

데이터 마이닝 기반의 수리부속 수요예측 연구

A Study on Forecasting Spare Parts Demand based on Data-Mining

김재동¹
Jaedong Kim

이한준*
Hanjun Lee

요약

수리부속 수요예측은 장비가동률 향상과 국방 운영 예산 효율화 제고를 위한 국방 군수 분야의 핵심 과제 중 하나이다. 현재 우리 군은 수리부속 소요 데이터를 활용한 시계열 기법으로 과거 데이터 분석을 통해 수리부속 수요예측에 활용하고 있으나 정확도 제고에 지속적인 노력이 요구되고 있는 실정이다. 이에 본 연구에서는 지난 5개년의 수리부속 18,476개 품목의 수요데이터를 수집하고 데이터마이닝 기법을 활용한 수리부속 수요예측 모델을 제안하였다. 제안한 모델에 따른 실험 결과는 기존 시계열 기법에 비해 개선된 수요예측 정확도를 보였다.

☞ 주제어 : 수요예측, 데이터 마이닝, 군수, 수리부속

ABSTRACT

Demand forecasting is one of the most critical tasks in defense logistics, because the failure of the task can bring about a huge waste of budget. Up to date, ROK-MND(Republic of Korea - Ministry of National Defense) has analyzed past component consumption data with time-series techniques to predict each component's demand. However, the accuracy of the prediction still needs to be improved. In our study, we attempted to find consumption pattern using data mining techniques. We gathered an 18,476 component consumption data first, and then derived diverse features to utilize them in identification of demanding patterns in the consumption data. The results show that our approach improves demand forecasting with higher accuracy.

☞ keyword : demand forecasting, data mining, logistics, spare part

1. 서론

일반적으로 수요예측(Demand Forecasting)이란 ‘시장 환경변화 예측을 토대로 판매 가능한 잠재시장의 규모를 도출하고 언제 얼마만큼 팔 것인가를 전략적으로 결정해 가는 창조적 작업 과정’으로 정의된다[1]. 다시 말해 미래의 수요를 준비하는 일련의 과정이라 할 수 있다.

최신 군은 급격히 증가하는 신형·장기 운영 무기체계의 고장 정비에 따른 수리부속 교체 및 수리에 대한 신속한 대응이 요구된다. 이에 따라 수많은 수리부속을 적기·적소에 공급하기 위해 정확한 수요예측을 필요로 하고 있다. 수리부속 수요예측 정확도는 재고 감축과 정비시

간 단축 등 군수지원 성과 및 전투 준비태세에 직접적인 영향을 미치는 중요한 요소이다. 수리부속 수요예측 정확도가 저조하여 재고 고갈 시, 장비 정비 지연에 따른 장비가동률 저하를 초래하며, 재고 과잉 시, 한정된 국방 예산의 효율적 활용으로 긴급한 타 장비 혹은 수리부속의 획득 차질 및 장기 비수용 수리부속 재고자산 증가를 야기할 수 있다[2]. 이 때문에 우리 군은 수리부속 수요예측 정확도 제고에 지속적인 노력을 기울여 왔다. 현재 사용되는 수리부속 수요예측 기법은 시계열 기법을 기반으로 하고 있으며, 이동평균법, 지수평활법 등 3~8개의 여러 기법이 활용되고 있다. 하지만 정확도 측면에서, 개선의 여지가 많은 실정이다.

이에 본 연구에서는 현재 공군에서 운영하는 항공기 수리부속에 대한 수요예측 정확도를 제고하고자, 최근 빅데이터 분석과 함께 주목을 받고 있는 데이터 마이닝 기법을 적용하여 분석하고자 한다[3]. 데이터 마이닝 기법은 수요예측 분야에 적용사례가 많지 않으며, 특히 군수 분야에서는 사례가 드물어 연구가 필요한 분야라 할 수 있다.

1 Center for Defense Management, Korea Institute for Defense Analyses, 02455, 37 Hoegi-ro, Dongdaem-gu, Seoul, Korea

2 Center for Defense Acquisition, Korea Institute for Defense Analyses, 02455, 37 Hoegi-ro, Dongdaem-gu, Seoul, Korea

* Corresponding author (hjlee1609@gmail.com)

[Received 27 October 2016, Reviewed 28 October 2016, Accepted 21 November 2016]

한편, 우리 군은 기존 정보시스템을 개선하여 수요 관련 자료 뿐 아니라 항공기 운영시간, 소티(sortie)수* 등과 같은 수리부속별 특성 자료까지 포함하여 관리하고 있다[4]. 그러나 아직까지 이러한 데이터들은 기존 수요 예측에는 활용되지 못하고 있다. 본 연구에서는 수리부속 수요 관련 데이터를 포함하여 앞서 언급한 수리부속별 특성 자료를 분석에 활용하였다.

