

# 전자상거래 고객의 클러스터링 분석방법 고찰을 통한 효과적인 군인체력 모형 RSC 클러스터링 분석방법 도출<sup>☆</sup>

## Derivation of an effective military fitness model RSC clustering analysis method through review of e-commerce customers clustering analysis methods

이 준 호<sup>1</sup>                      노 병 인<sup>2</sup>                      신 동 규<sup>3\*</sup>  
Junho, Lee                      Byung-in, Roh                      Dong-kyoo, Shin

### 요 약

군에서 군인들의 체력 관리와 전투력 향상을 위한 필수적인 요소로서, 군인들의 운동 능력, 체력, 건강 수준을 효과적으로 측정하고 모니터링하는 중요성을 강조한다. 체력의 효과적인 측정이 관리의 핵심이며, 이는 현대경영학의 원칙과 일치하고 있다. 특히, 전장의 동적인 변화에 대비하기 위해서는 강인한 체력을 가진 군인을 양성하는 것이 중요하다. 이 연구에서는 RFM(Recency, Frequency, Monetary) 고객 분석 및 클러스터링 방법을 전자상거래에서 검증된 방법으로 소개하며, 이를 기반으로 머신러닝 알고리즘이 적용된 AI 고객 분석 방법을 군인 체력 평가에 적용하기 위해, RSC(Reveal, Sustainable, Control) 분석 모형을 도입하여 군인 체력을 효과적으로 분류하고 모니터링하는 방법을 검토하였다. RFM 기법을 적용한 RSC 분석 모형을 통해 군인 체력을 계량화하고 모델링함으로써 지속적인 발전을 도모하고 체력 관리의 효과를 높일 수 있는 전략을 모색하였다. 이러한 방법을 통해 군인의 체력을 향상하고 유지하는데 있어서 AI 고객 분석 기법을 응용한 RSC 클러스터링 분석방법을 도출하였다.

☞ 주제어 : 군인, 체력, RFM 고객 분석 기법, RSC 분석 모형

### ABSTRACT

This study emphasizes the essential need in the military for effective measurement and monitoring of soldiers' physical fitness, health, and exercise capabilities to enhance both their overall fitness and combat effectiveness. The effective assessment of physical fitness is considered a core element of management, aligning with principles of modern management. Particularly, preparing soldiers with robust physical fitness is deemed crucial for adapting to dynamic changes on the battlefield. In this research, the RFM (Recency, Frequency, Monetary) customer analysis and clustering methods, validated in e-commerce, are introduced as a basis for applying an AI-driven customer analysis approach to assess military personnel fitness. To achieve this, the study explores the incorporation of the RSC (Reveal, Sustainable, Control) analysis model. This model aims to effectively categorize and monitor military personnel fitness. The application of the RFM technique in the RSC analysis model quantifies and models military fitness, fostering continuous improvement and seeking strategies to enhance the effectiveness of fitness management. Through these methods, the study develops an AI customer analysis technique applied to the RSC clustering analysis method for improving and sustaining military personnel fitness.

☞ keyword : military, fitness, RFM customer analysis, RSC analysis model

## 1. 서 론

군에서 군인들의 체력검정 기준을 수립하고 효과적으로 측정 및 모니터링하기 위해서는 체력검정 기준을 선정하는 것이 필요하다. 본 연구는 전투형 강군 육성에 기

1 Department of Computer Engineering, Sejong University, Seoul, 05006, Korea.

2 SOFIZION, Yeongdeungpo-gu, Seoul, 07271, Korea

3 Department of Convergence Engineering for Intelligent Drones, Sejong University, Seoul, 05006, Korea.

\* Corresponding author (shindk@sejong.ac.kr)

[Received 13 October 2023, Reviewed 24 October 2023 (R2 9 November 2023, R3 21 November 2023), Accepted 24 November 2023]

<sup>☆</sup> 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2022R1F1A1074773). 이 논문은 2023년도 세종대학교 교내연구비 지원에 의한 논문임.

여할 수 있는 모델을 개발하기 위해 검증된 RFM(Recency, Frequency, Monetary) 모델을 기반으로 한 RSC(Reveal, Sustainable, Control) 모형을 제안한다.

전자상거래에서 고객의 구매패턴을 이해하는 데 효과적인 RFM 모델은 거래의 최근성, 빈도, 규모 등을 고려하여 고객을 분류하는 간편한 모델링 방법이다. 이를 군체력 모형에 적용하기 위해, ICSPFT(국제체력테스트표준화 위원회)에서 제시한 체력요인을 기반으로 한 RSC 모형을 도입한다. RSC 모형은 운동 발현 능력(Reveal), 운동 지속 능력(Sustainable), 운동 조절 능력(Control)의 세 가지 인자를 통합하여 군인 체력을 효과적으로 분류하고 모니터링하는데 활용될 것으로 기대된다.

