

# 신용정보 및 공공 빅데이터 분석사례: 채납자 회수 가능성 예측모형 개발 및 채납정보 시스템 개선

김은찬 연구원(KCB 연구소), 유병준 교수(서울대학교)

김은찬 연구원

KCB 연구소

Tel: 02-708-6370, E-mail: eunchan.kim@koreacb.com

유병준 교수

서울대학교 경영대학

Tel: 02-880-2550, E-mail: byoo@snu.ac.kr

## 요약

KCB는 2017년 공공 빅데이터 표준분석모델 수행기관으로 선정되어, 표준 채납자 예측 모형을 개발하고 기존 채납 정보 시스템을 개선하였다.

해당 과제는 국가행정의 기반이 되는 세금의 비효율적인 관리 및 징수 프로세스와 시스템을 개선하고, 데이터 기반의 과학적 국정운영 및 의사결정을 도모하였다는 평가를 받았다.

본 고에서는 채납자 회수 가능성 예측 모형의 주요 내용과 시뮬레이션 분석을 통한 기대효과를 살펴보고자 한다. 첫째, 채납자의 신용정보와 채납 정보의 데이터 융합 및 탐색적 자료 분석(EDA). 둘째, 로지스틱 회귀분석과 세 가지 머신 러닝 방법론(Recursive Partitioning, Neural Network, Random Forest)의 예측 성능 비교를 통한 모형개발과정. 셋째, 예측 결과의 시스템 적용 및 시사점. 끝으로, 예측 시뮬레이션 분석을 통한 모형 적용 시 세금 조기 환수 및 비용절감 기대효과.

## 키워드:

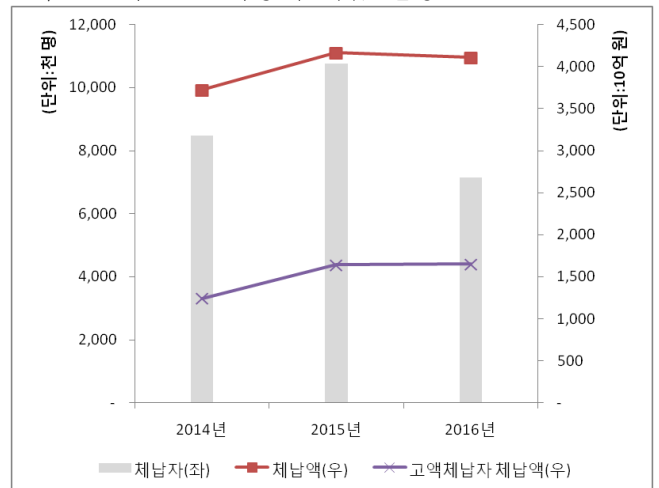
Machine Learning, Prediction Modeling, Neural Network, Random Forest, Credit Score System, Logistic Regression

## [사전자료] 채납자 및 채납액 현황

2015년 말, 지방세 채납자 수는 1천만 명, 총 채납 금액은 약 4조 1,654억 원을 기록하였다. 지방자치단체의 효율적 운영을 위해서는 재정운영의 기반이 되는 지방세를 효과적으로 회수할 수 있는 적절한 방안 수립이 매우 중요한 사안이다. 특히, 2015년까지 꾸준히 증가하던 채납자 수가 2016년에 급격히 감소한 반면 전체 채납 금액은 큰 변화가 없어, 채납자 1명당 채납 금액이 증가하고 있다. 실제, 고액 채납자 수와 고액 채납자의 채납액

비중은 매년 증가하고 있다. 2016년 기준 고액채납자 13,018명의 채납액은 1조 6,501억 원으로 전체 채납액의 40%를 차지한다. 채납 현황을 정확히 파악하고, 채납자의 특성을 적용한 차별화된 회수 조치 방안이 필요한 대목이라고 할 수 있다.