본 논문은 총 5장으로 구성하였다. 2장에서는 수요예측과 데이터 마이닝 분야의 관련 연구들을 소개한다. 3장에서는 분석을 위한 모델 수립에 대해 설명한다. 4장에서는 데이터 마이닝의 기법별 실험 결과를 분석하며, 5장에서는 논의를 요약하고 본 연구의 한계와 향후 연구 방향을 제안한다.

2. 문헌 연구

2.1 수요예측

본 연구에서는 그 동안의 주요 수요예측 분야 연구를 적용기법의 성격에 따라 정량적 분석 방법, 정성적 분석 방법과 시스템 접근 방법으로 세 가지로 크게 분류하였다. 각 분류별 분석기법의 종류와 특징은 다음과 같다.

첫째 정량적 분석 방법은 과거의 데이터로 패턴이 미래에도 지속된다고 판단되어질 때 적합한 방법이다. 대표적인 기법으로 회귀분석 기법, 시계열 기법과 확산모형 기법으로 세 가지 기법으로 분류할 수 있다.

먼저 회귀분석 기법은 변수 사이의 인과관계를 분석하는 통계방법으로 분석 대상의 정확한 데이터 확보가 중요한 기법으로 여러 분야에 적용되고 있다[1]. 기존 연구에서는 경상북도의 다도시의 과거 지역전력 수요량, 지역 경제 및 지역인구 등의 데이터를 입력변수로 사용하여, 정확도 분석을 통해 변동성이 심한 지역의 경우 적합한 전력수요예측을 하였다. 이처럼 외부 환경의 변수를 활용하여 민감도 분석뿐 아니라 불확실한 상황을 고려하는 시뮬레이션에도 활용이 가능하다[5].

시계열 기법은 과거 데이터 수집이 용이한 분야와 예측 목적으로 개발된 대표적인 모형으로 외부 충격에 의한 반응을 장기적으로 파악할 수 있는 장점이 있다. 과거 연구로는 과거 철도 수요 데이터를 활용한 X-12-ARIMA 모형을 활용하여 철도수요의 요일 및 계절 변동의 변수를 이용하여 변동성을 예측하였다[6]. 또한, ARIMA 기법

을 기반으로 1983년부터 2008년까지의 프로축구 관객 수 자료를 활용하여 한국 프로축구 관객 수 예측을 하였다[7].

끝으로 확산모형 기법은 신제품이나 신기술에 대한 수요 예측 및 과거 데이터 수집이 불가능한 분야나 초기 데이터만을 활용이 가능한 상황에서 적합한 기법이다. 최근 연구에는 전기자동차의 미래 수요예측을 현재 시장규모에서 시간의 흐름에 따라 변화하는 시장규모를 추정 및 설정하여 확산과정 모델링과 효과를 도출하였다. 이와 같이 초기 데이터, 전문가의 경험 및 유사 사례 적용 등으로 미래의 시장 규모의 예측이 가능하다[8].

