-	-	0
H1_D		-	-	-	0
H1_G	체납전체금액	-	-	-	0
H1_H	월평균사용량	-	-	-	0.0029215
H2	단수도여부	-	-	-	0
H2_D		-	-	-	0
H2_E		-	-	-	0
H2_I		-	-	-	0
H2_F		-	-	-	0
H3	단가스여부	35	3	32	0
H3_C		-	-	-	0
H3_D		-	-	-	0
H3_E		-	-	-	0
H3_F		-	-	-	0
H4	전기료체납여부	267	-	267	0.01836
H4_A		-	-	-	0
H4_C	체납개월수	-	-	-	0.0018416
H4_D		-	-	-	0
H4_E	월평균사용량	-	-	-	0.0331055
H5	국민연금체납여부	59	-	59	0.0011035
H5_A		-	-	-	0.0532443
H5_D	체납개월수	-	-	-	0.0044554
H5_E	체납보험료	-	-	-	0.0017904

전략복도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H5_A	월보험료	-	-	-	0.0532443
H6	건강보험료체납 여부	4,951	2,767	2,184	0.0028163
H6_C	체납개월수	-	-	-	0.0510448
H6_D	체납보험료	-	-	-	0.1041789
H6_E	월보험료	-	-	-	0.0772653
H8_new	의료비과다	-	-	-	0
H9	피부양 의무자 장기요양여부	3,194	2,798	396	0.0979049
H9_B	월보험료	-	-	-	0
H10	전세금액기준 이하가구여부	10,411	8,960	1,451	0.0091329
H10_B	세대원수	-	-	-	0.0256991
H10_A	임차보증금액	-	-	-	0.0433242
H11	월세금액기준 이하가구여부	16,356	13,363	2,993	0.0162303
H11_C	세대원수	-	-	-	0.0607764
H11_R	월세전환총액 변수	-	-	-	0.2577933
H13	고용보험실직 사유대상여부	113	-	113	0.0076608
H15	방문건강집중 관리군여부	129	117	12	0.0048713
H22	시설입퇴소여부	107	54	53	0.0016467
H23	기초생활긴급지 원수급탈락여부	1,739	-	1,739	0.0207982
SS1	소득인정액정보	-	-	-	0
SS1_A	중위소득비율	-	-	-	0
region_B	지역(중)	-	-	-	0.0158918
region_C	지역(소)	-	-	-	0.0080417
H18	자살예방관리 대상자여부	4	2	2	0

전략복도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H19	자살시도대상자 여부	6	3	3	0
H20	위기학생여부	315	-	315	0.0188545
H16	기저귀조제분유 지원대상자여부	1,682	1,218	464	0.0227009
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	17	12	5	0
H14	고용보험비대상 여부	4,731	4,348	383	0.0044692
H24	공공임대주택 채납자여부	156	44	112	0.0011628
H24_C		-	-	-	0
H25	산재요양종결후 근로단절자여부	8	-	8	0.0026691
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	168	-	168	0.002302
H27_A	현재연체금액	-	-	-	0.0259423
H27_B	연체등록건수	-	-	-	0
H27_C	연체발생기관수	-	-	-	0
H27_M	대부연체금액 포함여부	-	-	-	0
H28	의료비용과다 지출가구여부	246	152	94	0
H12	고용보험개별연 장급여대상여부	-	-	-	0
H17	신생아난청 확진자여부	-	-	-	0
H24_C	임대료채납금액	-	-	-	0
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0
합계		44,722	33,842	10,880	1.053244

다음으로 제주특별자치도 모형에 대한 1인가구, 다인가구에 대한 대상자 분포도와 가중치를 살펴본다. <표 4-13>는 1인가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 건강보험료체납여부 (H6) 변수로써 10,514명으로 가장 많은 분포를 보여주었으며, 다음으로 고용보험비대상여부 (H14)변수가 1,921명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 건강보험료체납여부 (H6)가 8,615명 그리고 고용보험비대상여부 (H14)변수가 1,567명으로 단독 변수 보유 대상자를 보였다. 중복 변수 기준으로는 총 변수와 동일하게 건강보험료체납여부 (H6)가 1,899명으로 제일 높으며, 국민연금 체납여부 (H5)변수가 760명으로 높은 것을 알 수 있었다. 가중치 기준으로 보면 피부양 의무자 장기요양여부 (H9) 변수가 0.7063, 월세 세대원수 (H11_C) 0.2029 순으로 가중치가 높은 것을 알 수 있다. 다른 지역과는 다르게 제주 지역의 경우 전체 인구 수가 작기 때문에 비율적으로 조절하여 대상자를 추출해야 된다는 사실을 알 수 있다.

〈표 4-13〉 제주특별자치도 모형 대상자 분포도와 가중치(1인가구)

제주특별자치도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	76	38	38	0
H1_C		-	-	-	0
H1_D		-	-	-	0
H1_G	체납전체금액	-	-	-	0
H1_H	월평균사용량	-	-	-	0
H2	단수도여부	-	-	-	0
H2_D		-	-	-	0
H2_E		-	-	-	0
H2_I		-	-	-	0
H2_F		-	-	-	0
H3	단가스여부	10	5	5	0
H3_C		-	-	-	0
H3_D		-	-	-	0
H3_E		-	-	-	0
H3_F		-	-	-	0
H4	전기료체납여부	925	694	231	0
H4_A		-	-	-	0
H4_C	체납개월수	-	-	-	0
H4_D		-	-	-	0
H4_E	월평균사용량	-	-	-	0
H5	국민연금체납여부	1,554	794	760	0
H5_A		-	-	-	0
H5_D	체납개월수	-	-	-	0
H5_E	체납보험료	-	-	-	0
H5_A	월보험료	-	-	-	0
H6	건강보험료체납여부	10,514	8,615	1,899	0
H6_C	체납개월수	-	-	-	0
H6_D	체납보험료	-	-	-	0.0717146
H6_E	월보험료	-	-	-	0

제주특별자치도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H8_new	의료비과다	-	-	-	0
H9	피부양의무자장 기요양여부	608	512	96	0.7063203
H9_B	월보험료	-	-	-	0
H10	전세금액기준이 하가구여부	737	632	105	0
H10_B	세대원수	-	-	-	0
H10_A	임차보증금액	-	-	-	0
H11	월세금액기준이 하가구여부	1,279	917	362	0
H11_C	세대원수	-	-	-	0.2029374
H11_R	월세전환총액변 수	-	-	-	0
H13	고용보험실직사 유대상여부	321	222	99	0
H15	방문건강집중관 리군여부	99	93	6	0
H22	시설입퇴소여부	109	65	44	0
H23	기초생활긴급지 원수급탈락여부	471	257	214	0
SS1	소득인정액정보	-	-	-	0
SS1_A	중위소득비율	-	-	-	0
region_B	지역(중)	-	-	-	0
region_C	지역(소)	-	-	-	0
H18	자살예방관리대 상자여부	-	-	-	0
H19	자살시도대상자 여부	18	10	8	0
H20	위기학생여부	3	2	1	0
H16	기저귀조제분유 지원대상자여부	1	1	-	0
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	4	4	-	0
H14	고용보험비대상 여부	1,921	1,567	354	0.0190276

제주특별자치도 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H24	공공임대주택체 납자여부	48	23	25	0
H24_C		-	-	-	0
H25	산재요양종결후 근로단절자여부	25	13	12	0
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	1,310	640	670	0
H27_A	현재연체금액	-	-	-	0
H27_B	연체등록건수	-	-	-	0
H27_C	연체발생기관수	-	-	-	0
H27_M	대부연체금액포 함여부	-	-	-	0
H28	의료비용과다지 출가구여부	109	83	26	0
H12	고용보험개별연 장급여대상여부	-	-	-	0
H17	신생아난청확진 자여부	-	-	-	0
H24_C	임대료체납금액	-	-	-	0
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0
합계		20,142	15,187	4,955	1

<표 4-14>는 다인 가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 1인 가구와 마찬가지로 건강보험료체납여부 (H6) 변수가 13,942명으로 가장 많은 분포를 보여주었으며, 다음으로 고용보험비 대상여부 (H14)가 7,590명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 건강보험료체납여부 (H6)가 12,084명으로 제일 대상자가 많이 분포된 변수가 제일 높으며, 두 번째로 높은 변

수는 고용보험비대상여부 (H14)변수가 6,832명으로 1인 가구와 동일한 분포를 보였다. 중복 변수 기준으로는 건강보험료체납여부 (H6)가 1,858명으로 제일 높으며, 월세금액기준이하가구여부 (H11)가 1,309명으로 두 번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액변수 (H11_R) 0.3035, 건강보험료 체납보험료 (H6_D) 0.128, 건강보험료 월보험료 (H6_E) 0.1165 순으로 가중치가 높은 것을 알 수 있었다. 이런 결과를 통해 월세전환총액변수로 인해 많은 대상자가 월세금액기준이하가구여부로 분포가 많이 된다는 사실을 알 수 있었다.

<표 4-14> 제주특별자치도 모형 대상자 분포도와 가중치(다인가구)

제주특별자치도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	134	106	28	0
H1_C		-	-	-	0
H1_D		-	-	-	0
H1_G	체납전체금액	-	-	-	0
H1_H	월평균사용량	-	-	-	0
H2	단수도여부	-	-	-	0
H2_D		-	-	-	0
H2_E		-	-	-	0
H2_I		-	-	-	0
H2_F		-	-	-	0
H3	단가스여부	15	10	5	0
H3_C		-	-	-	0
H3_D		-	-	-	0
H3_E		-	-	-	0
H3_F		-	-	-	0
H4	전기료체납여부	838	604	234	0
H4_A		-	-	-	0
H4_C	체납개월수	-	-	-	0
H4_D		-	-	-	0

제주특별자치도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H4_E	월평균사용량	-	-	-	0.0160652
H5	국민연금 채납여부	125	-	125	0.0269154
H5_A		-	-	-	0.032993
H5_D	채납개월수	-	-	-	0
H5_E	채납보험료	-	-	-	0
H5_A	월보험료	-	-	-	0.032993
H6	건강보험료채납 여부	13,942	12,084	1,858	0.0043631
H6_C	채납개월수	-	-	-	0.053138
H6_D	채납보험료	-	-	-	0.1280096
H6_E	월보험료	-	-	-	0.1165685
H8_new	의료비과다	-	-	-	0
H9	피부양 의무자 장기요양여부	578	443	135	0.0901708
H9_B	월 보험료	-	-	-	0
H10	전세금액기준 이하가구여부	2,409	2,075	334	0.000743
H10_B	세대원수	-	-	-	0.0174942
H10_A	임차보증금액	-	-	-	0.0120628
H11	월세금액기준이 하가구여부	5,581	4,272	1,309	0.0284582
H11_C	세대원수	-	-	-	0.1048125
H11_R	월세전환총액 변수	-	-	-	0.3034589
H13	고용보험실직 사유대상여부	1,133	845	288	0.0010398
H15	방문건강집중 관리군여부	49	44	5	0.0016131
H22	시설입퇴소여부	94	72	22	0.0004081
H23	기초생활긴급지 원수급탈락여부	2,648	1,843	805	0.0081545
SS1	소득인정액정보	-	-	-	0
SS1_A	중위소득비율	-	-	-	0
region_B	지역(중)	-	-	-	0