이와 같은 연구는 군인의 운동 능력을 정량화하고, RFM 모델과의 효과적인 결합을 통해 군 체력 측정과 관련 시스템을 고도화시킬 수 있도록 제안할 수 있다.

## 2. 관련 연구

RFM 분석 기법은 Hughes(1994)에 의해 개발되었으며, 이는 CRM(고객관계관리)에서 주로 사용되는 고객 세분화 방법 중 하나이다. 이 기법은 데이터베이스에 축적된 고객 데이터를 활용하여 고객들의 상품 구매의 최근성(Recency), 구매 빈도(Frequency), 구매 금액(Monetary) 등을 평가하고 이를 기반으로 각 고객에게 점수를 부여한다. 이를 통해 기업은 고객을 다양한 기준으로 등급화하고, 이를 활용하여 개인화된 맞춤형 마케팅 전략을 수립할 수 있다. RFM 분석은 Hughes(1994)의 제안 이후에도 여러 연구자들에 의해 계속적으로 활용되어 왔다[1].

RFM 분석 기법이 각광받는 이유는 고객의 다양한 니즈와 성향을 고려하여 고객 가치에 기반한 세분화가 필요해졌기 때문이다. 이전에는 기업이 주로 인구통계학적 분류를 사용하여 마케팅을 진행했지만, RFM 분석 기법을 통해 소비자의 과거 구매 이력 정보를 활용하면 더 나은 정확도로 미래의 구매 행동을 예측할 수 있음이 입증되었다[1][2].

RFM 모형은 고객의 구매 패턴을 기반으로 한 간단하고 효율적인 모델로서, 기계학습 기술이 보편화되기 전에도 수학적 통계적 기법을 사용하여 연구되었으며, 그 간단하고 직관적인 특성으로 널리 활용되었다[3][4][5][6].

종합적으로, RFM 분석 기법은 기업이 고객을 세분화하고 효과적인 마케팅 전략을 구사하기 위한 강력한 도구로 자리 잡았다.

4차 산업혁명을 주도하는 인공지능과 기계학습 기술의 발전에 따라, 전자상거래 시스템에서 수집한 데이터의 구매 패턴 분석 기술에 대한 연구와 개발이 다양하게 진행되고 있다[2]. 특히, 최근에는 컴퓨팅 파워의 증가와 기계학습 기술의 급격한 성장을 통해 기계학습을 활용한 K-Means나 DBSCAN 같은 알고리즘을 전자상거래 데이터에 적용하는 연구가 활발히 이루어지고 있다[7][8][9].

### 2.1 K-Means 클러스터링

K-Means 클러스터링은 대규모 데이터 마이닝에서 사용되는 비지도 학습의 대표적인 기술로, 클러스터 분석 방법 중에서도 가장 많이 사용되는 알고리즘 중 하나이다[3]. 이 알고리즘은 사용 편의성 때문에 많은 문제에 쉽게 적용할 수 있다.

각 클러스터에는 하나의 중심이 있다. 각 객체는 가장 가까운 중심에 할당되고 동일한 중심에 할당된 객체는 함께 그룹화되어 클러스터를 형성한다.

- K-Means 클러스터링 방법은 다음 절차로 수행된다.
- 몇 개의 클러스터로 클러스터링할지를 미리 정한다.
  - 정한 클러스터 개수만큼 중심점(Centroid)을 초기화한다.
  - 각 객체는 가장 가까운 중심점에 할당된다.
  - 중심점을 이동하면서 각 클러스터에 속한 객체와의 거리의 평균이 가장 작은 중심점의 위치를 찾는다.

K-Means 클러스터링의 장점은 쉽고, 빠르게 연산이 가능하다는 것이다. 그러나, 몇 가지 단점이 존재한다. 초기에 랜덤으로 설정하는 중심값(Initial point)에 따라 클러스터링 결과가 크게 영향을 받을 수 있고, 노이즈와 이상치(Outlier)의 데이터가 클러스터링 성능에 큰 영향을 미칠 수 있다는 점과 클러스터링할 클러스터 개수인 하이퍼파라미터(hyperparameter) K값을 설정하는 것에 어려움이 있다.

## 3. 전자상거래 고객 클러스터링

### 3.1 전자상거래 고객의 RFM 데이터 분석 방법

전자상거래 고객의 거래 데이터와 고객별 RFM 데이터 유형은 그림 1과 같다. RFM 데이터를 추출하는 과정은 머신러닝 기술이 필요하지는 않으며 데이터베이스의

쿼리나 Python 서버의 함수 호출로 연산이 가능하다.