둘째 정성적 분석 방법은 과거 데이터 수집이 불가능하거나 불충분할 때 적합한 방법으로 주로 전문가의 판단에 의존하여 결과를 도출한다. 전문가 의견 활용 분석 기법, 컨조인트 분석 기법과 인덱스 분석 기법 등이 이 분류에 포함될 수 있다. 먼저 전문가 의견 활용 분석 기법은 데이터 수집이 불가능한 상황 혹은 해당 분야의 적합한 전문가 확보가 가능 때 적용하는 기법이다. 대표적인 기법으로는 델파이 기법으로 1950년대 미국 랜드연구소(Rand Corporation)에서 개발되었으며, 냉전시대에 나타나는 문제점을 해결하기 위한 연구방법으로 주목받기 시작했다. 그 이후 미래를 전망하는 방법으로 CATV, TV회의와 같은 새로운 서비스를 예측 등과 같이 여러 분야에서 활용되고 있다. 이와 같이 방식으로 델파이 방법은 전문가가 있는 분야에서 다양하게 활용되고 있다[9]. 다음으로 컨조인트(Conjoint) 분석 기법은 제품의 기능이나 속성별 필요성을 파악하여 신제품의 시장 반응을 예측 시 사용되는 기법이다. 관련 연구는 소비자의 구매 제품에 대하여 우선순위가 반영된 자료를 토대로 소비자의 구매 과정을 분석한다. 관련 데이터를 기반으로 연속형 선호도 자료를 추출하여 분석하는 방식으로 1970년부터 소비자 와 관련된 다양한 분석에 활용될 수 있도록 이론적 기초를 제공하였다[10]. 이러한 특성을 갖고 있는 컨조인트 기법은 정확도가 높아 기업에서 널리 사용 중이며 특히 서비스 분야에서 고객 분석 시 사용된다[1]. 마지막으로 인덱스 분석 기법은 공산품보다 부동산, 프로젝트 등 희소 제품의 선택 가능성에 적합한 특징을 갖고 있다. 이러한 기법의 특성을 이용하여 당선 가능성이 높은 후보 예측에 많이 활용되었으며, 1896년부터 2008년까지의 29개의 미국 대선 후보 데이터로 59개의 평가지표를 적용하여 최종 영향 변수와 후보별 선택 가능성을 예측하여 제시하였다[11].

셋째 시스템 접근 방법은 여러 변수 데이터를 기반으로 지식을 학습하여 결과를 도출하는 방법으로 수요와

* 소티 수: 전투기 출격 횟수

영향 요인들 간의 관계가 매우 복잡한 경우 적용하는 방법이다. 대표적인 기법으로는 정보 예측 시장 기법, 시스템 다이내믹스 기법과 인공 신경망 기법으로 나누어 볼 수 있다. 먼저 정보 예측 시장(Prediction Market) 기법은 선물 및 주식시장에서 투자환경 및 투자자들의 행동을 토대로 의사결정 시스템을 구축하여 예측하는 기법이다. 특징으로 큰 비용 투자 없이 예측을 할 수 있으며, 행동을 하는 참여자가 많아야 의미가 있는 정보 추출이 가능하다[1]. 다음으로 다이내믹스 기법은 변수와 변수들 간에 연속적인 인과관계를 현실적인 모형으로 전환하고 시물레이션을 통해 시간의 흐름에 따른 변화과정 및 관련 정보를 분석하는 기법이다. 기존 연구에서는 항공기 수요를 둘러싼 요인들의 인과관계의 관계를 통하여 항공기 수요 예측을 하였으며, 생산능력 증감의 의사 결정을 지원하였다[12]. 또한, 1983년부터 2000년까지의 물동량 자료를 기반으로 인과관계를 통한 현실적인 모형 구축으로 홍콩항구의 물동량을 예측하였다[13]. 마지막으로 인공 신경망 기법은 복잡한 신경망 구조와 학습방법을 모방해 데이터들 간의 패턴을 찾아내고 적용하여 예측하는 기법이다. 적용된 논문으로는 도시의 전력 수요 패턴의 도출을 위해 온도, 습도, 풍속과 일사량을 변수로 활용하여 전력 수요 예측을 수행한 연구가 있다[14]. 이와 같이 인공 신경망 기법은 인과관계가 복잡하고 많은 데이터 분석이 필요한 수요예측 문제에 적합하다.

한편, 우리 군은 (표 1)과 같이 정량적 기법 중 시계열 기법 중심으로 1980년 이래 대부분 군수 분야에서 활용되고 있다. 특히 수리부속(9종) 수요예측 기법을 각 군별 현재까지 활용되고 있고, 현재 우리 군의 수요예측 정확도는 70~75% 수준이다[15]. 육·해군은 시계열 기법의 결과로 품목관리자가 수정하여 최종적인 수요예측 판단을 하며, 공군은 Y-1년도 수요예측 실시 후 Y년도 실제 수요와 비교하여 품목관리자들이 결정하는 기법을 적용하고 있다. 최근 해군의 기법 중 가중이동평균법 외 4개는 추가 개정되었다[16].