그림 1. 최근3년 지방세 채납 현황



행안부 지방세통계연감 자료[1]

## 서론

2016년, “정부 3.0”이라는 국정운영의 새로운 패러다임이 발표되면서, 데이터 기반의 과학적 행정이 강조되었다. 이후, 정부의 빅데이터 분석을 통한 현안 해결 수요가 증가하였으며, KCB는 지방자치단체(이하 지자체) 주요 과제 중 하나인 “수원시 지방세 채납자 분석 사업”을 수행하였다. 해당 과제를 통해 KCB는 최초로 채납자 대상 신용 데이터 융합을 시도하고, 회수 가능성 예측 모형을 개발하여 지자체의 회수 전략의 차별화를 도모하였다는 평가를 받았다.[2]

수원시 프로젝트에 대한 긍정적 평가로 KCB는

표준분석모델에서는 기존 수행 내역을 보다 고도화 하는 작업을 통해 체납 분석의 범위를 확장하였으며, 개별 지자체의 특성을 반영하면서도 타 지역으로 확산이 용이하도록 표준화 작업을 수행하였다. 기존 모형의 분석 범위는 지방세에 한정 되었으며 신용정보 만을 모형 개발에 활용하였으나, 표준 분석 모형에는 분석 범위에 세외수입 항목을 추가하였으며, 체납 정보도 모형 개발에 활용하였다. 본 고에서는 KCB가 수행한 ‘17년 표준분석모델의 주요 모형 개발내용과 시뮬레이션 분석을 통한 기대효과를 살펴보고자 한다.

그림 2와 같이 지방세는 부과 이후 기본적으로 최초 납입 기한이 존재하고, 해당 최초 납입 기한이 경과하면 시스템 상 체납자로 된다. 이후 지자체는 독촉 및 압류 등의 프로세스를 거쳐 체납 지방세를 회수한다. [3]

[illegible]

1 행정안전부에서 선정하는 빅데이터 우수분석 모델로서 전국 지방정부에 표준화된 시스템을 확산·보급하여 분석결과의 일관성을 유지하고 비용의 최소화를 목표로한

지방세 체납 분석의 주요한 목적은 보다 효율적인 지방세 징수활동을 지원하고 기존 징수 시스템을 개선하는 것이다. 체납자 별 회수가능등급을 토대로 체납 발생 후 일정기간 경과 전까지 납부 가능성이 높은 집단에는 부드러운 고지 및 독려 활동(SMS 및 이메일 안내 등) 을, 납부 가능성이 낮은 집단에는 강력한 독려 활동을 추진하도록 모형 및 시스템의 프로세스를 설계하였다.

지방세 및 세외수입의 데이터 추출기준은 과세년월 기준 포함 3년 (2014.07 - 2017.06)의 체납 건으로 정하여, 연도별 데이터 추이와 변동폭이 클 경우와 모형 개발 이후 데이터 상 통계·계량적 검증을 위한 작업을 대비하였다.[4]

지방세는 지방자치단체가 당해 년도의 재정수요를 충당하기 위하여 주민에게 부과 및 징수하는 조세로서 총 11개 세목<sup>2)</sup>으로 구성되어 있다. 본 분석에서는 현업 담당자 인터뷰를 통해 개인을 대상으로 부과하고 실제 체납이 발생 가능한 6개 세목(등록면허세, 자동차세, 주민세, 지방소득세, 취득세)으로 분석대상을 정의하였다.

세외수입은 국가·공공단체의 조세를 포함하는 공채 이외의 수입을 일컬으며, 수익자부담금, 재산수입, 사용료·임대료 수입, 수수료수입, 기부금·벌금·과태료 등으로 분류되고 있다. 하지만, 전국적으로 관리되고 있는 세목을 기준으로 살펴보면 약 2천여 개 이상의 세목이 존재하며, 각 지자체 별 특성에 따라 관리 대상 세목도 상이하다. 본 분석에서는 각 지자체에 공통적이고 체납 빈도 약 97% 이상을 차지하는 세목으로 총 7개 세목<sup>3)</sup>을 선정하였다.