분석한 데이터는 1년의 거래데이터 이므로 최근구매성의 R값은 분석데이터의 마지막날로부터 최근구매일이다. 따라서, R의 최솟값은 0이고 최댓값은 1년이 365일이므로 364가 된다. F값은 1년 동안의 총 구매 횟수이다. 한 번의 구매로 여러 개의 상품을 구매하였을 때는 구매한 상품의 수만큼 횟수를 산정하였다. M값은 1년 동안의 총 구매금액이다.

거래데이터 유형

구매일	고객아이디	상품아이디	구매금액
2020.2.1	pointmotion	374856948	37,000
2020.2.2	jacksonchoi	298923499	63,000
2020.2.2	pointmotion	893998204	27,000

2021.1.31	luckym79	982639889	36,000
-----------	----------	-----------	--------

고객별 RFM 데이터 유형

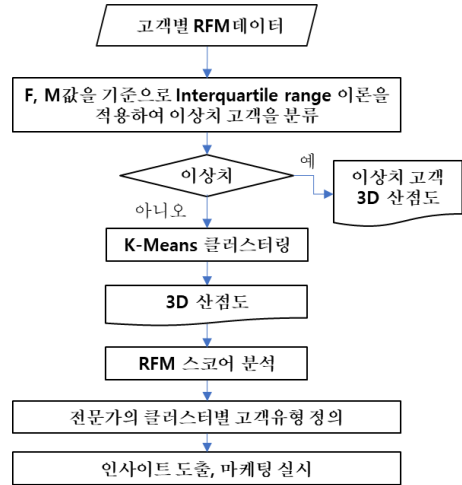
고객아이디	R	F	M
tomy273	100	3	68,000
jacksonchoi	3	9	293,000
pointmotion	273	11	364,000
luckym79	47	2	42,000

(그림 1) 전자상거래 고객의 거래데이터와 고객별 RFM 데이터 유형

(Fig. 1) Types of Transaction Data and Customer-Specific RFM Data in E-commerce (7)

RFM 분석절차는 그림 2와 같다. 생성된 고객별 RFM 데이터에서 F와 M값만 이용하여 IQR(Interquartile range) 이론을 적용하여 이상치고객을 분류해 낸다[9].

이상치고객은 하나의 클러스터가 되고, 이상치가 아닌 고객은 K-Means 클러스터링 한다. K-Means 클러스터링 한 결과를 3D 산점도로 표현하고, RFM 스코어와 비교하여 도메인 지식이 있는 전문가가 각 클러스터의 유형을 결정한다. 이렇게 결정된 각 클러스터는 고객등급으로 분류될 수 있으며 고객등급의 특징을 규정하고 각 등급에 따라 맞춤형 마케팅을 실시할 수 있다.



(그림 2) RFM 분석 절차

(Fig. 2) RFM Analysis Procedure (7)

### 3.2 이상치 제외의 필요성

전자상거래에서 고객 분포와 체력측정에서의 군인 분포는 특성에 따라 이상치의 존재 여부가 다를 수 있다. 전자상거래의 경우 이상치에 해당하는 소수의 고객이 전체 매출의 상당비율을 차지할 만큼 매우 높은 R, F, M 수치를 보일 수 있기 때문에 이상치 그룹 때문에 분포가 크게 퍼지거나 특이한 양상을 보인다.

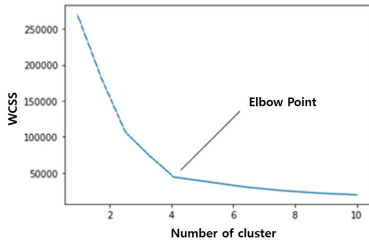
하지만 군인의 측정된 체력 분포는 마치 학교에서 학생의 체력 분포와 유사하게 정규 분포적인 특성을 띤다. 특별한 상황이나 이벤트에 의해 수치가 크게 치우치거나 특이한 양상을 보이기 어렵다. 따라서, 본 연구에서는 이상치는 다루지 않아도 된다.