그동안 군에서 진행된 연구로는 수리부속 수요예측 정확도 제고를 위해 1999년부터 2007년까지의 수요 자료를 활용하여 공군 수요예측 기법과 ARIMA 기법을 비교분석하여 정확도를 제고한 바 있다[3]. 또한, 창 소요자재 예측 연구로서 주정비 일정계획, 자재명세서와 교환율 등을 사용하여 운영수준량과 안전수준량을 예측한 연구도 진행되었다[17]. 그 밖에 수요예측 관련 연구로, 수리부속 수요의 특성을 분석하여 계층적 수요예측법과 직접수요 예측법의 대한 상대적 퍼포먼스와의 가변성, 왜도, 첨도

(표 1) 각 군의 수요예측 적용기법

(Table 1) Demand forecasting techniques of each service

군	적용 기법	판단방법
육군	산술평균법 이동평균법 최소자승법	회귀식 기술기와 편차율이 프로그램 상 제안 후 품목관리자 수정
해군	산술평균법 이동평균법 최소자승법 가중이동평균법 단순지수평활법 이중지수평활법 홀트지수평활법 윈터지수평활법	8개 기법 적용 후 연도별로 표시된 그래프를 통해 품목관리자 결정(적용 기법 일부 추가됨)
공군	가중이동평균법 선형이동평균법 추세분석법 단순지수평활법 선형지수평활법	Y-1년도 수요예측 실시 후 Y년도 실제 수요와 비교하여 수요예측 정확도 높은 기법 적용

의 상관관계를 분석으로 장비 특성별 수요예측기법의 선택 방법을 제시하였다[18]. 이와 같이 군에서도 수요예측 연구가 꾸준히 진행되어 왔다.

그러나 현재 군에서 적용되는 기법은 과거 수요 데이터 변수만으로 미래를 예측하는 시계열 기법을 사용 중이다. 현재 우리 군이 운용중인 정보시스템에는 수요자료 외에 많은 정보가 수록되어 있다. 예를 들어 항공기 장비의 경우에는 운영시간, 소티(sortie)수, 조달기간, 보급기간 등과 같은 데이터가 축적되어 있다. 이러한 데이터를 활용하여 수요 데이터 변수만을 의존하는 방식을 벗어나 시스템 접근 방법과 같이 수요와 영향 요인들 간의 관계를 분석한 방법론이 절실히 요구된다.

2.2 수요예측 분야 데이터마이닝

데이터 마이닝은 대규모의 데이터에서 탐색, 분석, 모델링의 과정을 거쳐 인자(feature)와 인자와의 관계나 패턴, 규칙 등을 추출함으로써 의사결정 지원에 도움을 주는 분석 방법이다[19,20]. 데이터 마이닝 분석 기법에는 의사결정 나무 기법, 인공신경망, 베이지안넷 등 여러 기법이 있으며 지금까지 다양한 분야에 폭넓게 활용되어 왔다[19-21].

반면, 수요예측 분야에서는 데이터마이닝 적용이 상대적으로 활발하지 못했다. 대표적인 논문을 살펴보면 고객의 주문정보로부터 주문한 선박 부품내역에 대한 데이터

를 통하여 재고관리 품목의 거래 빈도, 품목 간 연관성을 분석하여 품목의 출고준비에 소요되는 시간을 줄인 바 있으며[22], 물 사용량을 예측을 위하여 3년간의 일일 수요량의 데이터로 요일별 특성을 포함한 계절적 특성 등에 대한 상관분석 후 다중선형 자기회귀모델과 칼만필터 모델로 수요예측 결과를 도출하기도 했다[23]. 이와 비슷한 연구로 기상 데이터의 일조시간, 강수량, 적설량, 평균기온 등의 데이터로 요일별 특성, 계절적 특성 등에 대한 상관관계 분석을 하였다. 도출된 결과 데이터로 연관분석, 군집분석과 이상치 탐지를 통해 데이터 마이닝 기법을 적용하여 단기 물 수요예측의 결과를 도출하였다[24]. 최근에는 데이터마이닝기법을 이용한 수요예측 연구로 농산물 유통시장의 부정거래 예측하는 모형은 6개 농산물 도매시장법인으로부터 32개월 청과거래 정산데이터를 수집하여 경매장 구분, 경매구분, 품목코드 등 34개의 변수 중 16개 변수를 선정하여 데이터마이닝 기법인 의사결정 나무기법을 이용하여 허위거래를 탐지하는 모형을 제시하였다[25].