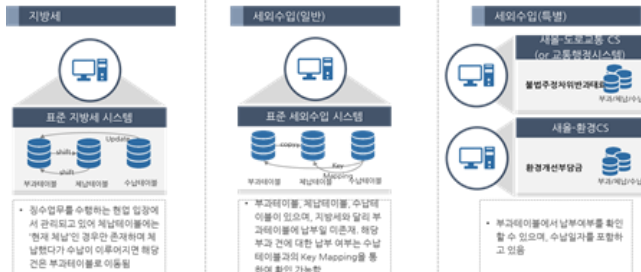
2 지방세 세목은 자동차세, 주민세, 취득세, 재산세, 지방소득세, 등록면허세, 담배소비세, 레저세, 지역자원시설세, 지방소비세, 지방교육세 11개로 분류할 수 있다

자동차검사지연과태료, 자동차손해배상보장법위반과태료  
(자동차의무보험미가입과태료), 장애인주차구역위반과태료,  
불법주정차위반과태료, 화물자동차운수사업법위반과징금,  
환경개선부담금, 그외수입이 해당된다.

1 행정안전부에서 선정하는 빅데이터 우수분석 모델로서 전국 지방정부에 표준화된 시스템을 확산·보급하여 분석결과의 일관성을 유지하고 비용의 최소화를 목표로한

지방세 연구원의 연구자료[5]에 따르면, 지방세와 세외수입 데이터는 각각의 독특한 관리 구조를 지니고 있는데 실제 시스템을 확인해 보면, 그림 3과 같이 공통적으로 부과·체납·수납 테이블로 구성되어 있고, 지방세와 세외수입은 각각의 다른 관리 시스템을 보유하고 있다.

그림 3. 지방세 및 세외수입 데이터 관리 현황

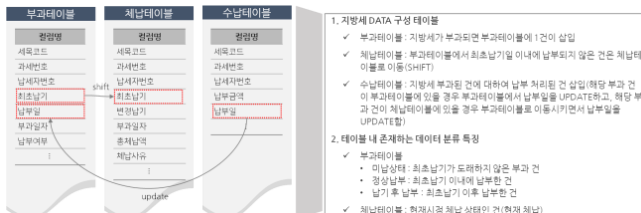


### ① 표준지방세시스템

표준지방세시스템은 전국 지자체의 공통적인 세금관련 시스템으로서, 크게 부과·체납·수납 테이블로 구성되어 있다. 시스템의 특징은 지역정보개발원에서 전국 시스템을 운영하나, 데이터의 원천이 되는 서버는 각 지자체에서 보유하고 있다는 점이다.

과제에서 필요한 지방세 체납 데이터의 추출을 위해서는 체납 테이블 이외에 부과테이블을 확인해야 한다. 체납테이블에서는 현재 시점에 관리되고 있는 체납 건만 존재하기 때문이다. 즉, 체납한 건을 추출 시 현재 체납 상태인 체납 건 이외에도 부과 테이블에서 최초 납기 후 납부한 건(최초납기일자 < 납부일)을 포함하여 확인해야 한다.

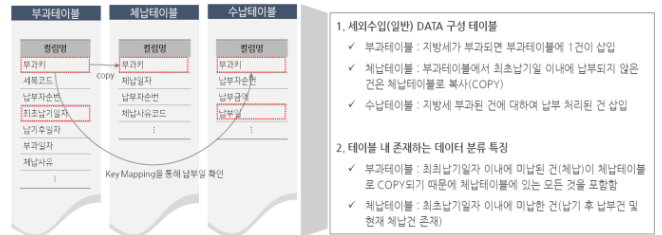
그림 4. 표준지방세시스템 데이터 구성(참고)



### ② 표준세외수입 시스템

표준세외수입시스템도 지방세와 동일하게 부과·체납·수납테이블로 구성되어 있는데, 표준지방세 시스템과의 차이점은 부과테이블에 납부 일자 항목이 존재하지 않아, 부과 건 별 납부 일자를 추출하기 위한 별도 작업을 수행해야 했다. 이외에도 표준세외수입시스템에서는 데이터 조회를 위한 SQL 쿼리를 사용할 수 없다는 차이가 있었다.