## 4. 고객분석 시스템 사례 분석

### 4.1 이상치 고객을 제외한 고객의 RFM 분석 결과

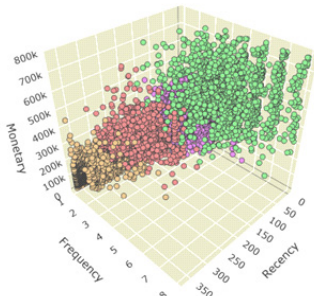
이상치 고객을 제외한 일반고객은 R값과 F값 M값을 Numpy의 percentile 백분위수 계산방식[10]을 사용하여 5점 척도를 구하였다. 클러스터의 개수(K) 결정은 엘보우(Elbow) 기법의 한 형태인 WCSS(Within-Cluster Sum of Square)를 활용하였다. 엘보우 기법은 1부터 10까지의 클러스터 수(K)에 대해 WCSS 값을 계산하며, WCSS는 클러스터 내 각 점과 중심 사이의 거리 제곱의 합이다. K 값에 따른 WCSS를 그래프로 표현할 때, 그림 3처럼 그

래프가 꺾이는 부분이 엘보우로 간주되어 최적의 클러스터 수를 결정한다. 엘보우포인트는 2~3에도 있으나, 본 연구에서는 기준선과 수선거리가 가장 먼 지점에 해당하는 4로 결정되었다.



(그림 3) 엘보우포인트 결정 그래프  
(Fig. 3) Elbow Point Determination Graph [7]

데이터와 파라미터가 준비되었으므로, K-Means 클러스터링을 수행하면 그림 4와 같은 3차원 산점도를 얻을 수 있다. 3차원 산점도를 보면 비슷한 유형끼리 군집을 이루고 같은 색깔로 표시된다. 예를 들어 Recency 가 낮고, Frequency 가 높고, Monetary 가 높은 군집은 우수고객의 특성을 띠므로 우수고객 군집으로 분류될 수 있다.

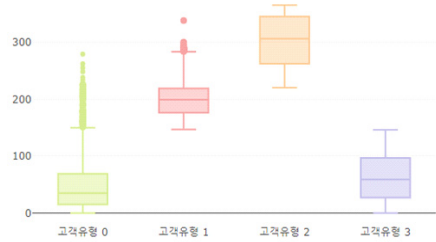


(그림 4) K-Means 클러스터링 이후의 산점도  
(Fig. 4) After K-Means Clustering Scatter Plot [7]

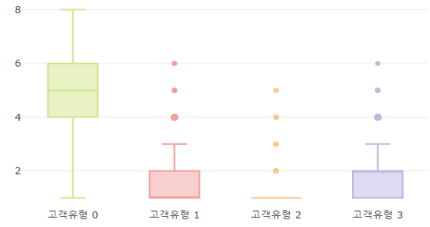
#### 4.2 클러스터별 RFM 범위

클러스터별 Recency, Frequency, Monetary의 범위는 그림 5, 그림 6, 그림 7과 같으며, 상자 수염 그림(Box Plot)으로 표현하였다. 상자 수염 그림은 데이터의 사분위수를 활용하는 특징이 있다. 그림의 중앙의 사각형 영역은 데이터의 50%가 포함되어 있으며, 상자의 위아래의 선은 각각 데이터의 상위 25%와 하위 25%를 가리킨다. 또한, 범위 바깥에 있는 이상치 데이터는 별도의 점으로 표시

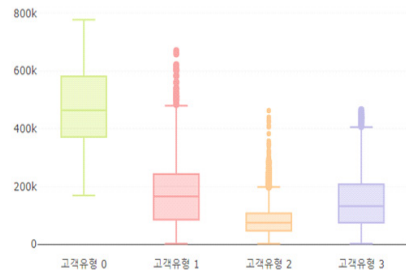
한다. 이 그림은 주로 산점도와 함께 사용되며 도메인 지식을 갖춘 전문가가 각 클러스터의 특성을 파악하는 데에 도움이 된다.



(그림 5) 클러스터별 Recency 범위  
(Fig. 5) Cluster-specific Recency Range [7]



(그림 6) 클러스터별 Frequency 범위  
(Fig. 6) Cluster-specific Frequency Range [7]



(그림 7) 클러스터별 Monetary 범위  
(Fig. 7) Cluster-specific Monetary Range [7]

#### 5. 군인체력 분석 지표

4장에서는 전자상거래 고객의 RFM 클러스터링 분석 방법에 대해 다루었으며, 본 장에서는 RFM 클러스터링 분석 방법을 군인체력 모형에 적용하기 위해서 전자상거래 고객 RFM 요인에 해당하는 군인체력 요인 맵핑 기준을 어떻게 산정할 것인지를 정하는 것이 중요하다. 본 장

에서는 군인체력 능력에 대한 클러스터링 분석을 위해 RFM에 대비되는 RSC의 기준요인에 대해서 알아보고 적절한 군인체력 분석 지표로 제시하고자 한다.