이와 같은 기존 연구처럼 데이터마이닝 기법은 고객의 주문정보로부터 거래빈도, 품목 간의 관계 분석 연구 및 기상과 일별·계절별 특성들 간의 연관성 분석을 통해 다양한 분야에서 미래 지향적 의사결정 지원 목적으로 활용되다가 최근 빅데이터와 함께 각광을 받고 있다. 그러나 현재 우리 군은 과거 수요 데이터만을 활용하는 시계열 기법에 의존도가 높다. 과거에는 수요 데이터 외에는 활용하기 어려운 정보시스템이었으나 2009년 이후 장비 정비정보시스템이 구축되면서 보다 많은 정보의 축적으로 수요 데이터를 제외한 변수도 활용이 가능해졌다. 본 연구에서는 이러한 변수들은 수요 예측 정확도 제고에 활용될 수 있을 것으로 판단, 아직까지 적용해 본 적이 없는 데이터 마이닝 기법을 군수 분야 수리부속 수요예측에 적용하고자 한다.

3. 수리부속 수요예측 모형 제안

3.1 데이터 수집 및 변수추출

우리 군에서는 2009년부터 무기체계의 주 장비 중심으로 수리부속의 보급 및 정비 관련 사항을 최신 정보 기술을 이용하여 개발한 정보시스템을 구축하여 육·해·공군의 편성 부대로부터 국방부에 이르기까지 정비 관련 부서에서 사용하는 통합정보지원 시스템인 장비정비정보시스템(Defense Logistics Integrated Information System)를 운영하고 있다.

본 연구 대상 항공기는 우리나라와 함께 여러 나라에서 운영 중인 주력기로 항공기 예산 중 큰 비중을 차지하고 있으며, 전쟁발발 시 공중우세와 작전지속 능력향상에 전장의 핵심 기능을 수행하는 공군 항공기인 F-XX를 연구 대상 장비로 선정하였다.

공군 장비정비정보시스템의 정비데이터베이스에서 1,053,422개 트랜잭션 데이터를 수집하였고, 수집된 데이터에는 정비날짜, 정비당일 소모된 수리부속품목의 개수, 대당구성수 등 22개 항목이 포함되어 있었다.

본 연구의 목적은 수리부속별 수요예측이므로 데이터를 품목별로 재정리하였으며 수리부속 품목수는 총 18,476개였다. 이 과정에서 수리부속별로 2011년부터 2015년까지 5개년 간의 연도별 소모 데이터 및 조달기간, 표준단가, 그리고 항공기 운영시간과 소티수를 추출하였다. 추가적으로 수리부속당 수리율(수리된 개수/(장비입고량 × 대당구성수))과 폐기율(폐기된 개수/(장비입고량 × 대당구성수))을 보급거래 자료와 함께 가공하여 분석에 추가하였다. 최종적으로 활용한 변수는 (표 2)와 같이 총 16개의 변수였다.

(표 2) 변수(feature) 설명
(Table 2) Feature description

변수	의 미
'10년 비율	'10~'14년 수리부속(품목별/연도별) 불출 개수 대비 소모 개수 비율(소모 개수 합/불출 개수 합)
'11년 비율	
'12년 비율	
'13년 비율	
'14년 비율	
'10년 소모개수	'10~'14년 수리부속 품목별 연도별 소모 개수 합
'11년 소모개수	
'12년 소모개수	
'13년 소모개수	
'14년 소모개수	
수리율	정비입고 된 요수리품 중 사용가능 재고로 환원된 품목의 비율
폐기율	정비입고 된 요수리품 중 폐기된 품목의 비율
조달기간	구매요청서를 낸 때로부터 해당 보급품을 100% 수령할 때까지의 행정소요시간
표준단가	군수품에 대한 자산액을 파악하거나 취득 또는 수령하는 물품의 가치판단에 적용하기 위해 군수품 표준단가 업무 훈령에 산정된 단가
항공기 운영시간	항공기 가동 시간(비행시간(hour)만 포함)
항공기 소티수	항공기 이·착륙 횟수

목표변수는 품목별 2015년 소요발생 여부 데이터로서 전체 18,476개의 품목 중 10,510건에 대해서는 소요가 발생하였고 7,965건에 대해서는 소요가 발생하지 않았다. 소요발생 여부에 대한 목표변수 갯수의 밸런싱을 위하여 소요가 발생한 품목 10,510건 중 무작위로 7,965건을 선택하여 소요가 발생하지 않은 7,965건과 함께 15,930개의 품목으로 구성된 데이터셋을 구성하였다. 이 때, 무작위 선택으로 인한 영향을 확인하기 위해 무작위 선택을 5회 반복하여 총 다섯 개의 데이터셋을 연구에 사용하였다.