제주특별자치도 모형		2인 이상 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
region_C	지역(소)	-	-	-	0
H18	자살예방관리 대상자여부	-	-	-	0
H19	자살시도대상자 여부	80	68	12	0
H20	위기학생여부	721	543	178	0.0099848
H16	기저귀조제분유 지원대상자여부	2,426	2,184	242	0.0087477
H21	범죄피해여부	-	-	-	0
H7	화재피해여부	28	28	-	0
H14	고용보험비대상 여부	7,590	6,832	758	0.0019288
H24	공공임대주택 채납자여부	107	72	35	0
H24_C		-	-	-	0
H25	산재요양종결후 근로단절자여부	56	47	9	0.0028123
H25_D	평균임금	-	-	-	0
H27	금융연체대상자 여부	329	161	168	0.0031176
H27_A	현재연체금액	-	-	-	0.0269402
H27_B	연체등록건수	-	-	-	0
H27_C	연체발생기관수	-	-	-	0
H27_M	대부연체금액 포함여부	-	-	-	0
H28	의료비용과다지 출가구여부	341	271	70	0
H12	고용보험개별연 장급여대상여부	-	-	-	0
H17	신생아난청 확진자여부	1	1	-	0
H24_C	임대료채납금액	-	-	-	0
H26	재난피해자여부	-	-	-	0
H28_A	본인부담금총액	-	-	-	0
합계		39,225	32,605	6,620	1.032993

제3절 지역별 예측 모델 간 성능 비교 분석

본 절에서는 지역별 예측 모형의 성능을 평가하기 위해 분류 행렬 (Confusion Matrix)와 AUC 값 그리고 Lift Chart를 활용하여 제일 우수한 모형을 판별하고자 한다. 분류 행렬의 Cut-off Point 는 0.4 이상으로 한정시킨다. 그리고 모형 간의 성능 비교를 수행할 때 기본 모형과 지역별 모형간의 성능 비교를 수행한다.

우선적으로 기본 모형과 지역별 모형인 서울특별시 모형의 분류행렬과 AUC 값을 비교하고자 한다. <표 4-15>는 1인 가구에서 기본 모형과 서울특별시 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 서울특별시 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 좋으나, 비대상자를 맞추는 비율은 서울특별시 모형에 비해 떨어지는 것을 알 수 있었다. 정확도 측면에서도 서울특별시 모형이 기본 모형에 비해 더 좋은 것을 보여주었다. 하지만 AUC값은 기본 모형이 더욱 더 좋기 때문에 분류행렬과 AUC 기준으로 기본 모형이 서울특별시 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-15> 기본 모형과 서울특별시 모형 분류행렬 및 AUC값 비교(1인가구)

구분	서울 (AREA=1)	
	1인	1인
	기본 모형	서울특별시 모형
TN	841	1646
FP	1382	577
FN	392	937
TP	1123	578
정확도	0.52541466	0.594970572
민감도	0.741254125	0.381518152
특이도	0.378317589	0.740440846
정밀도	0.448303393	0.5004329
AUC	0.5956	0.5851

다음은 1인 가구에서 기존 모형과 서울특별시 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹과 두 번째 그룹을 보면 기본 모형이 7.52, 1.5 서울특별시 모형이 7.13, 1.43으로 기본 모형이 서울특별시 모형에 비해서 약 0.4배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 1인가구 모형에서 기본 모형이 서울특별시 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-16> 기본 모형과 서울특별시 모형 Lift Chart값 비교(1인가구)

구분		기본 모형		서울 특별시 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	187	7.52	1.5	7.13	1.43
2	187	13.4	1.34	12.4	1.24
3	187	19.7	1.32	20.1	1.34
4	187	25	1.25	27.3	1.37
5	187	31.6	1.26	32.2	1.29
6	187	37.8	1.26	37.4	1.25
7	187	43.2	1.23	41.8	1.19
8	187	47.5	1.19	47.5	1.19
9	187	51.3	1.14	52.8	1.17
10	186	55.9	1.12	57.4	1.15
11	187	62.2	1.13	63	1.14
12	187	66.9	1.12	67.3	1.12
13	187	71.9	1.11	70	1.08
14	187	77.4	1.11	73.9	1.06
15	187	82	1.09	78.7	1.05
16	187	86.1	1.08	85.1	1.06
17	187	90	1.06	89.8	1.06
18	187	94.2	1.05	93.7	1.04
19	187	96.9	1.02	97.4	1.02
20	186	100	1	100	1

<표 4-17>는 다인 가구에서 기본 모형과 서울특별시 모형의 분류 행렬과 AUC 값을 보여주고 있다. 1인 가구와는 다르게 다인 가구에서는 기본 모형이 서울특별시 모형에 비해 대상자를 대상자로 잘 맞추는 민감도 측면에서는 약간 떨어지는 것을 알 수 있다. 특이도 측면에서도 약간 떨어진다. 또한 1인 가구와는 반대로 AUC값이 서울특별시 모형이 더욱 더 좋기 때문에 전반적으로 서울특별시 모형이 더욱 더 좋은 것을 알 수 있었다.

〈표 4-17〉 기본 모형과 서울특별시 모형 분류행렬 및 AUC값 비교(다인가구)

구분	서울 (AREA=1)	
	다인 가구	다인 가구
	기본 모형	서울특별시 모형
TN	7422	7336
FP	371	457
FN	2307	2196
TP	394	505
정확도	0.744806556	0.74718887
민감도	0.145871899	0.18696779
특이도	0.952393173	0.941357629
정밀도	0.51503268	0.524948025
AUC	0.6628	0.6742

다음은 다인 가구에서 기존 모형과 서울특별시 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹을 보면 기본 모형이 10.7, 2.14 서울특별시 모형이 12.2, 2.44으로 서울특별시 모형이 기본 모형에 비해 비해서 약 1.5배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 다인 가구 모형에서 서울특별시 모형이 기본 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-18> 기본 모형과 서울특별시 모형 Lift Chart값 비교(다인가구)

구분		기본 모형		서울특별시 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	525	10.7	2.14	12.2	2.44
2	525	18.7	1.87	19.8	1.98
3	525	25.6	1.71	27	1.8
4	524	33	1.65	33.9	1.69
5	525	39.1	1.56	40.6	1.62
6	525	45.5	1.52	46.8	1.56
7	524	51.1	1.46	52.2	1.49
8	525	56.7	1.42	57.6	1.44
9	525	62	1.38	62.9	1.4
10	524	66.7	1.33	67.3	1.35
11	525	71.4	1.3	72.3	1.32
12	525	75.7	1.26	76.8	1.28
13	525	80.2	1.23	81.1	1.25
14	524	84	1.2	85	1.21
15	525	87.9	1.17	88.3	1.18
16	525	91.4	1.14	91.6	1.14
17	524	94.7	1.11	94.9	1.12
18	525	97.4	1.08	97.5	1.08
19	525	99	1.04	99.1	1.04
20	524	100	1	100	1

<표 4-19>는 1인 가구에서 기본 모형과 경기도 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 경기도 모형에 비해 대상자를 대상자로 잘 맞추는 민감도 측면에서는 좋으나, 비대상자를 맞추는 비율은 경기도 모형에 비해 떨어지는 것을 알 수 있었다. 정확도 측면에서도 경기도 모형이 기본 모형에 비해 더 좋은 것을 보여주었다. 하지만 AUC값은 기본 모형이 더욱 더 좋기 때문에 분류행렬과 AUC 기준으로 기본 모형이 경기도 모형에 비해 더 좋은 것을 알 수 있다.

〈표 4-19〉 기본 모형과 경기도 모형 분류행렬 및 AUC값 비교(1인가구)

구분	경기도 군집	
	1인	1인
	기본 모형	경기도 모형
TN	1199	3096
FP	2431	534
FN	464	1752
TP	1876	588
정확도	0.515075377	0.617085427
민감도	0.801709402	0.251282051
특이도	0.33030303	0.852892562
정밀도	0.435570002	0.524064171
AUC	0.6071	0.6063

다음은 1인 가구에서 기존 모형과 경기도 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹과 두 번째 그룹을 보면 기본 모형이 7.56, 1.51 경기도 모형이 7.14, 1.43으로 기본 모형이 경기도 모형에 비해서 약 0.42배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 1인가구 모형에서 기본 모형이 경기도 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-20> 기본 모형과 경기도 모형 Lift Chart값 비교(1인가구)

구분		기본 모형		경기도 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	299	7.56	1.51	7.14	1.43
2	298	14	1.4	14.2	1.42
3	299	20.1	1.34	20.3	1.35
4	298	26.1	1.3	26.3	1.32
5	299	32.2	1.29	32.4	1.3
6	298	38	1.27	37.4	1.25
7	299	43.3	1.24	43.5	1.24
8	298	47.1	1.18	48.8	1.22
9	299	51.8	1.15	52.2	1.16
10	298	56.7	1.13	56.9	1.14
11	299	61.4	1.12	61.8	1.12
12	298	66.7	1.11	66.5	1.11
13	299	72.6	1.12	71.8	1.1
14	298	78.2	1.12	77.6	1.11
15	299	83	1.11	82.3	1.1
16	298	88.2	1.1	87.7	1.1
17	299	92.4	1.09	91.9	1.08
18	298	95.7	1.06	95.1	1.06
19	299	97.5	1.03	97.3	1.02
20	298	100	1	100	1

<표 4-21>는 다인 가구에서 기본 모형과 경기도 모형의 분류행렬과 AUC 값을 보여주고 있다. 1인 가구와는 다르게 다인 가구에서는 기본 모형이 서울특별시 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 0.17 정도 떨어지는 것을 알 수 있다. 반면 특이도 측면에서는 0.11 우수하다. AUC 측면에서는 경기도 모형이 약간 더 나은 것을 알 수 있다. 이와 같은 결과를 통해서 경기도 모형이 기본 모형에 비해 우수하다는 것을 알 수 있었다.