### 5.1 체력의 구성 요소 및 분류

체력 구성 요소의 정의는 사회적, 시대적, 문화적 배경에 따라 혹은 측정대상의 특성, 성별이나 연령에 따라 다양하게 해석했다. 체력의 정의를 명확하게 규명하기는 어렵지만 여러 학자의 의견을 종합해 볼 때, 체력이란 달리기, 뛰기, 던지기, 물건 옮기기 등의 인간 활동의 기초가 되는 신체능력이라 할 수 있다. 체력 구성 요인에 있어서 우리나라는 ICSPFT의 권고 기준에 의하여 Clarke(1971), Ishiko(1980), Cobin(2000)이 제시한 체력요인을 활용하여 표 1과 같이 체력을 8 개 요인으로 구분했다[11].

8 개 체력 요인은 3 개의 운동 능력으로 분류될 수 있기 때문에, 각 운동 능력에 따라 R, F, M과 쉽게 맵핑이 가능하도록 하단 표에 R(Reveal), S(Sustainable), C(control)로 표시하였다.

(표 1) 체력 요인 분류  
(Table 1) Classification of Physical Fitness Factors

운동능력 구분	체력 요인	연관계
운동 발현 능력 (Reveal)	근력	근육계
	순발력	
운동 지속 능력 (Sustainable)	근지구력	호흡계
	심폐지구력	
운동 조절 능력 (Control)	민첩성	신경계
	평형성	
	협응성	
	유연성	관절

### 5.2 체력 요인 분류와 군체력 검정 요인 연결기준

ICSPFT의 8 개 체력요인을 군인체력 분석지표로 활용하기 위해서 국군체육부대(2009)에서 외국군 체력검정요인의 운영종목으로 정리한 내용을 참조하여 매칭시키면, 군인체력 모형지표로 사용할 타당한 근거제시가 될 것으로 판단된다.

선진 외국군 체력검정 종목을 살펴보면 체력요인에 따라 3~5 종목을 운영하고 있으며, 체력검정요인은 표 2와 같다[12].

(표 2) 국가별 체력검정요인 운영종목 (국군체육부대, 2009)  
(Table 2) National Physical Fitness Test Factors and Disciplines

구분	근력	근지구력	전신지구력	순발력	민첩성
한국	팔굽혀펴기	윗몸일으키기	3km 달리기		
미국	팔굽혀펴기	윗몸일으키기	3.2km 달리기		
영국			4.8km 달리기		
프랑스	밧줄타기	수영	7km 달리기	제자리 멀리뛰기, 던지기	100m 달리기
독일	팔굽혀펴기	윗몸일으키기	12분 달리기	제자리 멀리뛰기, 던지기	10m 왕복달리기
일본	팔굽혀펴기		1.5km 달리기	제자리 멀리뛰기, 던지기	50m 달리기
러시아			3km 달리기		100m 달리기
브라질	팔굽혀펴기	윗몸일으키기	2.5km 달리기		100m 달리기
중국		턱걸이	1.5km 달리기	제자리 멀리뛰기, 던지기	60m 달리기
북한		턱걸이	1.5km 달리기	제자리 멀리뛰기, 던지기	100m 달리기

한국과 외국군의 체력검정요인 운영종목에 제시되어 있는 내용을 토대로 R, S, C 군인체력 분석지표 측정을 위한 실제 군인들을 대상으로 측정이 필요한 종목을 표 3과 같이 산정하였다[12].

(표 3) 운동능력별 군인체력측정을 위한 운영종목 기준  
(Table 3) Criteria for Operational Physical Fitness Assessment According to Military Personnel's Physical Fitness Levels

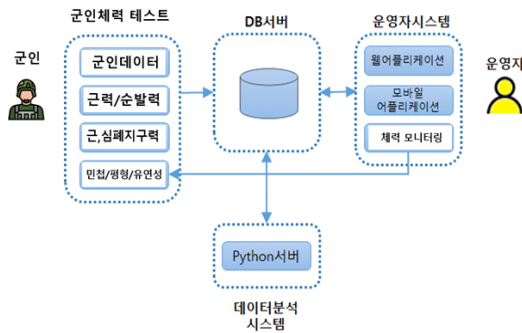
운동능력 구분	체력 요인	연관계
운동 발현 능력 (Reveal)	근력	팔굽혀펴기 제자리멀리뛰기
	순발력	
운동 지속 능력 (Sustainable)	근지구력	윗몸일으키기 턱걸이
	심폐지구력	
운동 조절 능력 (Control)	민첩성	50m달리기 100m달리기
	평형성	
	협응성	
	유연성	

상기 군인체력측정을 위한 운영종목 기준표에 따라 효과적인 군인체력 모형도출을 위해 전자상거래 고객지표인 R, F, M에 해당하는 항목을 R, S, C로 매칭하여, 운동발현능력에 해당하는 R지표로는 팔굽혀펴기와 제자리

멀리뛰기 측정치를 적용하고, 운동지속능력에 해당하는 S지표에는 윗몸일으키기와 턱걸이 측정치를, 운동 조절 능력에 해당하는 C지표에는 50M/100m달리기 측정치를 적용함으로써 군인체력 지표에도 충분히 뽑아낼 수 있음을 확인할 수 있다.