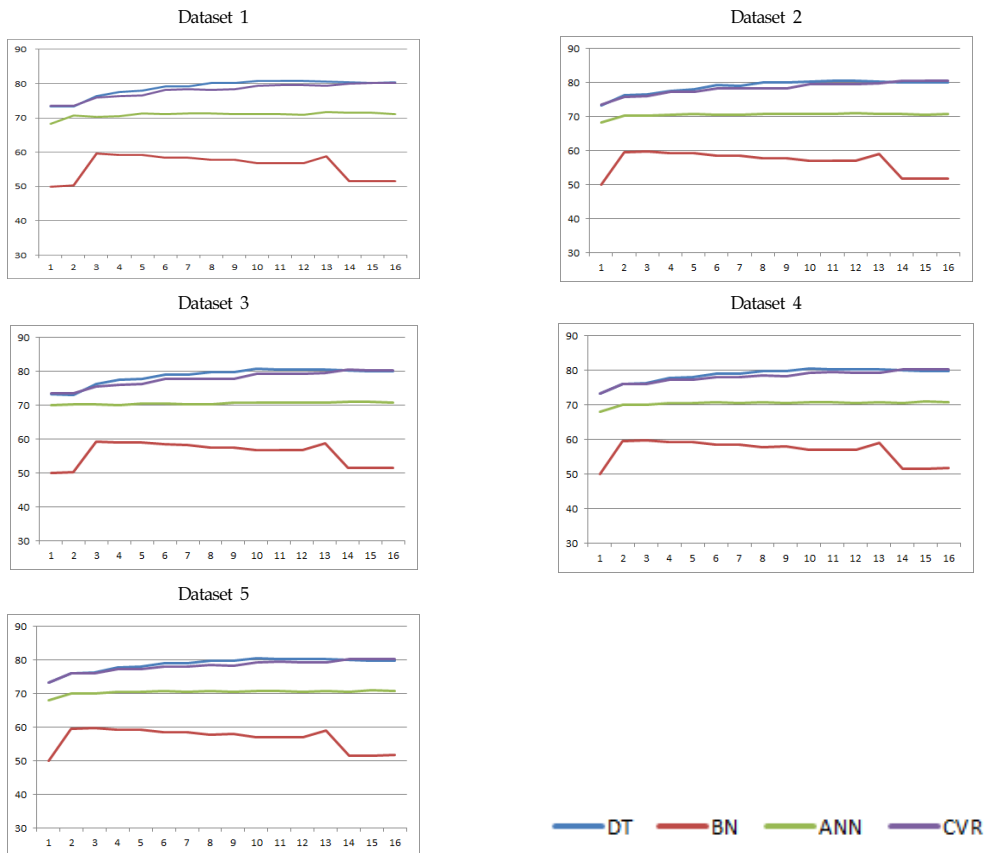
3.2 모델 구축

본 연구에서는 최적의 모델을 구축하기 위해 역행적 제거 래퍼(backward elimination wrapper) 방식을 이용하여 목표변수에 대한 영향도가 적은 변수부터 제거해가면서

모델을 구축하였다. 영향도 평가는 카이제곱 검정 결과를 이용하였으며 검정 결과, 항공기 운영시간, 항공기 소티수, '14년 소모개수 등의 변수가 목표변수에 영향도가 큰 것으로 나타났다. 모델은 DT(Decision Tree(J48)), BN(Bayesian Network), ANN(Artificial Neural Network), CVR(Core Vector Regression)의 네 가지 기법을 사용하였으며 각 모델 구축시 10겹 교차검증(10-folds cross validation)을 하였다. 모델 구축에는 공개소프트웨어인 Weka 3.6.1을 사용하였다.

