

## 데이터마이닝에 기반한 예비군훈련 입소율 예측에 관한 연구

조상준, 마정목\*  
국방대학교 국방과학학과

### A study on forecasting attendance rate of reserve forces training based on Data Mining

Sangjoon Cho, Jungmok Ma\*  
Department of Defense Science, Korea National Defense University

**요약** 예비군훈련을 담당하는 부대의 임무는 예비군이 평시에 실질적인 훈련을 받을 수 있는 환경을 만들어주는 것이다. 하지만 예비군훈련 담당부대의 특성상 운용 할 수 있는 병력부족의 문제로 실질적인 훈련환경을 만들어주는 예비군 훈련 지원 인력편성에 어려움이 많이 있다. 이러한 이유로 현재 군에서는 전년도 월 평균 예비군 입소율 결과로 당해연도 일일단위 예비군 입소율을 예측하면서 인력편성과 부대운영에 대한 계획을 수립하고 있다. 그러나 기존 예측방법은 실제 입소율과 비교 시 오차가 크게 발생할 수 있다는 문제점을 가지고 있다. 이 문제점은 훈련을 지원하는 교관과 조교 선정에 어려움을 주어 훈련성과 달성에 부정적으로 작용할 수 있다. 그러므로 실제 입소율과 오차를 최소화 할 수 있는 더 정확한 예측모형이 필요하다. 따라서 본 연구에서는 데이터마이닝을 기반으로 일일단위 예비군훈련 입소율을 예측한 모형을 제시하였다. 데이터마이닝 기반 모형의 검증을 위해 예비군훈련 담당부대에서 수집한 실제 데이터로 현재 군에서 사용하는 기존 예측방법과 비교하였다. 그 결과 본 연구에서 제시한 데이터마이닝 기반 예측모형이 기존 예측방법보다 오차를 줄이는 우수한 성능을 보였다.

**Abstract** The mission of the reserve forces unit is to prepare good training for reserve forces during peacetime. For good training, units require proper organization support agents, but they have difficulties due to a lack of unit members. For that reason, the units forecast the monthly attendance rate of reserve forces (using the x-1 year's result) to organize support agents and unit schedule. On the other hand, the existing planning method can have more errors compared to the actual result of the attendance rate. This problem has a negative effect on the training performance. Therefore, it requires more accurate forecast models to reduce attendance rate errors. This paper proposes an attendance rate forecast model using data mining. To verify the proposed data mining based model, the existing planning method was compared with the proposed model using real data. The results showed that the proposed model outperforms the existing planning method.

**Keywords** : Data Mining, Decision Tree, Random Forest, Multiple Linear Regression Analysis, Forecast, Reserve Forces Training, Attendance Rate

---

\*Corresponding Author : Jungmok Ma(Korea National Defense Univ.)

email: jxm1023@gmail.co.kr

Received December 4, 2020

Accepted March 5, 2021

Revised January 5, 2021

Published March 31, 2021

## 1. 서론

### 1.1 연구배경

국방개혁으로 상비병력의 규모가 감소되면서 군에서는 병력 부족 문제를 해결하기 위해 정예화된 예비군 양성 정책 방안을 수립하고 있다. 예비군은 전쟁 개시 전 억제력 확보와 전쟁지속능력 강화에 운용되는 병력으로서 많은 역할을 수행하고 있다[1].

예비군이 병력의 일부로 임무를 수행한 사례는 역사적인 전쟁에서 제시되고 있다. 중세시대부터 예비군은 병력으로 전투임무를 수행했다. 유럽에서 발생한 예나-우어슈테트 전투에서 예비군은 전체 병력의 20% 비율로 편성되어 수적 열세 상황을 극복하도록 도움을 주었다. 현대전 중 하나인 이라크전에서 예비군은 전체 병력의 40% 비율로 편성되어 지상군 부대, 특수작전부대, 항공부대 등 다양한 분야에서 전력으로 활용되었다.

우리나라의 예비군은 2017년 기준 상비병력 대비 4배의 규모이며 상비군의 전투와 지원에 중추적인 역할을 하고 있다[2]. 대규모의 예비전력이 전시에 원활하게 활용되기 위해서는 평시에 충분한 훈련으로 임무수행 능력을 구비해야 한다. 이에 따라 예비군훈련을 담당하는 부대는 예비군이 전투력 수준을 유지할 수 있도록 훈련환경 조성에 많은 노력을 기울이고 있다.