### 6. RSC 군인체력 분석 모형 맵핑

본 장에서는 고객 클러스터링 개념을 군인체력에 대한 클러스터링으로 변환하여, RFM 항목을 군인체력 기준에 따른 RSC 항목으로 맵핑하여 적용함으로써 효과적인 RSC 군인체력 분석모형을 제안하고자 한다.



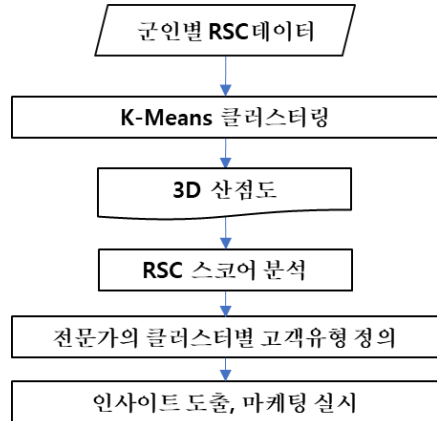
(그림 8) RSC 시스템 구성도  
(Fig. 8) RSC System Configuration Diagram

#### 6.1 RSC 시스템 구성과 분석

RSC분석시스템의 전체 구성도는 그림 8과 같다. 온오프라인 상거래 시스템에서 고객, 상품, 거래데이터 대신에 ICSPFT의 권고기준에 의하여 Clarke(1971), Ichiko(1980), Cobin(2000)이 제시한 체력요인을 측정데이터 항목으로 산정하여 군인들의 체력을 테스트하고, 운동 발현능력에 해당하는 R(Reveal) 값은 각 군인별 근력과 순발력에 대한 측정값을 사용하고, 운동 지속능력에 해당하는 S(Sustainable) 값은 근지구력, 심폐지구력에 대한 측정값을, C(Control) 값은 민첩성, 평형성, 유연성을 반영할 수 있는 측정값을 데이터로 활용하여 DB서버로 수집 및 통합하고, 수집된 데이터는 데이터분석 시스템에서 기계학습 분석을 실행한 후 그 결과를 DB서버에 저장한다.

운영자시스템은 운영자에게 분석결과를 제공하기 위해 웹어플리케이션과 모바일어플리케이션으로 구성된다. 운영자는 군인체력을 모니터링하여 지속적으로 고도화

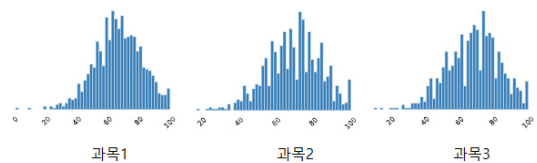
하기 위한 전문가집단으로 기계학습 분석 결과에 대해 라벨링과 적절한 의사결정을 하여 군인체력관리시스템에 반영한다. 이 시스템은 군인들에게 체력을 향상하기 위해 개별군인에게 정기적으로 측정값에 관한 메시지를 보내거나 각 군인에게 필요한 운동종목과 이상적인 측정값이나 향상시켜야 할 구체적 수치 등을 제공할 수 있다. RSC 분석절차는 그림 9와 같다.



(그림 9) RSC 분석 절차  
(Fig. 9) RSC Analysis Procedure

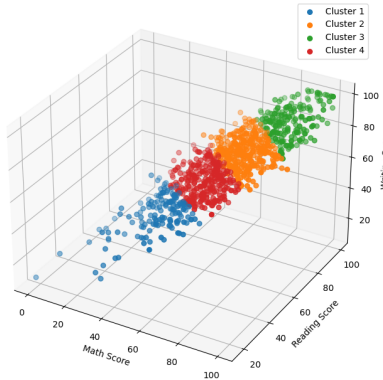
#### 6.2 군인체력별 클러스터 분석

본 연구에서는 군인체력 데이터는 확보하지 못하여 체력데이터와 가장 유사한 특성과 분포를 띄고 있는 체력데이터를 연구데이터로 활용하였다. 체력데이터는 Kaggle에서 공개하고 있는 “Students Performance in Exams” 데이터 세트이다. 이 데이터 세트는 1,000명의 학생이 3가지 과목에서 획득한 체력점수로 구성되어 있으며, 그림 10과 같이 과목별 점수 분포는 정규 분포(가우시안 분포)로 종 모양으로 좌우 대칭적인 분포를 가지고 있다[13].