그림 5. 표준세외수입시스템 데이터 구성(참고)



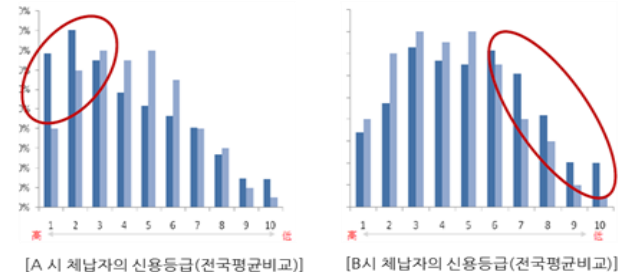
## 3) 탐색적 자료 분석(EDA) 결과

### ① 체납자 신용등급

지자체 별 체납자의 신용등급 분포를 확인하기 위한 분석을 통해 매우 흥미로운 결과를 확인하였는데, 각 지자체 별 체납자의 신용등급 분포가 매우 다른 형태를 보인다는 것이었다.

그림 6은 대표적인 지역 두 곳의 체납자 신용등급 분포도이다. B시의 체납자는 신용등급의 분포가 전국 신용등급 분포에 비해 하위등급(6 - 10등급)에 많이 몰려있음을 확인할 수 있는데, 예상대로 세금 체납자가 금융관련 연체도 많다고 확인해 볼 수 있다. 반면, A시는 일부 체납자들이 상위등급(1 - 3등급)에 다수 분포되어 금융 채무는 정상적으로 납부하나 세금을 체납하는 경우가 다수 있는 것을 확인하였고, 공적의무로서의 규칙을 지키지 않는 경우가 다수 있는 것을 확인할 수 있었다.

그림 6. 체납자에 대한 신용등급 현황



### ② 체납 건의 납부소요기간

회수 가능성 예측 모형 적용기간 설정 시 매우 중요한 요소로서, 납부소요기간을 확인한 결과 그림 7 및 8과 같이 체납 발생일로부터 30일 시점까지 급격히 증가하다가 완화된 패턴을 보였다. 이는 체납자들이 최초 납부기한(최초 30일)을 경과하여 독촉 기한(다음 30일) 경과 이전에 체납 세금을 납부하여 독촉 및 압류 등의 법적 절차를 회피하는 경향을 나타낸다.

그림 7. 체납 건의 납부소요기간 누적 현황

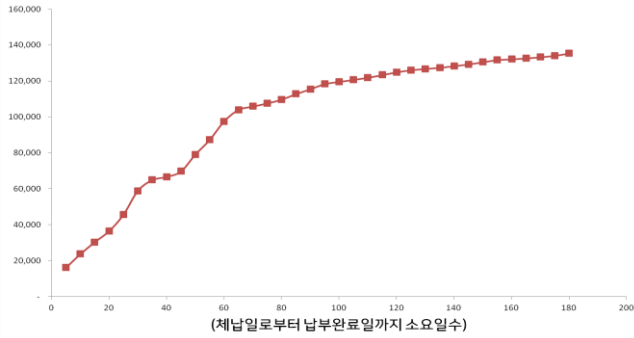


그림 8. 최초납기일로부터 납부 일까지 소요 일자



### ③ 지방세 및 세외수입 데이터 통합 분석

현재 전국의 각 지자체는 지방세와 세외수입을 별도의 시스템으로 운영·관리 하고 있기 때문에 징수 관련 업무 담당자도 별도로 구성되어있다. 체납자 개인의 입장에서는 납부해야 할 기관이 동일하지만 창구가 분리되어 있다. 금번 과제 진행 시, 이러한 문제점을 확인하고, 지방세와 세외수입 체납 데이터의 융합 분석을 하였다. 분석 결과, 표 1 과 같이 전체 체납자 중 약 17%의 체납자가 지방세와 세외수입을 동시에 체납한 것으로 나타났고, 이러한 중복 체납자의 체납 건수는 전체 체납 건수의 약 40%에 달하였다.