평시에 예비군은 훈련지원 요원(예비군 지휘관, 현역 병력)의 통제를 받아 훈련을 한다. 하지만 예비군훈련 담당부대의 평시 고유임무와 편제부족 문제로 인하여 통제 요원 편성에 어려움이 있다. 이러한 인력문제를 극복하기 위해 예비군훈련 담당부대에서는 사전에 예비군훈련 입소율을 예측하여 훈련지원 요원 규모를 편성하고 있다.

현재 군에서 사용하고 있는 당해연도(x년) 일일단위 예비군훈련 입소율 예측방법은 전년도 예비군훈련 월평균 입소율 결과(x-1년 0월 평균 입소율)를 반영하는 것이다. 기존 예측방법은 당해연도 0월 시작일부터 말일까지 입소율을 동일하게 적용하므로 일일단위 실제 입소율과 비교시 편차가 크게 발생 할 수 있다는 문제점을 가지고 있다. 낮은 예측력은 훈련지원 요원의 비효율적인 편성으로 훈련성과에 부정적인 결과를 줄 수 있다. 따라서 일일단위로 예비군훈련 입소율을 예측으로 예측력을 높일 수 있는 모형에 대한 연구가 필요하다.

### 1.2 관련연구

데이터마이닝은 데이터를 가지고 예측, 분류와 같은 의미있는 결과를 도출하는 분석방법으로 다양한 분야에

서 활용되고 있다[3].

그 중 특히 군에서는 주로 수요예측(수리부속 장비)과 동원병력 자원(예비군) 예측에서 활용되었다. 수리부속장비 수요예측과 관련된 기존연구를 보면 김형태[4]는 육군 ATCAMS 등 3개 장비에 대한 수리부속에 대하여 다중 선형회귀 등 4개의 기법을 적용하여 수요예측 모형을 제시하였고, 윤현민[5]은 해군 DDH-0 구축함의 수리부속에 대하여 회귀나무 등 4개의 기법을 적용하여 수요예측 모형을 제시하였다. 김태규[6]는 공군 항공통제기 E-737 수리부속 장비에 대해 랜덤포레스트 등 3개의 기법을 적용하여 수요예측 모형을 제시하였다.

예비군 병력자원 예측과 관련된 기존연구를 보면 광기효[7]는 2011년 예비군 262명이 참석한 동원훈련 결과를 의사결정나무 등 5개의 기법에 적용하여 동원령이 선포되는 전시를 가정한 응소율 예측모형을 제시하였고, 김세용[8]은 전국적으로 분포된 예비군 270명이 동원령 선포시 동원지정 지역으로 이동하는 경로와 시간을 군집분석 기법을 적용한 시뮬레이션 결과모형을 제시하였다.

동원병력 자원 예측과 관련된 기존연구들은 동원령이 선포되는 전시에 이루어지는 상황을 모의적인 결과로 예측한 것으로 전시의 실제 데이터와 비교하는데 어려움이 있다. 하지만 본 연구는 전시를 대비한 평시 훈련인 예비군훈련에 대한 입소율을 예측하는 것으로 기존연구들과 예측의 목적이 다르다. 또한, 실제 데이터인 예비군훈련 입소율 결과와 비교가 가능하다. 관련 연구들과 본 연구의 차이점을 비교하여 정리하면 Table 1과 같다.

Table 1. Related research

|             | Objective   | Data         |
|-------------|---|--------------|
| H.T.Kim[4]  | Demand of Forecast spare parts                      | Real data    |
| H.M.Yoon[5] |   |              |
| T.K.Kim[6]  |   |              |
| K.H.Kwak[7] | Forecast mobilization rate                          | Virtual data |
| S.Y.Kim[8]  | Analysis of personal mobilization                   |              |
| This paper  | Forecast attendance rate of reserve forces training | Real data    |

### 1.3 연구목적

본 연구의 목적은 기존 예측방법보다 정확한 예측을 하기 위하여 일일단위로 예비군훈련 입소율을 예측하는 것이다. 예비군훈련 입소율은 평균온도와 상대습도, 강우량 차이에 따라서 달라지는 경향이 있다. 하지만 기존 예

측방법은 x년도 0월 모든 훈련일자 의 입소율에 대한 예측을 x-1년 0월 평균 입소율로 적용하므로 일자별 평균 온도, 상대습도, 강우량과 같은 입소율의 영향을 주는 변수에 대한 반영이 어렵다. 그러므로 일일단위 실제 입소율과 비교 시 오차가 크게 발생할 수 있고, 인력편성과 부대운영에 부정적인 영향을 미칠 수 있다.