(그림 10) 과목별 점수 분포  
(Fig. 10) Subject-wise Score Distribution

이 데이터를 K-Means로 클러스터링 하고, 3D 산점도를 그려보면 그림 11과 같다. 각 클러스터를 분석해 보면 Cluster 3에 해당하는 녹색그룹이 체력이 높은 상위그룹이고, Cluster 1에 해당하는 청색그룹이 체력이 낮은 하위그룹이다.



(그림 11) 3과목 성적데이터의 3D 산점도

(Fig. 11) 3D Scatter Plot of Grades Data for Three Subjects

### 6.3 데이터 모형 RSC 클러스터링 분석 결과

클러스터링은 비지도학습이기 때문에 각각의 클러스터가 어떤 유형에 해당하는지는 도메인 지식이 있는 전문가가 판단해야 된다. 만약 분석데이터가 군인의 체력 데이터이고, 각 과목이 R, S, C 측정값에 해당한다면 표 4와 같은 분석결과를 얻을 수 있다.

(표 4) 군인 체력데이터 분석 결과 예시

(Table 4) Analysis Results of Military Personnel Physical Fitness Data

클러스터	군인분류	R	S	C	군인수	비율
Cluster 1	관심 체력	0~64	17~61	10~60	152	15%
Cluster 2	일반체력 상	51~94	60~89	61~87	361	36%
Cluster 3	우수체력	69~100	73~100	73~100	194	19%
Cluster 4	일반체력 하	39~76	47~76	47~75	293	29%

대비되는 RSC맵핑으로 군인을 클러스터링 해 보면 각 클러스터의 특징을 이해할 수 있고, 그 특징에 맞는 적절한 군인체력 모형 분석이나 적절한 운동종목 추천을 제공할 수 있다.

## 7. 결 론

본 연구에서 제시한 RSC클러스터링이라는 새로운 기준의 분석방법은 전자상거래에서 온라인 및 오프라인 거래데이터(최근성, 구매빈도, 구매금액)를 기반으로 한 RFM분석을 통해 고객을 클러스터링 한 기법과 맵핑하여 도출해 낼 수 있었다. 이를 통해 효과적인 군인체력 분류 및 모니터링이 가능하며, 각 군인 유형별 적절한 운동을 추천하고 체력을 향상할 수 있는 군인체력 모형의 효과적인 톨로 활용되기를 기대한다.

비록, 실질적인 군인체력데이터가 아닌 학생 체력데이터를 통해 수집된 RFM데이터를 가상의 RSC데이터로 맵핑하여 측정된 결과이지만, 효과적인 군인 체력 모형을 시뮬레이션하고, 클러스터링 할 수 있다는 것만으로도 많은 인사이트와 새로운 군인 체력 모형의 표본이 될 수 있다고 판단된다.

군에서 군인들의 체력을 측정함에 있어 각각의 능력에 대한 구체적인 측정값들이 측정되고 관리된다면, 본 연구에서와 같은 RSC 맵핑과 같은 관점으로 분석하는 것이 가능하며, 다양한 클러스터링 기술을 접목하여 군인들의 체력을 관리하는데 활용할 수 있게 된다. 이외에도 특히, 본 논문에서 사용한 군인체력 분류 기술은 군인들의 질병에 대한 면역능력과 환경변화에 대한 적응능력, 정신적인 저항능력을 관리함으로써 질병이나 사고를 예방하는 방법으로도 활용될 수 있을 것이다.

머신러닝 클러스터링 기법은 3차원의 데이터뿐만 아니라 매우 많은 차원의 데이터도 쉽게 분류할 수 있기 때문에, 체력데이터를 좀 더 세분화한 데이터와 기타 정신적인 저항능력 데이터도 추가하여 군인의 신체적 정신적 특성에 따라 좀 더 세분화된 정보를 제공할 수 있다.

따라서, 본 연구에서의 RSC 클러스터링 분석방법은 군인체력 관리뿐만 아니라 군인의 종합적인 건강과 신체적, 정신적 측면에서의 강화에 기여할 수 있는 다양한 응용 가능성을 제시하고 있다.