4. 실험 결과 및 해석

다섯 개 데이터셋에 대한 모델별 예측 정확도는 그림 1에서 보는 바와 같다. x축은 각 모델에 사용된 변수 개수를



(그림 1) 각 모델별/데이터셋별 예측정확도

(Figure 1) Prediction accuracy of each dataset and each model

(표 3) 모델별 예측정확도
(Table 3) Prediction accuracy of each model

변수갯수	DT	BN	ANN	CVR
1	73.340	49.969	68.261	73.528
2	73.245	50.370	70.785	73.453
3	76.265	59.623	70.402	75.913
4	77.615	59.209	70.565	76.347
5	78.035	59.209	71.406	76.516
6	79.165	58.475	71.205	78.186
7	79.084	58.481	71.287	78.349
8	80.213	57.759	71.274	78.192
9	80.182	57.759	71.168	78.311
10	80.804	56.868	71.149	79.454
11	80.703	56.886	71.124	79.592
12	80.753	56.899	70.954	79.498
13	80.540	58.795	71.783	79.385
14	80.370	51.557	71.450	80.044
15	80.283	51.557	71.576	80.107
16	80.396	51.607	71.186	80.107

뜻한다. '10년에서 '14년까지의 데이터를 기반으로 트레이닝하고 '15년 데이터로 테스트하였는데 '15년 수요에 대한 예측 정확도 측면에서 DT가 전반적으로 다른 모델에 비해 우수한 결과를 보여주었다. DT 모델의 경우 각 데이터셋의 예측정확도 최대값을 기준으로 평균 80.6152%의 정확도를 보여주었다. 데이터셋 별 예측정확도의 최대값과 최소값 간 차이는 0.308%였다. 첫 번째 데이터셋의 경우 80.8035%로 예측정확도가 가장 높았으며, (표 3)에서와 같이 이 때는 총 10개의 변수를 사용하였을 때였다. 사용된 변수는 항공기 운영시간, 항공기 소티수, 소모개수('10년~'14년), 수요/청구 비율('12년, '13년, '14년)였으며 수리율과 폐기율, 표준단가, 조달기간 및 수요/청구 비율('10년, '11년)은 제외되었다. 한편, 동일 5개 데이터 셋에 대해 시계열 기법(산술평균법, 이동평균법, 가중이동평균법, 최소자승법, 단순지수평활법, 이중지수평활법, 홀트지수평활법, 윈터지수평활법) 적용시 2번째 데이터 셋의 예측 정확도가 높은 것으로 나타났는데 이 때 정확도는 64.4%였다.

모델 구축시 15,930개의 각 품목에 대한 분류규칙이 생성된다. 그 중 첫 번째 데이터 셋에 대한 DT모델을 예시로 살펴보면 (표 4)와 같다. 예를 들어 15,930개 중 총 4,098개의 품목에 해당되는 규칙은 '14년 소모개수가 2개 이하이며 항공기운영시간이 2,267,063시간 미만이고 '14

(표 4) 분류 규칙
(Table 4) Classification Rule

빈도	규칙	분류결과
4,098회	'14년 소모개수 <= 2 항공기운영시간 <= 2,267,063, '14년 소모개수 <= 0	False
3,369회	'14년 소모개수 > 2 '13년 소모 대 불출 비율 > 0, '11년 소모개수 > 5, '12년 소모개수 > 3	True
1,139회	'14년 소모개수 > 2 '13년 소모 대 불출 비율 > 0, '11년 소모개수 <= 5 항공기운영시간 <= 887,487	True
901회	'14년 소모개수 <= 2 항공기운영시간 <= 2,267,063, '14년 소모개수 > 0, '11년 소모개수 <= 4, '10년 소모개수 <= 2, '12년 소모개수 <= 1	False
726회	'14년 소모개수 <= 2 항공기운영시간 > 2,267,063, '13년 소모개수 <= 0, '14년 소모개수 <= 0	False
...	...	

년 소모개수가 0개인 경우였으며, 이 때 '15년 수요는 발생하지 않는 것으로 분류되었다. 이는 전체 품목 대비 약 26%에 해당된다.

본 논문의 결과를 토대로 추후 항공기 수리부속 수요 예측의 유효 변수를 발견할 수 있었다. 먼저, 최근 '14년 소모개수와 직관적으로만 의미가 있다고 생각되었던 항공기운영시간이 항공기 수리부속의 소모패턴에 영향을 주는 변수임을 확인할 수 있다는 결과에 의미가 있다. 이러한 실험결과는 최근 수요와 항공기 운영시간은 다음해에도 수요패턴에 영향력이 크다는 것을 의미한다. 또한, 영향력 있는 변수인 최신 소모개수, 항공기운용시간 등을 반대로 활용하여 수리부속 수요예측도 가능하게 된다. 예를 들어 위에서 언급한 항공기운영실적을 활용한 운영시간 및 소티수에 비례한 항공기 수리부속 소요산출 방법론 개발도 가능할 것으로 판단된다.