본 연구에서 적용한 데이터마이닝 기반 예측모형은 예비군훈련 담당부대에서 수집한 예비군훈련 입소율 결과와 기상청에서 획득한 일자별 평균온도, 상대습도, 강우량 데이터를 사용하여 데이터마이닝 (의사결정나무, 랜덤포레스트, 다중회귀분석)에 기반한 일일단위 예비군훈련 입소율을 예측하는 것이다. 예측모형에 대한 성능은 실제 데이터인 예비군훈련 입소율 결과를 기준으로 기존 예측방법과 정확도를 비교하고, 데이터마이닝 기반 예측모형이 더 정확한 결과를 가져왔다는 것을 검증하고자 한다.

## 2. 데이터마이닝 예측모형 및 기법

### 2.1 데이터마이닝 기반 입소율 예측모형

데이터마이닝은 현재 데이터로부터 패턴과 규칙을 발견하여 미래를 예측을 하는 분석방법이다[9]. 당해연도 예비군훈련 입소율에 대한 예측은 전년도 예비군훈련 결과를 바탕으로 이루어지므로 데이터마이닝에 기반하여 입소율에 대한 예측이 가능하다. 데이터마이닝 기반 입소율 예측모형은 전년도 예비군훈련 결과 데이터 입력 단계를 시작으로 데이터 학습, 성능값 도출, 성능평가의 과정을 거치게 되고, 절차는 Fig. 1과 같다.

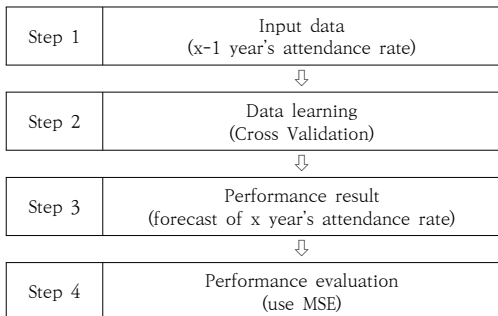


Fig. 1. Attendance rate forecast model using data mining

데이터마이닝 기반 예측모형의 첫 번째 절차는 데이터 입력 단계이다. x-1년도 예비군훈련 결과에 제시된 데이

터 내부 결측치를 제거하고, 실제 입소인원과 입소 대상 인원을 일일단위 입소율로 변환하는 전처리 과정을 통해 정제된 데이터를 입력한다.

두 번째 절차는 데이터 학습 단계이다. 데이터마이닝 기반 모형을 구축하기 위해 일반적으로 데이터 학습에 사용하는 방법은 데이터를 훈련, 테스트, 검증 3개의 집단으로 분류하는 것이다. 하지만 이 방법은 일부 유사 데이터가 특정 1개 집단으로 집중되는 과적합으로 부정확한 결과가 나올 수 있다는 문제점을 가지고 있다. 따라서 데이터를 훈련과 검증 2개 집단으로 분류하고, 모든 데이터가 한 곳에 집중되지 않도록 조정이 가능한 교차검증 방법을 사용했다. 5번의 검증과정을 거치는 교차검증을 예로 들면 Fig. 2와 같다.

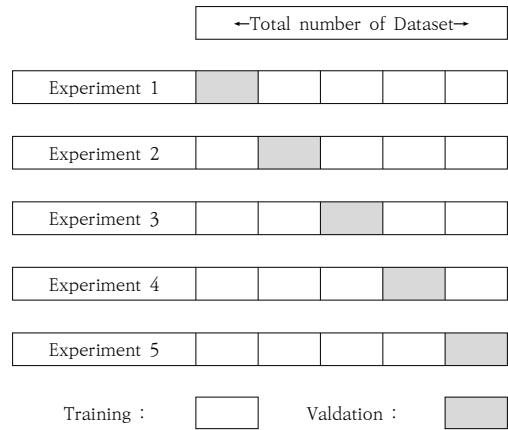


Fig. 2. 5 Folds Cross Validation

Fig. 2에서 제시된 교차검증 방법은 모든 데이터가 각 실험에서 훈련과 검증 비율이 4:1로 나누어지고, 5번의 실험을 통해서 특정집단으로 데이터가 중복되지 않으므로 정확한 학습이 가능하다.