〈표 4-21〉 기본 모형과 경기도 모형 분류행렬 및 AUC값 비교(다인가구)

구분	경기도 군집	
	다인 가구	다인 가구
	기본 모형	경기도 모형
TN	13028	11515
FP	690	2203
FN	4977	3981
TP	686	1682
정확도	0.707600227	0.680924617
민감도	0.121137206	0.297015716
특이도	0.949701123	0.839408077
정밀도	0.498546512	0.432947233
AUC	0.6463	0.6503

다음은 다인 가구에서 기존 모형과 경기도 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹은 비슷한 경향을 보이지만, 두 번째 그룹의 경우 기본 모형의 누적 반응률과 누적 Lift는 16.1, 1.61 경기도 모형이 16.4, 1.64로 경기도 모형이 기본 모형에 비해 비해서 약 0.3배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 다인 가구 모형에서 경기도 모형이 기본 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-22> 기본 모형과 경기도 모형 Lift Chart값 비교(다인가구)

구분		기본 모형		경기도 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	970	8.97	1.79	8.9	1.78
2	969	16.1	1.61	16.4	1.64
3	969	22.4	1.5	23.1	1.54
4	969	29.1	1.46	29.6	1.48
5	969	34.9	1.4	36.4	1.46
6	969	41.7	1.39	42.3	1.41
7	969	48	1.37	48.1	1.37
8	969	54.2	1.36	54.3	1.36
9	969	59.8	1.33	59.6	1.33
10	969	64.8	1.3	64.9	1.3
11	969	70	1.27	70.1	1.27
12	969	74.8	1.25	75	1.25
13	969	79.1	1.22	79.5	1.22
14	969	83.5	1.19	83.8	1.2
15	969	87.7	1.17	87.9	1.17
16	969	91.2	1.14	91.5	1.14
17	969	94.3	1.11	94.8	1.12
18	969	97.4	1.08	97.1	1.08
19	969	99.1	1.04	99	1.04
20	969	100	1	100	1

<표 4-23>는 1인 가구에서 기본 모형과 전라북도 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 전라북도 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 매우 좋으나, 반대로

비대상자를 맞추는 비율은 전라북도 모형에 비해 떨어지는 것을 알 수 있었다. 정확도 측면에서는 기본 모형이 전라북도 모형에 비해 더 좋은 것을 보여주었다. AUC값 또한 기본 모형이 전라북도 모형에 비해 더 좋기 때문에 분류행렬과 AUC 기준으로 기본 모형이 전라북도 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-23> 기본 모형과 전라북도 모형 분류행렬 및 AUC값 비교(1인가구)

구분	전라북도 군집	
	1인	1인
	기본 모형	전라북도 모형
TN	213	965
FP	776	24
FN	114	963
TP	918	69
정확도	0.559623949	0.511627907
민감도	0.889534884	0.066860465
특이도	0.21536906	0.975733064
정밀도	0.541912633	0.741935484
AUC	0.5872	0.5793

다음은 1인 가구에서 기존 모형과 전라북도 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹을 보면 기본 모형이 7.75, 1.55 전라북도 모형이 7.07, 1.41으로 기본 모형이 경기도 모형에 비해서 약 0.68배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 1인가구 모형에서 기본 모형이 전라북도 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-24> 기본 모형과 전라북도 모형 Lift Chart값 비교(1인가구)

구분		기본 모형		전라북도 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	102	7.75	1.55	7.07	1.41
2	101	13.2	1.32	12.3	1.23
3	101	19.1	1.27	17.9	1.2
4	101	24.4	1.22	21.1	1.06
5	101	29.5	1.18	26.6	1.07
6	101	32.8	1.09	31.5	1.05
7	101	38.7	1.1	37.1	1.06
8	101	44.8	1.12	44.6	1.11
9	101	51.2	1.14	48.4	1.08
10	101	55.1	1.1	53.2	1.06
11	101	58.6	1.07	59.5	1.08
12	101	62.9	1.05	65.2	1.09
13	101	67.7	1.04	70.8	1.09
14	101	73.2	1.05	75.1	1.07
15	101	79.9	1.07	80.4	1.07
16	101	85.9	1.07	84.9	1.06
17	101	90.2	1.06	89.8	1.06
18	101	94.2	1.05	93.8	1.04
19	101	97	1.02	97.1	1.02
20	101	100	1	100	1

<표 4-25>는 다인 가구에서 기본 모형과 전라북도 모형의 분류행렬과 AUC 값을 보여주고 있다. 1인 가구와 마찬가지로 다인 가구에서도 기본 모형이 전라북도 모형에 비해 대상자를 대상자로 잘 맞추

는 민감도 측면에서는 우수하다는 것을 알 수 있다. 반면 특이도 측면에서는 전라북도 모형이 0.01 정도 우수하게 나왔다. AUC 측면에서는 기본 모형이 약간 더 나은 것을 알 수 있다. 이와 같은 결과를 통해서 기본 모형이 전라북도 모형에 비해 우수하다는 것을 알 수 있었다.

〈표 4-25〉 기본 모형과 전라북도 모형 분류행렬 및 AUC값 비교(다인가구)

구분	전라북도 군집	
	다인가구	다인가구
	기본 모형	전라북도 모형
TN	3278	3292
FP	211	197
FN	1456	1481
TP	234	209
정확도	0.67812319	0.675999228
민감도	0.138461538	0.123668639
특이도	0.939524219	0.94353683
정밀도	0.525842697	0.514778325
AUC	0.6423	0.6303

다음은 다인 가구에서 기존 모형과 전라북도 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹에서 기본 모형의 누적 반응률과 누적 Lift는 9.35, 1.87 전라북도 모형이 8.52, 1.7로 기본 모형이 전라북도 모형에 비해 비해서 약 0.83배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 다인 가구 모형에서 기본 모형이 전라북도 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-26> 기본 모형과 전략복도 모형 Lift Chart값 비교(다인가구)

구분		기본 모형		전략복도 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	259	9.35	1.87	8.52	1.7
2	259	15.9	1.59	14.7	1.47
3	259	22	1.46	21.3	1.42
4	259	29.3	1.47	27.5	1.38
5	259	34.1	1.37	33.6	1.34
6	259	41.2	1.37	40	1.33
7	259	46.5	1.33	45.7	1.31
8	259	52	1.3	51.4	1.28
9	259	57.6	1.28	57.2	1.27
10	259	63.4	1.27	62.2	1.24
11	259	69.1	1.26	68.2	1.24
12	259	73.7	1.23	73.4	1.22
13	259	78.3	1.2	77.5	1.19
14	259	82.5	1.18	81	1.16
15	259	86.6	1.15	85.7	1.14
16	259	90	1.12	89.9	1.12
17	259	93.5	1.1	93.4	1.1
18	259	96.7	1.07	96	1.07
19	259	98.8	1.04	98.3	1.04
20	258	100	1	100	1

<표 4-27>는 1인 가구에서 기본 모형과 제주도 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 제주도 모형에 비해 대상자를 대상자로 잘 맞추는 민감도 측면에서는 매우 좋으나, 반대로 비대상

자를 맞추는 비율은 제주도 모형에 비해 떨어지는 것을 알 수 있었다. 정확도 측면에서는 제주도 모형이 기본 모형에 비해 더 좋은 것을 보여주었다. AUC값은 기본 모형이 제주도 모형에 비해 더 좋기 때문에 분류행렬과 AUC 기준으로 기본 모형이 제주도 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-27> 기본 모형과 제주특별자치도 모형 분류행렬 및 AUC값 비교(1인가구)

구분	제주특별자치도 군집	
	1인	1인
	기본 모형	제주도 모형
TN	81	187
FP	217	111
FN	31	103
TP	184	112
정확도	0.516569201	0.582846004
민감도	0.855813953	0.520930233
특이도	0.271812081	0.627516779
정밀도	0.458852868	0.502242152
AUC	0.604	0.5948

다음은 1인 가구에서 기존 모형과 제주도 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹을 보면 기본 모형이 6.05, 1.21 제주도 모형이 4.19, 0.837으로 기본 모형이 제주도 모형에 비해서 약 1.86배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 1인가구 모형에서 기본 모형이 제주도 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-28> 기본 모형과 제주특별자치도 모형 Lift Chart값 비교(1인가구)

구분		기본 모형		제주도 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	26	6.05	1.21	4.19	0.837
2	26	10.7	1.07	9.3	0.93
3	25	17.2	1.15	15.3	1.02
4	26	23.3	1.16	25.6	1.28
5	26	27.9	1.12	32.1	1.28
6	25	34	1.13	39.5	1.32
7	26	41.4	1.18	44.7	1.28
8	26	50.2	1.26	48.8	1.22
9	25	52.1	1.16	52.1	1.16
10	26	57.2	1.14	54	1.08
11	26	61.9	1.12	60	1.09
12	25	65.6	1.09	64.2	1.07
13	26	72.1	1.11	68.8	1.06
14	26	80.5	1.15	74	1.06
15	25	83.7	1.12	81.9	1.09
16	26	87.9	1.1	87.4	1.09
17	26	92.1	1.08	92.6	1.09
18	25	95.8	1.06	94.9	1.05
19	26	97.2	1.02	97.7	1.03
20	25	100	1	100	1

<표 4-29>는 다인 가구에서 기본 모형과 제주도 모형의 분류행렬과 AUC 값을 보여주고 있다. 1인 가구와 마찬가지로 다인 가구에서도 기본 모형이 제주도 모형에 비해 대상자를 대상자로 잘 맞추는 민

감도 측면에서는 우수하다는 것을 알 수 있다. 반면 특이도 측면에서는 제주도 모형이 0.04 정도 우수하게 나왔다. AUC 측면에서는 제주도 모형이 약간 더 나은 것을 알 수 있다. 이와 같은 결과를 통해서 1인 가구와 마찬가지로 다인 가구에서도 기본 모형이 제주도 모형에 비해 우수하다는 것을 알 수 있었다.