지방세 및 세외수입 둘 중 한 종류만 체납하고 있는 경우는 인당 체납 건수가 1.4건 이하인 반면, 이중 체납자의 인당 체납 건수는 4.6건으로 약 3배 이상 높음을 확인할 수 있었다. KCB는 이러한 분석내용을 토대로 현재의 지방세 및 세외수입이 별도의 관리 시스템으로 운영되는 것보다 통합 시스템으로 운영 시 징수 업무의 효율성이 증대될 것으로 기대하고 행정안전부와 지역정보개발원에 관련 의견을 개진한 상태이다.

표1. 지방세 VS. 세외수입 체납의 연관성 분석

### • 지방세 체납자 VS. 세외수입 체납자 Matrix

구분		지방세 체납자 여부	
		X	O
세외수입 체납자 여부	X		61%
	O	22%	17%

### • 지방세 체납 건 VS. 세외수입 체납 건 Matrix

구분		지방세 체납 여부	
		X	O
세외수입 체납 여부	X		45%
	O	16%	39%

※ 지방세와 세외수입 체납의 연관성 분석은 8시의 과세연월 기준 2016년(1년) 체납 데이터 활용

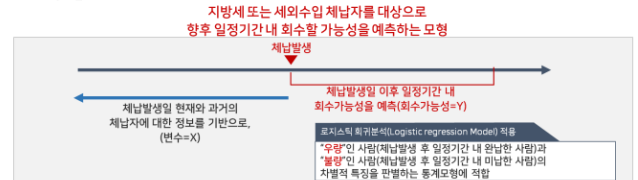
## 체납자 회수 가능성 예측모형 개발

회수 가능성 예측 모형은 체납 발생 시 일정기간 내 체납자가 자발적으로 납부할 가능성을 예측하는 모형[6] 으로서 체납 정보와 신용정보 데이터를 융합하여 모형을 개발하였다. 그림 9 와 같이 예측 모형은 우량과 불량 의 차별적 특징을 판별할 수 있도록 기본적인 신용평가 모형과 같이 설계 되었다. 기본적인 신용평가 모형은 1970년대 정립 된 것으로 [7], 부도·금융 연체 확률을 로지스틱 모형으로 추정하여 이를 점수화 및 등급화 하는 것을 의미한다. 모형의 예측 성능을 높이기 위한 학계의 다양한 노력 덕분에 최근에는 전통적인 통계적 기법을 벗어난, 머신러닝 방법론 및 앙상블 등 하이브리드 모형도 개발되었고 [8, 9, 10], 데이터 상의 다중공산성 및 내생성 문제를 극복하기 위한 방법론의 개선이 이루어지고 있다.

체납자 회수 가능성 예측모형의 방법론은 이러한 다양한 통계적 분석기법 중 이항 로지스틱 회귀분석<sup>4</sup>을 선택·적용하였는데, 이는 예측면에서 여타 방법론들이 조금 우세하나 각 변수의 설명력(explanability)를 고려하여 최종 선택 하였다. 모형 개발 이후에는 연체 확률을 점수화하고, 부여된 점수를 토대로 구간화(calibration) 하여 등급을 산출을 차후 진행하여 등급별 일정한 우량률과 불량률을 유지 할 수 있도록 개발 하였다.

그림 9. 모형 설계 내용

• 스코어 모형 개발 Frame



모형적용 대상 세목: 지방세 6개 세목, 세외수입 7개 세목

※ 개인 대상 부과 지방세는 11개 세목이 있으나, 현업담당자 인터뷰 결과 담배세 외 4개 세목은 체납이 발생하지 않는 세목이므로 분석에서 제외기로 결정하였으며, 모형적용 대상 세목 6개는 재산세, 자동차세, 지방소득세, 등록면허세, 주민세, 취득세에 해당함  
※ 세외수입은 체납 발생건수 기준으로 A시 전체 38개의 관리 세목 중 7개 세목이 체납건 전체의 96% 이상에 해당함