세 번째 절차는 예측모형 성능결과를 확인하는 단계이다. 데이터마이닝에 기반한 예측과 관련된 많은 연구에서는 다양한 기법들이 적용되었지만, 본 연구에서는 의사결정나무(CART), 다중회귀분석, 랜덤포레스트 3가지 기법을 활용하여 x년도 일일단위 예비군훈련 입소율을 예측했다.

마지막 절차는 예측모형 성능평가 단계이다. x년도 예비군훈련 입소율 결과를 기준으로 기존예측 방법과 데이터마이닝 기반 입소예측 모형의 정확도를 비교했다. 정확도 비교는 다양한 지표 중 직관적이고 보편적으로 사용되는 평균제곱오차(MSE)방법을 사용하여 정확도가 가장

높은 예측모형을 선정했다.

## 2.2 데이터마이닝 기법

데이터마이닝은 종속변수의 존재유무에 따라 지도학습, 비지도 학습으로 나뉜다. 지도학습에는 대표적으로 선형회귀분석, 의사결정나무, 인공신경망 등의 기법이 있고 비지도학습에는 연관성분석, 군집분석의 기법이 있다.

Fig. 1의 절차를 적용한 예측모형에서는 9종류의 데이터를 적용하였다. 독립변수는 8종류로 년, 월, 일, 대상인원, 입소인원, 평균온도, 상대습도, 강우량이고, 종속변수는 1종류로 입소율이다. 종속변수가 존재하고 수치형임에 따라 지도학습과 관련된 연구에서 많이 사용한 기법들을 후보로 선정하였다. 그중에서도 변수 영향 정도를 제시할 수 있어서 설명력이 좋은 의사결정나무(CART), 다중회귀분석, 그리고 변수 영향 여부를 알 수 없어 설명력은 다소 낮지만, 예측력이 우수한 랜덤 포레스트 3가지 기법을 데이터마이닝 기반 예측모형에 적용하였다.

의사결정나무(CART) 기법은 독립변수를 이진 분리로 하면서 예측 기준점을 형성하는 방법이다. 변수의 분산정도를 나타내는 지니지수로 독립변수의 중요도를 선정하고 종속변수에 대한 예측과 설명이 가능하다. 데이터마이닝 기반 예측모형은 8개의 독립변수를 가지고 있으므로 지니지수 I(A)에 대한 불순도 지수는 Eq. (1)과 같다.

$$I(A) = 1 - \sum_{k=1}^8 p_k^2 \quad (1)$$

Where, k denotes independent variable,

$p_k$  denotes rate of including class k

다중회귀분석 기법은 다수의 독립변수와 종속변수의 인과관계를 함수식의 형태로 나타내어 예측하는 방법이다. 각 독립변수가 가지는 회귀계수로 중요도를 선정하고, 종속변수에 대한 예측과 설명이 가능하다. 데이터마이닝 기반 예측모형은 8개의 독립변수를 가지고 있으므로 종속변수 Y에 대한 다중회귀분석 식은 Eq. (2)와 같다.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_8 X_8 + \epsilon \quad (2)$$

Where,  $\beta_i$  denotes regression coefficient (i=0 to 8),

$X_i$  denotes independent variable (i=1 to 8),

$\epsilon$  denotes noising coefficient

랜덤 포레스트 기법은 데이터의 분포에 따라 성능 차이가 크고 이진 분리로 과적합의 결점을 가진 의사결정 나무 모형을 보완한 예측방법이다. 무작위로 조합된 다수의 의사결정나무 기법의 예측결과를 평균내는 배깅방법을 사용한다. 조합과정을 알 수 없어서 기법에 대한 설명은 어렵지만, 예측결과와 분산이 작아서 모형의 예측력을 높일 수 있다.

## 3. 데이터마이닝 기반 예측모형 검증

### 3.1 데이터 및 변수선정

예비군훈련 담당부대에서는 전년도 월 평균 입소율을 기반으로 일일단위 입소율을 예측하고 있으며, 일일단위 예비군훈련 결산간 훈련 대상인원, 입소인원, 평균온도, 강우여부 등 관련 내용은 수집 및 기록만 하고 있다. 그러나 이러한 관련 내용들도 예비군 입소율에 영향을 줄 수 있으므로 입력변수(8가지)는 년, 월, 일, 대상인원, 입소인원, 평균온도, 상대습도, 강우량으로 선정하였고, 예측변수는 입소율로 선정하였다.