〈표 4-29〉 기본 모형과 제주특별자치도 모형 분류행렬 및 AUC값 비교(다인가구)

구분	제주특별자치도 군집	
	다인 가구	다인 가구
	기본 모형	제주도 모형
TN	1214	1261
FP	67	20
FN	386	436
TP	60	10
정확도	0.737695426	0.735958309
민감도	0.134529148	0.022421525
특이도	0.947697112	0.984387198
정밀도	0.472440945	0.333333333
AUC	0.6088	0.6101

다음은 다인 가구에서 기존 모형과 제주도 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹을 살펴보면 누적 반응률과 누적 Lift는 9.87, 1.97 제주도 모형이 5.83, 1.17로 기본 모형이 제주도 모형에 비해 비해서 약 4.04배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로

판단하였을 때 다인 가구 모형에서 기본 모형이 제주도 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-30> 기본 모형과 제주특별자치도 모형 Lift Chart값 비교(다인가구)

구분		기본 모형		제주도 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	87	9.87	1.97	5.83	1.17
2	86	17	1.7	14.6	1.46
3	87	22.9	1.52	21.5	1.43
4	86	29.1	1.46	27.6	1.38
5	86	33.6	1.35	34.8	1.39
6	87	38.6	1.29	41	1.37
7	86	43.7	1.25	46.4	1.33
8	86	49.3	1.23	53.1	1.33
9	87	56.5	1.26	57.6	1.28
10	86	61.7	1.23	61	1.22
11	86	66.8	1.21	65.5	1.19
12	87	73.5	1.23	71.7	1.2
13	86	76.7	1.18	76.5	1.18
14	86	80.5	1.15	80.3	1.15
15	87	83.9	1.12	84.8	1.13
16	86	87.2	1.09	87.4	1.09
17	86	90.4	1.06	90.8	1.07
18	87	93.7	1.04	94.4	1.05
19	86	98.4	1.04	98.4	1.04
20	86	100	1	100	1

제4절 파생 변수 투입에 따른 효과 분석

본 절에서 지역 파생 변수가 추가 된 지역 파생 모형에 대해서 살펴본다. 추가되는 파생 변수는 지역변수로서 시군구 지역변수로 구성된 <표 4-31>과 같이 볼 수 있다.

<표 4-31> 지역파생 모형(지역변수)

시군구	값	시군구	값
서울특별시	1	강원도	10
부산광역시	2	충청북도	11
대구광역시	3	충청남도	12
인천광역시	4	전라북도	13
광주광역시	5	전라남도	14
대전광역시	6	경상북도	15
울산광역시	7	경상남도	16
세종특별자치시	8	제주특별자치도	17
경기도	9		

우선적으로 파생 변수를 넣었을 때의 대상자 분포도와 가중치에 대해서 살펴본다. 대상자 분포도와 가중치는 2절에서 살펴본 것과 같이 1인 가구와 다인가구로 분리해서 살펴본다.

<표 4-32>는 1인 가구에 대한 지역 파생 모형 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 월세금액기준이하가구여부 (H11) 변수로써 15,153명으로 가장 많은 분포를 보여주었으며, 다음으로 피부양 의무자장기요양여부 (H9)변수가 5,405명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 월세금액기준이하가구여부 (H11) 변수가 12,805명 그리고 피부양 의무자장기요양여부 (H9)변수가 3,168

명으로 단독 변수 보유 대상자를 보였다. 중복 변수 기준으로는 총 변수와 동일하게 월세금액기준이하가구여부 (H11)가 2,348명으로 제일 높으며, 피부양의무자장기요양여부 (H9)변수가 3,168명으로 두번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 피부양의무자장기요양여부 (H9) 변수가 0.1647로 제일 높으며, 두 번째로 새로 추가된 지역 파생변수 (AREA)가 0.1233 순으로 가중치가 높은 것을 알 수 있다. 이런 결과를 통해 대상자 선정하는데 있어 새로 추가된 지역 파생변수의 역할이 크다는 것을 알 수 있었다.

<표 4-32> 지역파생 모형 대상자 분포도와 가중치(1인가구)

지역파생 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	22	0	22	0
H1_C		0	0	0	0
H1_D		0	0	0	0
H1_G	체납전체금액	0	0	0	0
H1_H	월평균사용량	0	0	0	0.000417
H2	단수도여부	0	0	0	0
H2_D		0	0	0	0
H2_E		0	0	0	0
H2_I		0	0	0	0
H2_F		0	0	0	0
H3	단가스여부	46	0	46	0.000188
H3_C		0	0	0	0
H3_D		0	0	0	0
H3_E		0	0	0	0
H3_F		0	0	0	0
H4	전기료체납 여부	36	0	36	0.014543
H4_A		0	0	0	0
H4_C	체납개월수	0	0	0	0.020726

지역파생 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H4_D		0	0	0	0
H4_E	월평군사용량	0	0	0	0.022615
H5	국민연금체납 여부	0	0	0	0.002382
H5_A		0	0	0	0.005898
H5_D	체납개월수	0	0	0	0
H5_E	체납보험료	0	0	0	0.000844
H5_A	월 보험료	0	0	0	0.005898
H6	건강보험료 체납여부	646	0	646	0.003826
H6_C	체납개월수	0	0	0	0.009857
H6_D	체납보험료	0	0	0	0.063796
H6_E	월 보험료	0	0	0	0.022558
H8_new	의료비과다	0	0	0	0
H9	피부양의무자 장기요양여부	5,405	3,168	2,237	0.164701
H9_B	월 보험료	0	0	0	0
H10	전세금액기준 이하가구여부	626	172	454	0.003399
H10_B	세대원수	0	0	0	0.028904
H10_A	임차보증금액	0	0	0	0.080616
H11	월세금액기준 이하가구여부	15,153	12,805	2,348	0.028578
H11_C	세대원수	0	0	0	0.050004
H11_R	월세전환총액 변수	0	0	0	0.108137
H13	고용보험실직 사유대상여부	175	0	175	0.000814
H15	방문건강집중 관리군여부	31	0	31	0.070733
H22	시설입퇴소 여부	180	0	180	0.01129
H23	기초생활긴급 지원수급탈락 여부	1,100	0	1,100	0.117059

지역파생 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
SS1	소득인정액 정보	0	0	0	0
SS1_A	중위소득비율	0	0	0	0
region_B	지역(중)	0	0	0	0.012034
region_C	지역(소)	0	0	0	0.015919
H18	자살예방관리 대상자여부	6	0	6	0
H19	자살시도대상 자여부	4	0	4	0
H20	위기학생여부	1	0	1	0
H16	기저귀조제분 유지원대상자 여부	0	0	0	0
H21	범죄피해여부	0	0	0	0
H7	화재피해여부	1	0	1	0
H14	고용보험비 대상여부	663	0	663	0.004176
H24	공공임대주택 채납자여부	9	0	9	0.001849
H24_C		0	0	0	0.002876
H25	산재요양종결후 근로단절자여부	14	0	14	0.000184
H25_D	평균임금	0	0	0	0
H27	금융연체대상 자여부	45	0	45	0.002428
H27_A	현재연체금액	0	0	0	0.004793
H27_B	연체등록건수	0	0	0	0
H27_C	연체발생 기관수	0	0	0	0
H27_M	대부연체금액 포함여부	0	0	0	0
H28	의료비용과다 지출가구여부	139	0	139	0.000194
H12	고용보험개별 연장급여대상 여부	0	0	0	0

지역파생 모형		1인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H17	신생아난청확 진자여부	0	0	0	0
H24_C	임대료 채납금액	0	0	0	0.002876
H26	재난피해자 여부	0	0	0	0
H28_A	본인부담금 총액	0	0	0	0.000394
AREA	지역 파생변수	0	0	0	0.123269
합계		24,302	16,145	8,157	1.008774

<표 4-33>는 다인 가구에 대한 대상자 분포도와 가중치를 보여주고 있다. 총 변수 기준으로 대상자 분포에서 가장 많은 분포를 가지는 변수는 1인 가구와 마찬가지로 월세금액기준이하가구여부(H11) 변수로써 17,812명으로 가장 많은 분포를 보여주었으며, 다음으로 피부양 의무자 장기요양 여부 (H9) 변수가 16,921명으로 두 번째 높은 대상자 분포를 보여주었다. 단독 변수 기준으로는 1인 가구와는 반대로 피부양 의무자 장기요양 여부(H9)가 13,614명으로 제일 대상자가 많이 분포된 변수가 제일 높으며, 두 번째로 높은 변수는 월세금액기준이하가구여부(H11) 변수가 11,840명으로 1인 가구와 다른 양상을 보였다. 중복 변수 기준으로는 월세금액기준이하가구여부(H11)이 5,972명이 제일 높으며, 건강보험료채납 여부 (H6) 4,696명으로 두 번째로 높은 것을 알 수 있었다. 가중치 기준으로 보면 월세전환총액변수 (H11_R) 0.1793, 피부양 의무자 장기요양 여부 (H9) 0.09, 지역 파생변수 (AREA) 0.083 순으로 가중치가 높은 것을 알 수 있었다. 이런 결과를 통해 1인 가구와 마찬가지로 새로 생성된 지역 파생변수가 대상자를 분리하

는데 있어 큰 역할을 하는 것을 볼 수 있었다.

<표 4-33> 지역파생 모형 대상자 분포도와 가중치(다인가구)

지역파생 모형		다인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H1	단전여부	41	2	39	0.00114
H1_C		0	0	0	0
H1_D		0	0	0	0
H1_G	체납전체금액	0	0	0	0.004959
H1_H	월평균사용량	0	0	0	0.005864
H2	단수도여부	0	0	0	3.12E-05
H2_D		0	0	0	0
H2_E		0	0	0	0
H2_I		0	0	0	0
H2_F		0	0	0	0
H3	단가스여부	71	0	71	0.000826
H3_C		0	0	0	0
H3_D		0	0	0	0
H3_E		0	0	0	0
H3_F		0	0	0	0
H4	전기료체납 여부	574	2	572	0.006101
H4_A		0	0	0	0
H4_C	체납개월수	0	0	0	0.007791
H4_D		0	0	0	0
H4_E	월평균사용량	0	0	0	0.019575
H5	국민연금체납 여부	356	0	356	0.001205
H5_A		0	0	0	0.015994
H5_D	체납개월수	0	0	0	0.004653
H5_E	체납보험료	0	0	0	0.024419
H5_A	월보험료	0	0	0	0.015994
H6	건강보험료체 납여부	4,696	0	4,696	0.002749
H6_C	체납개월수	0	0	0	0.036722

지역파생 모형		다인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
H6_D	채납보험료	0	0	0	0.073026
H6_E	월보험료	0	0	0	0.059096
H8_new	의료비과다	0	0	0	0
H9	피부양의무자 장기요양여부	16,921	13,614	3,307	0.090089
H9_B	월보험료	0	0	0	0
H10	전세금액기준 이하가구여부	5,958	2,642	3,316	0.014686
H10_B	세대원수	0	0	0	0.030569
H10_A	임차보증금액	0	0	0	0.07198
H11	월세금액기준 이하가구여부	17,812	11,840	5,972	0.060145
H11_C	세대원수	0	0	0	0.044561
H11_R	월세전환총액 변수	0	0	0	0.179326
H13	고용보험실직 사유대상여부	160	0	160	0.003584
H15	방문건강집중 관리군여부	102	0	102	0.003226
H22	시설입퇴소 여부	233	0	233	0.003411
H23	기초생활긴급 지원수급탈락 여부	3,140	0	3,140	0.031669
SS1	소득인정액 정보	0	0	0	0
SS1_A	중위소득비율	0	0	0	0
region_B	지역(중)	0	0	0	0.012096
region_C	지역(소)	0	0	0	0.0086
H18	자살예방관리 대상자여부	16	0	16	0.000769
H19	자살시도대상 자여부	16	0	16	9.64E-05
H20	위기학생여부	1,307	0	1,307	0.020766
H16	기저귀조제분	1,684	0	1,684	0.024954