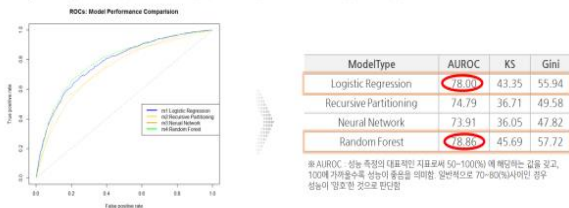
<sup>4</sup> 판별분석에 주로 사용되는 통계적 방법론으로 독립변수의 선형관계를 통해 종속변수를 설명할 수 있기에 높은 설명력을 가짐



추가적으로 세 가지 머신러닝 알고리즘<sup>5</sup>을 활용하여 모형 성능 비교를 진행하였다. 그림 10과 같이 통계적 방법론 별 모형의 성능으로는 AUROC<sup>6</sup> 기준 Random Forest 방법론이 가장 뛰어난 성능을 보였고 근소한 차이로 로지스틱 방법론이 뒤를 이었다. 모형의 예측 성능만 고려한다면 Random Forest 모형을 선택하여야 하나, 로지스틱 방법론은 각 변수의 영향도를 설명할 수 있을 뿐만 아니라 주어진 데이터에 매몰되는 과 적합 문제도 상대적으로 낮은 수준이기에, 도입 시 징수 업무 담당자의 입장에서 결과 해석과 활용이 용이한 로지스틱 회귀분석 모형을 최종 선정하였다.

그림 10. 방법론 별 모형 성능 비교(세외수입 A 모형 기준)

· 로지스틱 회귀(Logistic Regression) 방법론과 머신러닝 방법론을 사용한 체납자 회수 가능성 예측 성능 비교



- Recursive Partitioning(반복분할): 데이터 상에서 확인 가능한 분류 및 세분화 학습을 통해 향후 관측 값을 예측 및 분류하는 알고리즘
- Neural Network(신경망): 인체 신경망의 복잡계를 차용한 학습 알고리즘으로서 2개이상의 은닉층을 보유하면 Deep Learning (또는 Deep Neural-Net) 이라 칭함
- Random Forest(랜덤포레스트): 여러 개의 결정 트리들을 임의적으로 학습하는 다중 트리모형으로서 앙상블 알고리즘

### 참고) 예측 모형 별 수식 [8]

각 예측 모형 별 수식은 아래와 같다. 모형에 최종 선정된 변수는 신용정보 변수 7개와 체납정보 변수 3개이며, 각 모형 별 사용 변수는 상이하다. 해당 변수의 정보는 비공개로 한다.

#### 1) 로지스틱 회귀모형 수식

$$Y = \alpha + \beta X_1 + \beta X_2 + \dots + \beta X_n + \varepsilon$$

(Y 는 부도 · 연체 확률 이항수식 上 동일)

#### 2) 반복분할 트리모형 수식

$$Y = W_1 X_1 + W_2 X_2 + \dots + W_n X_n$$

(반복 트리모형으로,  $W \sim C(\pi_1, \pi_2)/2$  로 이분판별)

<sup>5</sup> Recursive Partitioning(반복분할), Neural Network(신경망), Random Forest(랜덤포레스트)

<sup>6</sup> 산출된 ROC의 면적으로서 판별된 적중확률과 오경보확률의 계산을 통해 값을 산출하며, 1에 가까울수록 우수한 예측력을 지닌 모형으로 판단

### 3) 신경망 모형 수식

$$Y = h\theta(X) = a_1(X_1 + X_2 + \dots + X_n)$$

( $h\theta(X)$  은 한 개의 hidden layer(은닉층) 와 20개의 node(은닉연결지점) 를 갖는 모형으로,  $\tanh(x) = 2\sigma(2x) - 1$  의 활성화함수를 따른다.)