데이터는 국방동원정보체계에 제시된 예비군훈련 담당부대(0부대)가 '17년(부대개편)부터 '19년(COVID-19 상황 '20년 훈련중지)까지 3년간 실시한 일일단위 예비군훈련 입소결과와 기상청에 제시된 일일단위 평균온도, 상대습도, 강우량을 수집하였다. 수집한 520개 데이터 중 독립변수의 결측치가 발생한 8건은 제거하여 최종적으로 정제된 512개의 데이터를 획득하였다.

### 3.2 검증 시나리오

데이터마이닝 기반 예측모형 검증을 위해 수집한 데이터(512개)의 8종류 입력변수를 기존예측 방법과 동일하게 Table 2와 같이 적용하였다.

Table 2. Original data

| Data | Y   | M   | D   | AT(°C) | H(%) | R(mm) | People |        |
|------|-----|-----|-----|--------|------|-------|--------|--------|
|      |     |     |     |        |      |       | Plan   | Result |
| 1    | '17 | 3   | 2   | 2.1    | 66.8 | 2.6   | 290    | 197    |
| 2    | '17 | 3   | 3   | 2.7    | 67.8 | 0.0   | 204    | 140    |
| 3    | '17 | 3   | 6   | 1.5    | 43.9 | 0.0   | 237    | 164    |
| ...  | ... | ... | ... | ...    | ...  | ...   | ...    | ...    |
| 510  | '19 | 11  | 27  | 4.2    | 65.0 | 0.0   | 250    | 138    |
| 511  | '19 | 11  | 28  | 2.7    | 77.5 | 0.0   | 220    | 119    |
| 512  | '19 | 11  | 29  | 0.8    | 67.6 | 0.0   | 190    | 112    |

Y : Year / M : Month / D : Day

AT : Average Temperature

H : Humidity

R : Rainfall

그리고 기존 예측방법과 동등한 조건에서 성능비교가 되도록 데이터마이닝 기반 예측모형은 x년도 예비군훈련 입소율 예측시 x-1년 예비군훈련 입소결과 데이터만 학습하였다.

데이터와 검증 시나리오를 Fig. 1에서 제시한 데이터마이닝 기반 예측모형으로 구현하기 위하여, Intel(R) Core(TM) i5-8250U CPU와 Windows 10, Python 3.8.2v sklearn 패키지 시스템 환경에서 '18, '19년 예비군훈련 입소율을 예측하였다.

### 3.3 검증 결과

기존 예측방법과 데이터마이닝 기반 예측모형의 x년 일일단위 입소율 예측결과는 Table 3-4와 같다. 실제 x년 예비군훈련 입소율 결과를 기준으로 기존 예측방법과 데이터마이닝 기반 예측모형에 대한 예측결과와 정확도 비교결과는 Table 5-6과 같다.

Table 3은 기존 예측방법의 입소율 예측결과이다. Data 2-3의 결과로 예를 들면, 기존 예측방법은 x-1년 0월 평균 입소율을 x년 0월 모든 훈련일정의 입소율이라고 예측하므로 '18년 3월 5일과 6일 모두 예비군훈련 입소율을 70%로 예측한다는 것을 알 수 있다.

Table 3. Existing forecast model

| Data | Year | Month | Day | Forecast rate(%) |
|------|------|-------|-----|------------------|
| 1    | '18  | 3     | 2   | 70               |
| 2    | '18  | 3     | 5   | 70               |
| 3    | '18  | 3     | 6   | 70               |
| ...  | ...  | ...   | ... | ...              |
| 346  | '19  | 11    | 27  | 75               |
| 347  | '19  | 11    | 28  | 75               |
| 348  | '19  | 11    | 29  | 75               |

Table 4는 데이터마이닝 기반 예측모형의 입소율 예측결과이다. Data 2로 예를 들면, '18년 3월 5일 예비군훈련 입소율을 의사결정나무 모형은 69.4%, 다중회귀분석 모형은 65.0%, 랜덤 포레스트 모형은 67.7%로 예측한다는 것을 알 수 있다.