지역파생 모형		다인 가구			
		총합(a) (소득인정제외)	단독(b) (소득인정제외)	중복(c) (a-b)	가중치
	유지원대상자 여부				
H21	범죄피해여부	0	0	0	0
H7	화재피해여부	3	0	3	0
H14	고용보험비대 상여부	478	0	478	0.0036
H24	공공임대주택 채납자여부	191	0	191	0.000593
H24_C		0	0	0	0.005325
H25	산재요양종결 후근로단절자 여부	8	0	8	0.004125
H25_D	평균임금	0	0	0	0
H27	금융연체대상 자여부	682	0	682	0.015285
H27_A	현재연체금액	0	0	0	0.01887
H27_B	연체등록건수	0	0	0	0
H27_C	연체발생 기관수	0	0	0	0
H27_M	대부연체금액 포함여부	0	0	0	0
H28	의료비용과다 지출가구여부	248	0	248	0.000345
H12	고용보험개별 연장급여대상 여부	3	0	3	0
H17	신생아난청 확진자여부	0	0	0	0
H24_C	임대료채납 금액	0	0	0	0.005325
H26	재난피해자 여부	0	0	0	0.000426
H28_A	본인부담금 총액	0	0	0	0.004241
AREA	지역 파생변수	0	0	0	0.082512
합계		54,700	28,100	26,600	1.021318

다음은 분류행렬과 AUC, 그리고 Lift Chart를 통하여 새로운 지역 파생 모형을 평가한다. 평가를 수행할 때 3절에서 지역별 모형과의 비교에서 우수한 모형과의 비교를 수행하고자 한다.

우선적으로 서울특별시 데이터 군집의 기본 모형과 서울특별시 모형 비교에서 기본 모형이 우수한 것으로 판단되어서 기본 모형과 지역 파생 모형과 비교를 수행한다.

<표 4-34>는 1인 가구에서 기본 모형과 지역 파생 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 지역 파생 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 매우 좋으나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 떨어지는 것을 알 수 있었다. 정확도 측면에서는 지역 파생 모형이 기본 모형에 비해 더 좋은 것을 보여주었다. AUC값은 지역 파생 모형이 기본 모형에 비해 더 좋기 때문에 분류행렬과 AUC 기준으로 지역 파생 모형이 기본 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-34> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(서울시 1인가구)

구분	서울특별시 군집	
	1인 가구	1인 가구
	기본 모형	지역 파생 모형
TN	841	1252
FP	1382	971
FN	392	674
TP	1123	841
정확도	0.52541466	0.559925094
민감도	0.741254125	0.555115512
특이도	0.378317589	0.563202879
정밀도	0.448303393	0.464128035
AUC	0.5956	0.5963

다음은 1인 가구에서 기본 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 기본 모형과 지역 파생 모형이 비슷한 결과를 보여주었다. 반면에 두 번째 그룹의 경우 기본 모형이 13.4, 1.34 지역 파생 모형이 13.7, 1.37으로 지역 파생 모형이 기본 모형에 비해서 약 0.3 배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같이 전반적으로 판단하였을 때 1인 가구에서 지역 파생 모형이 기본 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-35> 기본 모형과 지역파생 모형 Lift Chart값 비교 (서울시 1인가구)

구분		기본 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	187	7.52	1.5	7.46	1.49
2	187	13.4	1.34	13.7	1.37
3	187	19.7	1.32	19.6	1.31
4	187	25	1.25	25.6	1.28
5	187	31.6	1.26	31.4	1.26
6	187	37.8	1.26	38	1.27
7	187	43.2	1.23	43.2	1.23
8	187	47.5	1.19	47.9	1.2
9	187	51.3	1.14	52.1	1.16
10	186	55.9	1.12	57.1	1.14
11	187	62.2	1.13	62.3	1.13
12	187	66.9	1.12	65.9	1.1
13	187	71.9	1.11	70.7	1.09
14	187	77.4	1.11	76.9	1.1
15	187	82	1.09	81.1	1.08
16	187	86.1	1.08	86.3	1.08
17	187	90	1.06	90.2	1.06
18	187	94.2	1.05	93.6	1.04
19	187	96.9	1.02	97.1	1.02
20	186	100	1	100	1

<표 4-36>는 다인 가구에서 기본 모형에 비해 우수한 서울특별시 모형과 지역 파생 모형과의 비교를 한다. 서울특별시 모형이 지역 파생 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 약

간 높으나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 약간 낮은 것을 알 수 있었다. 정확도 측면에서는 서울특별시 모형이 약간 좋은 결과를 보였다. AUC값 또한 서울특별시 모형이 지역 파생 모형에 비해 우수한 결과를 보였다.

<표 4-36> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(서울시 다인가구)

구분	서울특별시 군집	
	다인 가구	다인 가구
	서울특별시 모형	지역 파생 모형
TN	7336	7391
FP	457	402
FN	2196	2279
TP	505	422
정확도	0.74718887	0.744520678
민감도	0.18696779	0.15623843
특이도	0.941357629	0.948415244
정밀도	0.524948025	0.512135922
AUC	0.6742	0.662

다음은 다인 가구에서 서울특별시 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 서울특별시 모형의 값은 12.2, 2.44이며 지역 파생 모형의 경우 11.2, 2.24로 서울특별시 모형이 약 1% 배 정도 좋은 결과를 보여주었다. 이와 같은 결과를 전반적으로 판단하였을 때 다인 가구에서 서울특별시 모형이 지역 파생 모형에 비해서 우수하다는 것을 알 수 있었다.

<표 4-37> 서울특별시 모형과 지역파생 모형 Lift Chart값 비교 (다인가구)

구분		서울특별시 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	525	12.2	2.44	11.2	2.24
2	525	19.8	1.98	18.7	1.87
3	525	27	1.8	25.4	1.7
4	524	33.9	1.69	32.7	1.64
5	525	40.6	1.62	39.2	1.57
6	525	46.8	1.56	45.5	1.52
7	524	52.2	1.49	50.9	1.46
8	525	57.6	1.44	56.4	1.41
9	525	62.9	1.4	60.6	1.35
10	524	67.3	1.35	65.9	1.32
11	525	72.3	1.32	70.9	1.29
12	525	76.8	1.28	76.3	1.27
13	525	81.1	1.25	80.7	1.24
14	524	85	1.21	84.4	1.21
15	525	88.3	1.18	87.7	1.17
16	525	91.6	1.14	91	1.14
17	524	94.9	1.12	94.9	1.12
18	525	97.5	1.08	97.5	1.08
19	525	99.1	1.04	99.3	1.05
20	524	100	1	100	1

<표 4-38>는 경기도 데이터 기반으로 1인 가구에서 기본 모형과 지역 파생 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 지역 파생 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 약간 좋으나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 떨어지는 것을 알 수 있었다. 정확도 측면에서는 지역 파생 모형이 기본 모형에 비해 더 좋은 결과를 보여주었다. AUC값은 지역 파생 모형이 기본 모형에 비해 더 좋기 때문에 분류행렬과 AUC 기준으로 지역 파생 모형이 기본 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-38> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(경기도 1인가구)

구분	경기도 군집	
	1인 가구	1인 가구
	기본 모형	지역 파생 모형
TN	1199	1206
FP	2431	2424
FN	464	470
TP	1876	1870
정확도	0.515075377	0.515242881
민감도	0.801709402	0.799145299
특이도	0.33030303	0.332231405
정밀도	0.435570002	0.435491383
AUC	0.6071	0.6106

다음은 1인 가구에서 기본 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 기본 모형과 지역 파생 모형이 비슷한 결과를 보여주었다. 반면에 두 번째 그룹의 경우 기본 모형이 14, 1.4 지역 파생 모형이 13.8, 1.38으로 기본 모형이 지역 파생 모형에 비해서 약 0.02%배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 하지만 세 번째 그룹의 경우 지역 파생 모형이 같은 비율로 더 좋은 결과를 보인다. 이와 같은 결과를 통해 기본 모형과 지역 파생 모형의 성능이 거의 비슷하다는 것을 알 수 있다.

<표 4-39> 기본 모형과 지역파생 모형 Lift Chart값 비교 (경기도 1인가구)

구분		기본 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	299	7.56	1.51	7.56	1.51
2	298	14	1.4	13.8	1.38
3	299	20.1	1.34	20.4	1.36
4	298	26.1	1.3	26.4	1.32
5	299	32.2	1.29	32.5	1.3
6	298	38	1.27	38.2	1.27
7	299	43.3	1.24	43.7	1.25
8	298	47.1	1.18	48.5	1.21
9	299	51.8	1.15	52.4	1.16
10	298	56.7	1.13	57.1	1.14
11	299	61.4	1.12	62.2	1.13
12	298	66.7	1.11	66.9	1.11
13	299	72.6	1.12	71.9	1.11
14	298	78.2	1.12	78.5	1.12
15	299	83	1.11	83.5	1.11
16	298	88.2	1.1	87.9	1.1
17	299	92.4	1.09	92.1	1.08
18	298	95.7	1.06	95.5	1.06
19	299	97.5	1.03	97.7	1.03
20	298	100	1	100	1

<표 4-40>는 다인 가구에서 기본 모형에 비해 우수한 경기도 모형과 지역 파생 모형과의 비교를 한다. 경기도 모형이 지역 파생 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 약간 높으

나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 약간 낮은 것을 알 수 있었다. 정확도 측면에서는 지역 파생 모형이 약간 좋은 결과를 보였다. AUC값은 경기도 모형이 지역 파생 모형에 비해 우수한 결과를 보였다. 이와 같은 결과를 통해 전반적으로 경기도 모형이 지역 파생 모형에 비해 우수하다는 것을 알 수 있다.