### 4) 랜덤포레스트 모형 수식

$$Y = R_1 X_1 + R_2 X_2 + \dots + R_n X_n$$

(앙상블 트리모형으로,  $R \sim C(\pi_1, \pi_2)/2$  무작위판별임)

#### 1) 지방세 모형

지방세 체납자 회수 가능성 예측 모형은 체납발생일로부터 30일 이내에 자발적 납부 가능성을 예측한다. 모형에서 우량의 기준은 체납 발생 후 30일 이내에 완납한 경우로 정의하였다. 先 진행한 EDA 분석 결과를 반영하여 세목을 기준으로 구분된 2개의 모형을 개발하였다. 지방세 모형 A는 자동차세와 재산세 세목으로 구성하였고, 지방세 모형 B는 나머지 4개의 세목인 등록면허세, 주민세, 지방소득세, 취득세로 구성하였다.

#### 2) 세외수입 회수 모형

세외수입 체납자 회수 가능성 예측 모형은 체납 발생일(감면기간경과일)로부터 45일 이내에 자발적 납부 가능성을 예측하는 것으로서, 우량의 기준을 체납 발생 후 45일 이내에 완납한 것으로 설정하였다. 지방세와 마찬가지로 세목을 기준으로 모형을 2개로 구분하여 개발하였다. 세외수입 모형 A는 자동차 관련 세목<sup>7</sup>으로 구성되었고, 모형 B는 자동차 非 관련 세목<sup>8</sup>으로 구성하였다.

### 3) 모형 개발 결과

지방세와 세외수입 각각 2개씩 총 4개의 모형을 모두 체납 정보(내부정보)와 신용정보(외부정보)를 모두 활용하여 검증한 결과 성능이 매우 양호한 것으로 확인 되었다. 그림 11 과 같이 4개의 모형 모두 보통(AUROC 기준 70) 이상의 성능을 유지하였고, 이는 ‘16년 수원시 지방세 체납자 분석 사업 수행 시 개발하였던 모형 대비 예측 성능이 약 5% 포인트 정도 상승한 것이다. 수원시 모형 개발 시 외부 정보인 신용정보만으로 개발하였기에, 모형 개발 시 내 · 외부 정보의 융합을 통해 예측 모형의 성능을 한층 개선시킬 수 있음을 확인하였다.

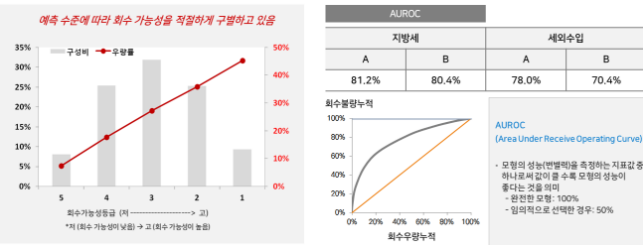
모형 개발 이후, 징수 업무 실무자가 업무상 활용에 용이하도록 5등급 체계로 점수화 · 등급화를 진행

<sup>7</sup> 자동차검사지연과태료, 자동차손해배상보장법위반과태료(자동차의무보험미가입과태료), 장애인주차구역위반과태료, 불법주정차위반과태료, 화물자동차운수사업법위반과징금으로 분석대상 7개 세목 중 5개 세목에 해당한다.

<sup>8</sup> 환경개선부담금 및 그외수입

하였다. 또한, 등급별 일정한 우량·불량률을 유지하여 등급에 따른 징수 업무 전략을 마련할 수 있도록 하였다.

그림 11. 모형 별 성능



### 결론 및 시사점

그림 12 와 같이 현재 전국 지방정부 대부분은 체계적인 기준이 없이 담당자의 주관적인 판단을 통한 비효율적 징수업무를 수행하고 있다. 기본적인 세금 관련 정보를 시스템을 통해 확인 할 수 있으나 징수 활동을 하는 주요 판단을 하는 지표·정보가 부족하였다. 하지만, 향후 징수 활동 시 표준분석모델이 적용된 체납 정보 시스템을 활용하여 회수 등급별 징수 활동에 차이를 둔다면 민원을 감소시키고 세금의 효율적 회수가 가능하다. 이는 시스템과 인프라의 개선을 통해 국가 기관의 재정적 경영의 효율을 가져오는 주요 사례라고 할 수 있다. [11]