Table 5는 Table 3-4의 예측결과와 실제 입소율을 비교한 것이다. Data 2로 예를 들면, '18년 3월 5일 실제 예비군훈련 입소율은 68.4%이고, 기존 예측방법은 70%, 데이터마이닝 기법을 적용한 의사결정나무 모형은 69.4%, 다중회귀분석 모형은 65.0%, 랜덤 포레스트 모형은 67.7%로 입소율을 예측하였다.

Table 4. Data mining forecast model

| Data | Year | Month | Day | Forecast rate(%) |      |      |
|------|------|-------|-----|------------------|------|------|
|      |      |       |     | DT               | MLRA | RF   |
| 1    | '18  | 3     | 2   | 58.5             | 64.7 | 67.1 |
| 2    | '18  | 3     | 5   | 69.4             | 65.0 | 67.7 |
| 3    | '18  | 3     | 6   | 66.0             | 64.5 | 68.1 |
| ...  | ...  | ...   | ... | ...              | ...  | ...  |
| 346  | '19  | 11    | 27  | 57.8             | 60.7 | 61.2 |
| 347  | '19  | 11    | 28  | 56.1             | 68.1 | 65.7 |
| 348  | '19  | 11    | 29  | 61.1             | 63.3 | 59.3 |

DT : Decision Tree(CART)  
MLRA : Multiple Linear Regression Analysis  
RF : Random Forest

Table 5. Performance result

| Data | Year | Month | Day | True value (%) | Forecast rate(%) |      |      |      |
|------|------|-------|-----|----------------|------------------|------|------|------|
|      |      |       |     |                | E                | DT   | MLRA | RF   |
| 1    | '18  | 3     | 2   | 67.7           | 70.0             | 58.5 | 64.7 | 67.1 |
| 2    | '18  | 3     | 5   | 68.4           | 70.0             | 69.4 | 65.0 | 67.7 |
| 3    | '18  | 3     | 6   | 68.8           | 70.0             | 66.0 | 64.5 | 68.1 |
| ...  | ...  | ...   | ... | ...            | ...              | ...  | ...  | ...  |
| 346  | '19  | 11    | 27  | 62.9           | 75.0             | 57.8 | 60.7 | 61.2 |
| 347  | '19  | 11    | 28  | 64.1           | 75.0             | 56.1 | 68.1 | 65.7 |
| 348  | '19  | 11    | 29  | 52.1           | 75.0             | 61.1 | 63.3 | 59.3 |

E : Existing forecast model  
DT : Decision Tree(CART)  
MLRA : Multiple Linear Regression Analysis  
RF : Random Forest

Table 6은 Table 5에서 음영으로 제시한 실제 예비군훈련 입소율을 기준으로 기존 예측방법과 데이터마이닝 기반 예측모형의 정확도를 비교한 결과이다. 평균제곱오차(MSE) 지표를 사용하여 기준값과 정확도가 우수한 예측모형 순으로 순위를 선정하였다.

Table 6. Comparison of performance

| Model  | MSE   |       |       |       |
|--------|-------|-------|-------|-------|
|        | RF    | DT    | MLRA  | E     |
| Result | 0.007 | 0.013 | 0.014 | 0.017 |
| Rank   | 1     | 2     | 3     | 4     |

RF : Random Forest  
DT : Decision Tree(CART)  
MLRA : Multiple Linear Regression Analysis  
E : Existing forecast

그 결과, 실제 예비군훈련 입소율과 가장 근접한 예측을 한 모형은 랜덤 포레스트 기법으로 성능이 가장 우수

하였고, 기존 예측모형이 가장 낮은 성능을 보였다. 따라서 본 연구에서 제시한 데이터마이닝 기반 예측모형이 더 우수한 성능을 보인다는 것을 알 수 있었다.

#### 4. 결론

예비군은 전시에 상비군을 지원하기 위해 평시에 훈련을 한다. 하지만 현역병력과 다르게 예비군훈련이라는 특정 기간에만 훈련이 가능하다. 하지만 현재 예비군훈련을 담당하는 현역부대에서는 인력부족 문제로 예비군이 높은 훈련성과를 얻을 수 있도록 도움을 주는 훈련지원 인원 편성에 어려움이 있다.

현재 군에서 적용하고 있는 x년 일일단위 예비군훈련 입소율 예측은 x-1년 0월 예비군훈련 평균 입소율 결과를 0월 모든 훈련일자에 적용하고 있다. 이 방법은 일일 단위 실제 입소율과 비교시 오차 발생으로 인력편성에 비효율적인 결과를 가져올 수 있다.