<표 4-40> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(경기도 다인가구)

구분	경기도 군집	
	다인 가구	다인 가구
	경기도 모형	지역 파생 모형
TN	11515	13041
FP	2203	677
FN	3981	4997
TP	1682	666
정확도	0.680924617	0.707239049
민감도	0.297015716	0.117605509
특이도	0.839408077	0.950648783
정밀도	0.432947233	0.495904691
AUC	0.6503	0.6476

다음은 다인 가구에서 경기도 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 경기도 모형의 값은 8.9, 1.78이며 지역 파생 모형의 경우 8.95, 1.79로 지역 파생 모형이 약 0.05% 배 정도 좋은 결과를 보여주었다. 하지만 두 번째 그룹과 세 번째 그룹 모두 경기도 모형이 더 좋은 반응을 보여주고 있기 때문에 전반적으로 보았을 때 경기도 모형이 지역 파생 모형에 비해 우수하다는 것을 알 수 있다.

<표 4-41> 경기도 모형과 지역파생 모형 Lift Chart값 비교 (다인가구)

구분		경기도 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	970	8.9	1.78	8.95	1.79
2	969	16.4	1.64	16.2	1.62
3	969	23.1	1.54	22.6	1.51
4	969	29.6	1.48	28.9	1.45
5	969	36.4	1.46	35.3	1.41
6	969	42.3	1.41	41.9	1.4
7	969	48.1	1.37	48.2	1.38
8	969	54.3	1.36	54.1	1.35
9	969	59.6	1.33	59.5	1.32
10	969	64.9	1.3	65.1	1.3
11	969	70.1	1.27	70.2	1.28
12	969	75	1.25	74.9	1.25
13	969	79.5	1.22	79.4	1.22
14	969	83.8	1.2	83.7	1.2
15	969	87.9	1.17	87.4	1.17
16	969	91.5	1.14	91.4	1.14
17	969	94.8	1.12	94.6	1.11
18	969	97.1	1.08	97.5	1.08
19	969	99	1.04	99	1.04
20	969	100	1	100	1

<표 4-42>는 전라북도 데이터 기반으로 1인 가구에서 기본 모형과 지역 파생 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 지역 파생 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에

서는 약간 낮으나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 약간 높은 것을 알 수 있었다. 정확도 측면에서는 기본 모형이 지역 파생 모형에 비해 더 좋은 결과를 보여주었다. AUC값은 지역 파생 모형이 기본 모형에 비해 더 좋기 때문에 분류행렬과 AUC 기준으로 지역 파생 모형이 기본 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-42> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(전북 1인가구)

구분	전라북도 군집	
	1인 가구	1인 가구
	기본 모형	지역 파생 모형
TN	213	151
FP	776	838
FN	114	73
TP	918	959
정확도	0.559623949	0.549233053
민감도	0.889534884	0.929263566
특이도	0.21536906	0.152679474
정밀도	0.541912633	0.533667223
AUC	0.5872	0.5898

다음은 1인 가구에서 기본 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 기본 모형과 지역 파생 모형이 비슷한 결과를 보여주었다. 반면에 두 번째 그룹의 경우 기본 모형이 13.2, 1.32 지역 파생 모형이 14, 1.4으로 기본 모형이 지역 파생 모형에 비해서 약 0.8%배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같은 결과를 통해 지역 파생 모형이 기본 모형의 성능이 우수하다는 것을 알 수 있다.

<표 4-43> 기본 모형과 지역파생 모형 Lift Chart값 비교 (전북 1인가구)

구분		기본 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	102	7.75	1.55	7.75	1.55
2	101	13.2	1.32	14	1.4
3	101	19.1	1.27	18.2	1.21
4	101	24.4	1.22	23.5	1.18
5	101	29.5	1.18	27.4	1.1
6	101	32.8	1.09	33.3	1.11
7	101	38.7	1.1	39.2	1.12
8	101	44.8	1.12	45.9	1.15
9	101	51.2	1.14	49.9	1.11
10	101	55.1	1.1	54.3	1.09
11	101	58.6	1.07	58.5	1.06
12	101	62.9	1.05	63.5	1.06
13	101	67.7	1.04	68.8	1.06
14	101	73.2	1.05	75.7	1.08
15	101	79.9	1.07	81.3	1.08
16	101	85.9	1.07	85.8	1.07
17	101	90.2	1.06	89.9	1.06
18	101	94.2	1.05	93.8	1.04
19	101	97	1.02	97.3	1.02
20	101	100	1	100	1

<표 4-44>는 다인 가구에서 기본 모형과 지역 파생 모형과의 비교를 한다. 지역 파생 모형이 기본 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 약간 높으나, 반대로 비대상자를 맞추는

비율은 기본 모형에 비해 약간 낮은 것을 알 수 있었다. 정확도 측면에서는 기본 모형이 약간 좋은 결과를 보였다. AUC값은 기본 모형이 지역 파생 모형에 비해 우수한 결과를 보였다. 이와 같은 결과를 통해 전반적으로 기본 모형이 지역 파생 모형에 비해 우수하다는 것을 알 수 있다.

<표 4-44> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(전북 다인가구)

구분	전라북도 군집	
	다인 가구	다인 가구
	기본 모형	지역 파생 모형
TN	3278	2631
FP	211	858
FN	1456	1026
TP	234	664
정확도	0.67812319	0.636223209
민감도	0.138461538	0.392899408
특이도	0.939524219	0.754084265
정밀도	0.525842697	0.436268068
AUC	0.6423	0.6368

다음은 다인 가구에서 기본 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 기본 모형의 값은 9.35, 1.87이며 지역 파생 모형의 경우 8.76, 1.75로 지역 파생 모형이 약 0.59% 배 정도 좋은 결과를 보여주었다. 전반적으로 보았을 때 기본 모형이 지역 파생 모형에 비해 우수하다는 것을 알 수 있다.

<표 4-45> 기본 모형과 지역파생 모형 Lift Chart값 비교 (전북 다인가구)

구분		기본 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	259	9.35	1.87	8.76	1.75
2	259	15.9	1.59	15.6	1.56
3	259	22	1.46	21.2	1.42
4	259	29.3	1.47	28.4	1.42
5	259	34.1	1.37	33.6	1.34
6	259	41.2	1.37	40.1	1.34
7	259	46.5	1.33	45.8	1.31
8	259	52	1.3	51.7	1.29
9	259	57.6	1.28	57.5	1.28
10	259	63.4	1.27	63.4	1.27
11	259	69.1	1.26	67.8	1.23
12	259	73.7	1.23	73.1	1.22
13	259	78.3	1.2	78	1.2
14	259	82.5	1.18	82.3	1.18
15	259	86.6	1.15	86.8	1.16
16	259	90	1.12	90.4	1.13
17	259	93.5	1.1	93.9	1.1
18	259	96.7	1.07	96.9	1.08
19	259	98.8	1.04	98.5	1.04
20	258	100	1	100	1

<표 4-46>는 제주특별자치도 데이터 기반으로 1인 가구에서 기본 모형과 지역 파생 모형의 분류행렬과 AUC 값을 보여주고 있다. 기본 모형이 지역 파생 모형에 비해 대상자를 대상으로 잘 맞추는 민감도 측면에서는 약간 높으나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 약간 높은 것을 알 수 있었다. 정확도 측면에서는 지역 파생 모형이 기본 모형에 비해 더 좋은 결과를 보여주었다. AUC

값은 기본 모형이 지역 파생 모형에 비해 더 좋기 때문에 분류행렬과 AUC 기준으로 기본 모형이 지역 파생 모형에 비해 더 좋은 것을 알 수 있다.

<표 4-46> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(제주 1인가구)

구분	제주특별자치도 군집	
	1인 가구	1인 가구
	기본 모형	지역 파생 모형
TN	81	96
FP	217	202
FN	31	42
TP	184	173
정확도	0.516569201	0.524366472
민감도	0.855813953	0.804651163
특이도	0.271812081	0.322147651
정밀도	0.458852868	0.461333333
AUC	0.604	0.573

다음은 1인 가구에서 기본 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 기본 모형이 6.05, 1.21 지역 파생 모형이 4.19, 0.837로 기본 모형이 지역 파생 모형에 비해서 약 1.86%배 정도 더욱 더 반응도가 좋은 것을 알 수 있다. 이와 같은 결과를 통해 기본 모형이 지역 파생 모형에 비해 성능이 우수 하다는 것을 알 수 있다.

<표 4-47> 기본 모형과 지역파생 모형 Lift Chart값 비교 (제주 1인가구)

구분		기본 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	26	6.05	1.21	4.19	0.837
2	26	10.7	1.07	8.84	0.884
3	25	17.2	1.15	13.5	0.899
4	26	23.3	1.16	20.5	1.02
5	26	27.9	1.12	28.4	1.13
6	25	34	1.13	33.5	1.12
7	26	41.4	1.18	40.9	1.17
8	26	50.2	1.26	44.7	1.12
9	25	52.1	1.16	48.8	1.09
10	26	57.2	1.14	54.4	1.09
11	26	61.9	1.12	57.7	1.05
12	25	65.6	1.09	61.9	1.03
13	26	72.1	1.11	70.7	1.09
14	26	80.5	1.15	76.3	1.09
15	25	83.7	1.12	82.8	1.1
16	26	87.9	1.1	87.4	1.09
17	26	92.1	1.08	92.6	1.09
18	25	95.8	1.06	95.3	1.06
19	26	97.2	1.02	97.7	1.03
20	25	100	1	100	1

<표 4-48>는 제주특별자치도 데이터 군집에서 다인 가구에서 기본 모형과 지역 파생 모형과의 비교를 한다. 기본 모형이 지역 파생 모형에 비해 대상자를 대상자로 잘 맞추는 민감도 측면에서는 약간 높

으나, 반대로 비대상자를 맞추는 비율은 지역 파생 모형에 비해 약간 낮은 것을 알 수 있었다. 정확도 측면에서는 지역 파생 모형이 약간 좋은 결과를 보였다. AUC값은 지역 파생 모형이 지역 파행 모형에 비해 우수한 결과를 보였다. 이와 같은 결과를 통해 지역 파생 모형이 약간은 기본 모형에 비해 우수한 것을 알 수 있다.

<표 4-48> 기본 모형과 지역파생 모형 분류행렬 및 AUC값 비교(제주 다인가구)

구분	제주특별자치도 군집	
	다인 가구	다인 가구
	기본 모형	지역 파생 모형
TN	1214	1238
FP	67	43
FN	386	407
TP	60	39
정확도	0.737695426	0.739432542
민감도	0.134529148	0.087443946
특이도	0.947697112	0.966432475
정밀도	0.472440945	0.475609756
AUC	0.6088	0.6162

다음은 다인 가구에서 기본 모형과 지역 파생 모형의 Lift Chart 값을 보여주고 있다. 누적 반응률과 누적 Lift 값을 기준으로 첫 번째 그룹의 경우 기본 모형의 값은 9.87, 1.97이며 지역 파생 모형의 경우 9.42, 1.88로 기본 모형이 약 0.4% 배 정도 좋은 결과를 보여주었다. 하지만 5번째 그룹이 후 지역 파생 모형이 더욱더 좋은 반응을

보여주었다. 이것은 대상자가 많이 뽑을 경우에 지역 파생 모형이 더욱 유리하다는 것을 의미한다.