## 참고문헌(Reference)

[1] HeuiJu Chun, "Customer Lifetime Value Model Using Segment-Based Survival Analysis," Communications for Statistical Applications and Methods, Vol. 18, No. 6, pp. 687-696, 2011.  
<https://doi.org/10.5351/CKSS.2011.18.6.687>

- [ 2 ] DongSeok Kim, “Data driven selection methods of weights in RFM Model,” 62, Jeju National University Master’s Thesis, 2021.  
<https://oak.jejunu.ac.kr/handle/2020.oak/23663>
- [ 3 ] TaeJin Kim, “Proposal Methodology for Disaster Risk Analysis by Region Using RFM Model,” *Journal of the Society of Disaster Information*, Vol. 16, Issue 3, 2020, pp. 493-504, 2020.  
<https://doi.org/10.15683/kosdi.2020.09.30.493>
- [ 4 ] YoungSung Cho, “Implementation of Personalized Recommendation System using RFM method in Mobile Internet Environment,” *Journal of The Korea Society of Computer and Information*, Vol. 13, No. 2, 2008, pp. 41-50, 2008.  
<http://journal.kci.go.kr/jksci/archive/articleView?artiId=ART001237406>
- [ 5 ] Birant, Derya. “Data mining Using RFM Analysis,” *International Journal of Scientific & Engineering Research*, Vol. 4, Issue 12, 2013, pp. 940-943, 2013.  
<http://doi.org/10.13140/RG.2.2.24229.04328>
- [ 6 ] YoungHwan Choi, “Studies on Types of Korean Professional Baseball Spectators Using RFM,” *Journal of Sport and Leisure Studies*, Vol. 50, 2012, pp. 289-299, 2012.  
<https://doi.org/10.51979/kssls.2012.12.50.289>
- [ 7 ] ByungIn Roh, “Method of extraction and processing of outliers in the customer’s RFM model”, 48, Sejong University Master’s Thesis, 2022.  
<https://library.sejong.ac.kr/search/DetailView.ax?sid=1&cid=1769889>
- [ 8 ] HyunJung Ji, “A Study on Customer rating using RFM and K-Means,” *Advances in Computer Science and Ubiquitous Computing*, pp. 823-828, 2017.  
[http://doi.org/10.1007/978-981-10-7605-3\\_131](http://doi.org/10.1007/978-981-10-7605-3_131)
- [ 9 ] Siti Monalisa, “Analysis Outlier Data on RFM and LRFM Models to Determining Customer Loyalty with DBSCAN Algorithm,” *IEEE*, pp. 1-5, 2018.  
<https://doi.org/10.1109/SAIN.2018.8673380>
- [10] JongSuk Lee, “Sequential Percentile Estimation for Sequential Steady-State Simulation,” *The KIPS Transactions:PartD*, Vol. 10D, Issue 6, pp. 1025-1032, 2003.  
<https://doi.org/10.3745/KIPSTD.2003.10D.6.1025>
- [11] DuckHyun Nam, “Study on Settlement of Army Fitness Test for the Purpose of Barracks Physical Training Stimulation”, *The Korean Journal of Defense Analysis*, Vol. 27, No. 4, 2012, pp. 149-179, 2012.  
<http://doi.org/10.22883/jdps.2012.27.4.006>
- [12] DuckHyun Nam, “Study on the Selection of the Korean Combat Fitness Factors: Based on the Case of American Military Combat Fitness,” *The Korean Journal of Defense Analysis*, Vol. 28, No. 3, pp. 175-196, 2012.  
<http://doi.org/10.22883/jdps.2012.28.3.006>
- [13] Lewin, Daniel R., “What Can we Learn from Exam Grade Distributions?,” *International Journal Scholarship of Teaching & Learning*, 12, Vol. 15, No. 2, 2021.  
<https://doi.org/10.20429/ijstol.2021.150207>



● 저 자 소 개 ●



**이 준 호(Jun-ho Lee)**

1994년 경희대학교 전자계산공학(이학사)  
2012년~현재 소프트자이온 대표이사  
2023년~현재 세종대학교 대학원 컴퓨터공학과(박사과정)  
관심분야 : 고객관리시스템, 머신러닝, 빅데이터, 클라우드 서비스, etc.  
E-mail : ssing@softzion.com



**노 병 인(Byung-in Roh)**

1994년 경희대학교 전자공학과(이학사)  
2022년 세종대학교 컴퓨터공학과(석사)  
2014년~현재 소프트자이온 연구소장  
관심분야 : 고객관리시스템, 머신러닝, 빅데이터, 클라우드 서비스, etc.  
E-mail : rohbi@softzion.com



**신 동 규(Dong-kyoo Shin)**

1986년 서울대학교 계산통계학과(이학사)  
1992년 Illinois Institute of Technology 대학원 컴퓨터과학과(공학석사)  
1997년 Texas A&M University 대학원 컴퓨터과학과(공학박사)  
1998년~현재 세종대학교 컴퓨터공학과 교수  
관심분야 : 머신러닝, 유비쿼터스 컴퓨팅, 생체신호 데이터처리, 정보보호, etc.  
E-mail : shindk@sejong.ac.kr