따라서 본 연구에서는 데이터마이닝에 기반한 일일단위 예비군훈련 입소율 예측모형을 제시했다. 연구의 검증 을 위해 실제 입소율을 기준으로 기존 예측방법과 비교 해 본 결과, 데이터마이닝 기반 예측모형이 더 높은 정확 도를 보여주었다.

하지만 본 연구는 부대개편과 위험도가 높은 전염병 (COVID-19)으로 인하여 많은 양의 데이터를 획득하는 데 어려움이 있었다. 그러므로 데이터 분석 측면에서는 신뢰수준이 높지 않다는 한계점을 가지고 있다. 따라서 분석 신뢰수준을 높일 수 있도록 예비군훈련 결과 데이 터를 지속적으로 축적해야 한다.

데이터마이닝 기반 예측모형을 예비군훈련 담당 현역 부대에서 적용한다면, 정확한 예비군훈련 입소율 예측으 로 현역부대의 안정적인 부대운영과 예비군의 훈련성과 달성에 모두 도움이 될 것으로 판단된다.

#### References

[1] Ministry of National Defense Republic of Korea, Working manual of Reserve Forces, pp.379, Republic of Korea Department of Defense, 2018, pp.4

[2] Republic of Korea Department of Defense, DEFENSE WHITE PAPER, pp.320, Republic of Korea Department of Defense, 2018, pp.104-105

[3] Galit Shmueli, Nitin R. Patel, Peter C. Bruce, Data

Mining for Business Intelligence, pp.460, 2012, pp.24

[4] Hyung Tae Kim, Suhwan Kim, "Data mining based army repair parts demand forecast", *The Korean Data and Information Science Society*, Korea, Vol.30, No.2, pp.429-444, 2019.03  
DOI : <http://dx.doi.org/10.7465/jkdi.2019.30.2.429>

[5] Hyunmin Yoon, Suhwan Kim, "Naval Vessel Spare Parts Demand Forecasting Using Data Mining", *Journal of the Society of Korea Industrial and Systems Engineering*, Korea, Vol.40, No.4, pp.253-259, 2017.12  
DOI : <http://dx.doi.org/10.11627/jkise.2017.40.4.253>

[6] Taegyu Kim, Jungmok Ma, "A Data Mining Approach for Intermittent Demand Forecasting of Aircraft Spare Parts - Focusing on the E-737(AEW&C: Airborne Early Warning & Control) Spare Parts -", *Journal of the Aviation Management Society of Korea*, Korea, Vol.16, No.4, pp.155-164, 2018.08  
DOI : <http://dx.doi.org/10.30529/amsok.2018.16.4.008>

[7] Ki-Hyo Kwak, "Development of the Prediction Model for Mobilization rate Using Data Ming", *Review of Korean Military Studies*, Korea, Vol.1, No.2, pp.148-166, 2012.12  
DOI : <http://dx.doi.org/10.34166/rokms.2012.1.2.148>

[8] Se-Yong Kim, Hoon Young Koo, "Analysis of Wartime Personal Mobilization Using Big-data", *Journal of Korea Society for Simulaion*, Korea, Vol.28, No.4, pp.57-65, 2019.12  
DOI : <http://doi.org/10.9709/JKSS.2019.28.4.057>

[9] Chi-Hyuck Jun, Data mining Techniques, pp.488, Hannarae Publishing, 2012, pp.15-23  
DOI : <http://doi.org/10.9709/JKSS.2019.28.4.057>

조 상 준(Sangjoon Cho)

[준회원]



- 2013년 2월 : 육군사관학교 무기 시스템 공학과 (무기 시스템 공학 학사)
- 2021년 1월 : 국방대학교 국방과 학과(무기체계 석사)

<관심분야>

국방 모델링, 데이터 분석학

마 정 목(Jungmok Ma)

[정회원]



- 2002년 2월 : 육군사관학교 운영 분석학과 (운영분석 학사)
- 2008년 8월 : 미국 펜실베이니아 주립대(PSU) (산업공학 석사)
- 2015년 5월 : 미국 일리노이대(UIUC) (산업공학 박사)
- 2015년 9월 ~ 현재 : 국방대학교 국방과학학과 부교수

<관심분야>

국방 모델링 및 데이터 분석학, 무기체계 획득관리