<표 4-49> 기본 모형과 지역파생 모형 Lift Chart값 비교 (제주 다인가구)

구분		기본 모형		지역 파생 모형	
Rank Group	Group Size	누적반응률	누적 Lift	누적반응률	누적 Lift
1	87	9.87	1.97	9.42	1.88
2	86	17	1.7	15.2	1.52
3	87	22.9	1.52	21.7	1.45
4	86	29.1	1.46	28	1.4
5	86	33.6	1.35	34.3	1.37
6	87	38.6	1.29	39.9	1.33
7	86	43.7	1.25	45.1	1.29
8	86	49.3	1.23	49.6	1.24
9	87	56.5	1.26	56.1	1.25
10	86	61.7	1.23	63	1.26
11	86	66.8	1.21	70	1.27
12	87	73.5	1.23	74.2	1.24
13	86	76.7	1.18	78.3	1.2
14	86	80.5	1.15	81.6	1.17
15	87	83.9	1.12	85.2	1.14
16	86	87.2	1.09	89	1.11
17	86	90.4	1.06	91	1.07
18	87	93.7	1.04	94.8	1.05
19	86	98.4	1.04	98.4	1.04
20	86	100	1	100	1

결론

제1절 결론 및 제언

제5장 결론

제1절 결론 및 제언

본 연구에서는 지역별 예측모델의 효용성을 살펴보기 위해 국내외 사례 선행 연구 조사부터, 데이터 탐색과 예측모델 개발 및 평가까지 수행하였다.

국내외 사례 연구 파트에서 국내 사례의 경우 복지사각지대에 대한 개념이 과거에는 단순히 저소득층을 기준으로 복지사각지대 대상자에 초점이 맞춰져 복지사각지대 발굴 수행을 하는 사례에 대해 검토를 하였다. 국외에는 non-take-up, blind spot 개념에서 보건복지부에서 최근에는 주소득원, 우울, 정신 질환 등으로 인해 복지 서비스를 받지 못하는 대상자까지 복지사각지대에 대한 개념을 확대하였다. 보건복지부에서는 복지사각지대 대상자를 경제적 어려움이 처한 사람 외에도 빈곤을 넘어 고립, 관계단절, 정신적 인지적 문제가 있는 경우까지 복지사각지대 대상자를 확대하고 있는 상황이다. 국외 사례의 경우 복지사각지대에 대한 영향이 미치는 요인을 오스트리아에서는 고용 상태, 가구, 교육을, 독일에서는 한부모 가정과 연령으로 보았다. 그리고 스위스에서는 도시의 규모, 프랑스에서는 연령으로 보았다. 지역정보를 활용한 머신러닝 연구에서는 지역정보와 머신러닝 기법을 활용한 연구 중 교통사고를 예측하기 위한 모형과 범죄발생을 예측하기 위한 연구 그리고 심정지 발생 지역에 대한 연구를 살펴보았다.

머신러닝 알고리즘에 대한 부분에서는 본 연구에서 머신러닝 기법으로 활용하는 XGBoost에 대한 세부적인 설명과 모형 평가를 위한 분류행렬, ROC커브, AUC, 리프트 도표에 대한 설명에 대하여 살펴 보았다.

3장 데이터 탐색에서는 복지사각지대 19년 1차와 2차 데이터를 기반으로 전체 대상자의 속성분석, 단독변수와 가구 수에 따른 속성분석, 지역별 대상자 속성분석까지 진행하였다. 데이터 탐색 결과 전체 가구에서 경기도 지역이 대상자가 제일 많았고, 연령으로는 노년층이 제일 많은 복지사각지대 대상자 분포를 보였다. 그리고 피부양장기요양여부 변수가 복지사각지대 대상자에 많은 영향을 미쳤으며, 1개 변수를 가진 대상자 그리고 총 추출된 대상자 대비 조치완료가 가장 높은 변수는 자살예방관리대상자여부 변수였다. 지역별로 분석한 결과 연령으로 보면 노년층이 서울, 경기, 전북에서 제일 높은 비율을 보였고, 제주도의 경우 중장년층이 제일 높은 비율을 보였다. 변수 개수의 경우 모든 지역에서 1개 변수가 가장 높은 비율을 보였다. 연계 변수로 보면 월세금액기준이하여부가 서울, 경기, 제주에서 가장 높은것으로 나타났으며, 전북에서는 피부양장기요양여부 변수가 복지사각지대 대상자를 뽑는데 가장 큰 역할을 한 것으로 보였다. 이와 같은 결과 국외 사례와 같이 주거와 연령이 복지사각지대 대상자에 영향을 많이 미친 것으로 나타났다.

예측 모형 구현 및 분석에서는 기본모형, 지역별 모형 그리고 파생 변수를 투입했을 때의 효과를 살펴보았다. 서울특별시 모형 기준으로 1인 가구의 경우 기본 모형이 더욱 더 좋으며, 다인 가구의 경우 서울특별시 모형이 더 좋았다. 경기도 모형도 마찬가지로 1인 가구의 경우 기본 모형이 더욱 좋았으며, 다인 가구는 경기도 모형이 더 우수한 결과를 보여주었다. 전라북도와 제주특별자치도의 경우 1인 가

구 다인 가구 모두 기본 모형이 더 우수한 결과를 보여주었다.

파생 변수를 투입 하였을 때 성능 비교를 한 결과는 다음과 같이 볼 수 있다. 서울특별시 데이터 기반으로 1인 가구에서는 파생 변수를 투입한 모형이 기본 모형에 비해 우수하였으며, 다인 가구에서는 서울특별시 모형이 파생변수 투입 모형에 비해 우수하였다. 경기도 데이터에서도 마찬가지로 1인 가구에서는 기본 모형에 비해 지역 파생 모형이 더욱 더 우수한 결과를 보여주었다. 다인 가구에서도 경기도 모형이 지역 파생 모형에 비해 우수한 결과를 보여주었다. 전라북도 데이터에서는 1인 가구에서 지역 파생 모형이 기본 모형에 비해 우수하였지만 다인 가구에서는 기본 모형이 지역 파생 모형에 비해 우수하다는 것을 보여 주었다. 제주특별자치도 데이터에서는 1인 가구와 다인 가구 모두 기본 모형이 더욱 더 우수하다는 것을 보여주었다. 결론적으로 본 연구에서 설정하였던 가설인 예측 모형을 활용하였을 때 지역별 예측 모형이 효과가 있는지 그렇지 않은지에 대한 결과는 서울, 경기 지역같이 인구수가 많은 지역에서는 효과가 보인다고 결론을 낼 수 있었다.

이와 같은 다양한 분석을 기반으로 본 연구에서는 다음과 같은 정책적 제언을 하고자 한다.

첫째, 사회보장정보원의 복지사각지대발굴시스템은 약 3년간 안정적으로 운영해온 시스템으로 많은 훈련데이터가 누적되었다. 이에 머신러닝에 대한 효과를 발휘하기에는 충분한 데이터로 활용될 수 있다. 하지만 현재 정확도에서 더욱 더 정교한 정확도를 가지기 위해서는 데이터 입수기관과 수집되는 데이터에 대한 정합성, 수집되는 데이터가 분석대상자로 구성되는 세부적인 프로세스, 분석대상자에서 예

측모형의 입력 데이터로 전환하기 위한 전처리 과정 및 최종 발굴 대상자가 되기까지에 존재하는 각 필터링 등을 재정비 하여 더욱 더 정확도를 재고해야 될 것이다. 또한 훈련데이터 정비 수행시 시계열 구성을 한다면 더욱 더 정교하고 정확한 예측 모델을 만들 수 있을 것이다. 하지만 시계열 대상자 데이터 구성을 위해서는 모든 연계변수 입수 주기를 1달 이내로 줄이고 상시 데이터 수집과 적제가 이뤄져야 되는 부분이 존재할 것이다.

둘째, 서울특별시와 경기도 같이 타 지역에 비해 2배 이상 인구가 많고 수급자 수가 많은 지역에서는 지역 군집을 수행하면 복지사각지대 적중률 향상에 도움이 될 수 있다. 하지만 전라북도 및 제주특별자치도와 같이 인구와 수급자가 상대적으로 적은 지역에서는 지역 군집의 효과가 떨어진다. 이와 같은 결과를 통해서 인구 수에 따라 지역을 분리하고 분리된 지역이 특정 인구 수 이상이 될 경우, 지역 모형을 도입하고, 그렇지 않을 경우 지역 파생 변수 또는 지역 융합 데이터 기반의 발굴 등으로 복지사각지대 발굴시스템 운영을 수행한다면 적중률 향상에 도움일 될 것이다. 본 연구에서 지역 파생변수는 시군구 데이터 기반으로 수행하였지만 만약 지리 정보 시스템 (GIS) 과 같은 마이크로 지역 데이터를 활용하여 공간자료 (Spatial Data) 와 벡터자료 (Vector Data)를 추가적으로 활용한다면 지역별 모형의 특색을 더 살릴 수 있을 것이다. 예를 들어, 복지 대상자가 거주하는 지역에서 가장 가까운 복지 서비스 건물과 건물 수, 그리고 주변에 상권 등을 별도 변수로 활용하는 것이다.

셋째, 본 연구에서는 데이터 분석을 통하여 시도, 연령, 연계변수, 변수갯수, 조치결과에 대해서 세부적으로 살펴보았다. 연령 기준으로

노년층이 다른 연령에 비해 많은 것을 보여주었으며, 연계변수의 경우도 연령과 같은 맥락으로 피부양장기요양여부 변수가 변수 개수의 경우 1개가 제일 많은 대상자 분포인 것을 보여주었다. 지역 단위로 쪼개서 연계변수를 보았을 때에는 월세기준이하가구여부가 제일 높은 연계변수를 보여주었다. 이런 결과를 통해 연령이 높은 노년기 층에 속하는 대상자의 경우 별도의 모형 또는 별도의 복지서비스 전달을 하기 위한 방안을 마련한다면 더욱 더 노년기층에 속하는 대상자에게 복지 서비스를 많이 주면서 복지사각지대 관리시스템에서는 타 연령 층에 많은 대상자에게 복지 서비스 기회를 줄 수 있을 것이다. 또한 주거유형을 전세, 월세로만 분리 하는 것을 넘어서 자가, 보증부월세, 영구임대주택, 국민임대, 장기전세임대, 무료임대, 가정위탁, 보장기관 거주자, 그룹홈 거주자 등 주거 유형을 더욱 더 세밀하게 분리하여 복지사각지대 대상자를 판별한다면 더욱 더 주거 취약 계층을 발굴하는데 도움이 될 것으로 판단된다.