그림 12. 현재와 향후의 징수 활동 기준 비교

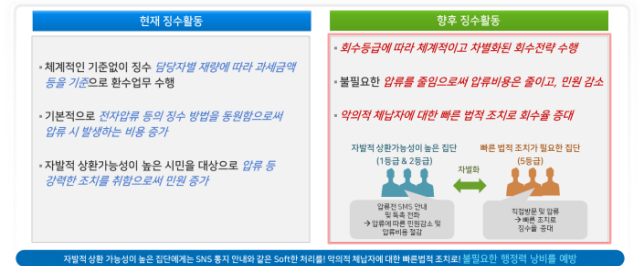


그림 13 과 같이 회수가능등급에 따른 차등적 징수 활동이 가능할 것으로 보이는데, 독촉납입기한 기준 등급별 회수 전략을 세워 징수 활동의 효율성을 높일 수 있다. 즉, 징수 업무 담당자의 시스템 화면에서 회수 가능성 등급 정보를 통해 합리적인 징수업무를 수행할 수 있다. 또한, 독촉 납입 기한 전, 자발적 상환가능성이 높은 집단에는 사전 문자 발송을 통해 민원을 감소시키고 자발적 상환가능성이 낮은 집단에는 독촉전화를 통한 납세 인식 제고가 가능하다.

결과적으로, 독촉 납입 기한 후 자발적 상환가능성이 낮은 집단에 한해 압류 조치와 같은 빠른 법적 조치를 통해 기존방식 대비 조기에 세금을 회수 할 수 있다.

그림 13. 회수 가능 수준에 따른 차등적 징수 활동 예시



### [참고자료] 모델 적용 시 기대효과

최종 산출된 회수 등급을 토대로 실제 징수 업무에 적용 시, 절감 비용과 조기 환수 금액을 계산해 볼 수 있다. 그림 14는 ○시의 시뮬레이션 결과인데, 지방세에서 약 3400만원의 비용절감 효과를 기대할 수 있는 것으로 나타났으며, 5700만원을 환수비용으로 사용하여 약 17억 원의 체납 금액 조기 환수 효과를 기대할 수 있을 것으로 보인다. 세외수입의 경우에는 약 500만원의 비용절감 효과와 500만원을 사용하여 2억 원의 체납 금액 조기 환수를 기대할 수 있을 것으로 확인 되었다.

그림14. 비용 및 환수 금액 시뮬레이션 분석



## 참고문헌

- [1] 지방 통계연감 지방세 자료. (2018)
- [2] 행정안전부 공공 빅데이터 홈페이지. URL: <http://bigdata.go.kr> (2018)
- [3] 변혜정, 체납지방세 관리방안에 관한 연구. 지방세연구원 (2016)
- [4] Koh, Hian Chye, Wei Chin Tan, and Chwee Peng Goh. "A two-step method to construct credit scoring models with data mining techniques." International Journal of Business and Information (2015)
- [5] 손철주, 지방세 체납세금의 효율적인 관리방안 연구. 지방세연구원 (2016)
- [6] Hoelzl, Erik, Maria Pollai, and Bernadette Kamleitner. "Experience, prediction and recollection of loan burden." Journal of Economic Psychology (2009)
- [7] Orgler, Yair E. "A credit scoring model for commercial loans." Journal of money, Credit and Banking (1970)
- [8] Huang, Cheng-Lung, Mu-Chen Chen, and Chieh-Jen Wang. "Credit scoring with a data mining approach based on support vector machines." Expert systems with applications (2007)
- [9] Soltan, A., and M. Mohammadi. "A hybrid model using decision tree and neural network for credit scoring problem." Management Science (2012)
- [10] Thomas, Lyn, Jonathan Crook, and David Edelman. Credit scoring and its applications. Vol. 2. Siam (2017)
- [11] Ravichandran, Thiagarajan, and Arun Rai. "Quality management in systems development: an organizational system perspective." MIS quarterly (2000)