넷째, 일반적으로 예측 모형을 평가하는 방법으로 분류행렬, ROC 커브, AUC, Lift 도표 등을 활용하여 예측 모형끼리 비교를 하여 제일 성능 좋은 예측 모형을 도출하게 된다. 물론 특정 데이터를 기반으로 예측 모형을 평가하는 데 활용될 수 있고, 고전적으로 많이 활용되는 예측 모형평가 방법이다. 하지만 사회보장정보원의 복지사각지대발굴시스템의 경우 예측 모형으로부터 선정된 대상자가 최종적으로 지자체에 분배되고 지자체 공무원이 복지 서비스를 연결 시켜주는 구조이기 때문에 모형끼리 비교 결과가 정확하게 예측 모형의 정확도를 설명하기는 어렵다. 이에 따라 지자체 공무원이 처리되는 실적까지 고려된 새로운 모형 평가 방법이 추가적으로 고려되어 가상 환경에서 실험한 결과가 지자체 공무원이 처리된 실적까지 예측이 될 수 있는

지표가 만들어져야 정확한 실험이 가능할 것으로 보인다. 예를 들어, 분류 행렬의 민감도 + 정확도 + 지자체 공무원 수 + 처리량 대비 정확도 등을 모형 평가에 넣는 것이다.


다섯째, 복지사각지대 발굴시스템의 정확도 제고 및 서비스 범위를 넓히기 위해서는 비정형 데이터의 확대를 수행해야 될 것이다. 특히 비정형 데이터 중 복지와 보건 사례관리용 비정형 데이터는 유용하게 활용될 것으로 판단된다. 2010년부터 공공영역에서 사례관리가 본격적으로 수행하였으며, 2012년에는 전국 시군구에 희망복지지원단을 통해 통합사례관리가 수행되어 2016년부터 읍면동까지 통합사례관리 사업이 확대되어 2017년에는 보건서비스까지 확대되어 추진되고 있다. 사례관리라고 하는 것은 복합적이면서 다양한 욕구가 있는 클라이언트와 그 가족을 대상으로 사회적 기능 회복을 돕는 통합적 실천 방법으로써 사례관리에서 수행되는 데이터는 다양한 욕구를 가진 클라이언트의 특성을 반영할 수 있다. 이와 같은 데이터를 복지사각지대 발굴시스템에서 활용한다면 더욱 더 정확하게 대상자의 욕구를 파악하여 서비스 연결을 할 수 있는 시스템으로 발전될 수 있을 것으로 판단된다.

여섯째, 머신러닝 알고리즘의 지도 학습을 보완하기 위한 Weakly-supervised learning 기법 활용은 인공지능을 더욱 더 강화하는데 도움을 줄 것이다. 현재 복지사각지대 관리시스템은 2달에 1번씩 지자체에 최종 대상자를 내려 보낸 후에 피드백 데이터가 들어오면 훈련하는 방식으로 진행된다. 하지만 훈련데이터 생성되는 기간이 짧게는 2달 길게는 6달 넘게 걸리는 점을 감안한다면 지도 학습을 통해 인공지능의 추론 능력을 향상시키는데 한계가 존재한다. 만약

연계 데이터가 실시간으로 들어오고, 지자체에서 실적이 실시간으로 들어오는 상황을 가정하면 실시간 연계 데이터가 들어오는 데이터를 가지고 인공지능은 Labeling을 하기 위한 별도 인공지능을 통해 훈련 데이터를 만들 수 있다. 또한 지자체 자체실적으로 들어오는 데이터를 계속해서 추적하여 어떤 연계데이터가 지자체 처리된 결과와 매칭되는지 파악하여 역으로 연계 데이터 또한 추론할 수 있을 것이다. 이런 구조가 된다면 2달에 한번씩 데이터가 쌓이는 구조가 아닌 강화 학습처럼 계속해서 강화되는 인공지능으로 거듭날 수 있을 것이다.

마지막, 복지사각지대 발굴관리시스템은 행정데이터와 최신 머신러닝 기법을 활용한 대표적인 시스템으로써 사회보장정보원의 다양한 시스템과의 접목이 가능하다. 사회보장정보원의 7대 대표 시스템인 사회보장정보시스템 (행복e음, 범정부), 지역보건의료정보시스템, 사회서비스전자바우처시스템, 보육통합정보시스템, 사회복지시설정보시스템, 취약노인지원시스템과 접목을 통해 다양한 응용 어플리케이션 개발이 가능할 것이다. 또한 기존 복지사각지대 발굴시스템의 벤치마킹을 수행한 e아동행복지원시스템과 같이 그 밖의 다양한 분야에서도 벤치마킹을 통해 업무 효율화 및 복지서비스의 확대를 기대할 수 있을 것으로 판단된다.

이상 7가지 정책적인 제언은 사회보장정보원의 복지사각지대 시스템의 고도화 및 차세대 사회보장정보시스템으로 발전하는데 큰 도움이 될 것으로 판단된다. 특히 본 연구는 차세대 사회보장정보시스템에서 예측모형의 다각화를 위한 기초 자료로 활용될 수 있으며, 연령별 모형, 가구유형 모형 등을 만들 때 참고자료로 활용될 수 있을 것으로 판단된다.



참고문헌

- 구인회·백학영(2008). “사회보장의 사각지대: 실태와 영향요인”. 사회보장연구, 24(1): 175-204.
- 김은하 외(2015). 사회보장정보시스템을 활용한 복지 사각지대 발굴방안 연구. 보건복지부·사회보장정보원.
- 김은하 외(2016). 사회보장정보시스템을 활용한 복지 사각지대 발굴 방안 후속연구. 보건복지부·사회보장정보원·한국보건사회연구원·서울대학교산학협력단.
- 보건복지부 (2017). 지난 겨울, 42만 3000여 명에게 온기를 나누다
- 보건복지부 (2019). 포용적 사회보장 정보 기반 마련을 위해 사회보장정보시스템 새로 구축한다!
- 최균 외(2012). 복지사각지대 발굴 체계화 및 복지정보 접근성 강화방안. 보건복지부·한국사회복지협의회.
- 사회보장위원회(2014). 복지사각지대 발굴 및 지원 종합대책(안). 보건복지부.
- 김은하·추병주(2014). 복지 사각지대 발굴 및 지원체계를 위한 DB 설계. 사회보장정보원.
- 최현수 외(2016). 1인 취약가구 위험분석 및 맞춤형 정책지원 방안 연구. 보건복지부·한국보건사회연구원.
- 최현수 외(2018). 사회보장정보시스템을 활용한 복지사각지대 발굴방안 확대 연구

- 박병호 (1995). 지역별 교통사고 예측모형에 관한 연구
- 허선영 외(2018). 머신러닝기반 범죄발생 위험지역 예측
- 박일수 외(2012). 지리적 가중회귀모형을 이용한 지역별 심정지 발생요인에 관한 연구
- 허진경 (2018). 처음 시작하는 빅데이터 분석 및 활용 : R 프로그래밍, 데이터 분석, 시각화, 머신러닝
- 허선영 외(2018). 머신러닝기반 범죄발생 위험지역 예측. 한국지리정보학회지
- 채경재 외(2018). 머신러닝과 샘플링을 이용한 강원도 지역 산불발생예측모형 개발. 한국빅데이터학회지
- 최창현 외(2018). 머신러닝 기법을 이용한 수도권 지역의 호우피해 예측함수 개발. 한국방재학회
- 강경희 외(2018). Support Vector Machine을 이용한 머신러닝 기반의 상주 지역 산사태 취약성 분석. 대한지질공학회
- 박호정 외(2018). 광역시 지역주민의 원전계속운전 수용성 결정요인 분석-머신러닝과 CVM 기법의 적용. 한국지방행정학보
- 신정훈 외(2018). 딥러닝 알고리즘을 활용한 가뭄 취약 지역 분석. 대한토목학회
- Neal JEan. et. al.(2016). Combining stellite imagery and machine learning to predict poverty. Science
- Linden McBride and Austin Nichols (2015). Improved poverty targeting through machine learning: USAID Poverty Assessment Tools.econthatmatters.com/wp-content/uploads/2015/01/improvedtargeting_21jan2015.pdf
- Anthony Perez. et. al.(2019). Semi-Supervised Multitask Learning on Multispectral Satellite Images Using Wasserstein Generative Adversarial Networks (GANs) for Predicting Poverty. Cornell University

- Harnisch, M.(2019). Non-Take-Up of Means-Tested Social Benefits in Germany. DIW: Berlin.
- Fuchs M. et al.(2019). Falling Through the Social safety Net? Analysing Non-take-up of Minimum Income Benefit and Monetary Social Assistance in Austria. EUROMOD Working Paper Series.
- Fuchs M. et al.(2019). Falling Through the Social Safety Net? The Case of Non-Take-Up in Austria. European Centre for Social Welfare Policy and Research.
- Heckman, J.(1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement*, 5(4): 475-492.
- Hümbelin O.(2019). Non-Take-Up of Social Assistance: Regional Difference and the Role of Social Norms. *Swiss Journal of Sociology*, 45(1): 7-33.
- Crettaz, E. et al.(2009). Sozialhilfe und Armutsstatistik im Vergleich. Neuchâtel: Bundesamt für Statistik.
- Arrighi Y. et al.(2015). The Non-Take-Up of Long-Term Care Benefit in France: A Pecuniary Motive? *Health Policy*, 119: 1338-1348.
- Tianqi Chen and Carlos Guestrin (2016). XGBoost: A Scalable Tree Boosting System.
- Fawcett, Tom (2006). An Introduction to ROC Analysis
- Stehman, Stephen V. (1997). Selecting and interpreting measures of thematic classification accuracy
- Tufféry, Stéphane (2011). Data Mining and Statistics for Decision Making

지역별 예측모델을 통한 사각지대 발굴 연구

발 행 일 : 2019. 12

저 자 : 이우식, 윤기찬, 박규범, 이인수

발 행 인 : 임 희 택

발 행 처 : 사회보장정보원

주 소 : 서울 중구 퇴계로 173(충무로3가) 남산스퀘어빌딩 15층(우 04554)

전 화 : 02) 6360-6114

홈페이지 : <http://www.ssis.or.kr>

인 쇄 : (주)필코in

서울 중구 퇴계로 187 (02)2269-2466

정 가 : 비매품

ISBN 979-11-86408-45-2

* 잘못된 책은 바꿔 드립니다.