수리부속 수요예측 정확도 제고 측면으로는 시계열 기법의 64.4% 수요예측 정확도보다 DT(Decision Tree)와 CVR(Core Vector Regression)가 좋은 정확도를 보였으며, 데이터마이닝 기법에서는 DT가 가장 높은 정확도를 보였다. 본 결과와 같이 시계열 기법을 벗어나 다양한 기법을 통해 정확도 개선의 노력이 요구된다. 이는 십여 년간 군에서 사용해 오던 수요예측 기법을 본 연구 방법론 결과를 토대로 민간 수요예측 기법을 벤치마킹하여 적용의 가치가 있다고 판단된다.

6. 결 론

본 연구에서는 공군 항공기의 수리부속 수요예측 정확도 제고를 위한 지난 5개년의 수리부속 수요데이터를 분석하고 데이터마이닝 기법을 활용하여 수리부속 수요예측 모델을 제안하였다. 이 과정에서 수요예측 정확도 제고에 연관성이 높은 변수들을 식별할 수 있었으며, 제안한 모델을 통한 실험결과는 기존의 수요예측 방식들에 비해 향상된 정확도를 나타내었다. 추후 장비 노후화에 따른 소모패턴의 변화를 고려하여 지속적인 모델의 업데이트가 요구된다. 아울러, 수리예측 정확도 제고를 위해서는 원천데이터의 신뢰성이 중요하므로, 현재 사용되고 있는 수리부속 청구실적 자료를 실제 소모된 소모실적 자료로 대체할 필요가 있다.

끝으로 본 연구에서는 수리부속 소모이력과 항공기운 영시간 등의 데이터를 분석에 활용하였으나, 향후 기상 정보 등의 외부 데이터나 정비기록 등의 비정형화된 텍스트 데이터를 추가적으로 활용한다면 본 연구에서 제시한 모델의 정확도 이상의 수요예측 결과를 얻을 수 있을 것으로 기대되며, 이에 대한 연구는 추후 연구로 남겨둔다.

참 고 문 헌(Reference)

- [1] Kim. Kui-Won, "Forecasting and research for Bigdata Analysis", Joeuddang, pp. 92-137, 2014. http://book.naver.com/bookdb/book_detail.nhn?bid=8143296
- [2] Yu. Hyeong-Gon, "Improving the management of spare parts", Security Management Institute, pp. 1-37, 2011. http://policy.nl.go.kr/search/searchDetail.do?rec_key=SH1_UMO20150143072&kwd=%EA%B5%AD%EB%B0%A9%EB%B6%80
- [3] Létourneau, S., Famili, F., Matwin, S., "Data mining for prediction of aircraft component replacement", IEEE Intelligent Systems Journal, pp. 59-66, 1999. <https://doi.org/10.1109/5254.809569>
- [4] Park. Young-Jin, Jeon. Geon-Wook, "A Demand Forecasting for Aircraft Spare using ARMA", Journal of the Military Operations Research Society of Korea, 34, pp. 79-101, 2008. <http://www.dbpia.co.kr/Journal/ArticleDetail/NODE01055240>
- [5] Nam. Bong Woo, Song. Kyung Bin, Kim. Kyu Ho, Cha. Jun Min, "The Spatial Electric Load Forecasting Algorithm using the Multiple Regression Analysis Method", The Korean Institute of Illuminating and Electrical Installation Engineers, pp. 63-70, 2008. <http://www.dbpia.co.kr/Journal/ArticleDetail/NODE00959324>
- [6] David F. Findely, Brian C. Monsell, William R. Bell, Mark C. Otto and Bor-Chung Chen, "New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program", Journal of Business and Econometric Statistics, vol. 16, 1988. <https://doi.org/10.2307/1392566>
- [7] Choi. Jae-Il, Jeong. Young-Rak, "Predication of the Number of the Korea Professional Soccer Spectators Using Time-series Data Analysis", Journal of Sport and Leisure Studies, 39, pp. 921-928, 2010. http://www.riss.kr/search/detail/DetailView.do?p_mat_type=1a0202e37d52c72d&control_no=ba5c357985d5b1386aae8a972f9116fb
- [8] Chae. Hyun Seok, Chung. Jea Woo, Kim. Jong Dall, "Demand Forecast for Electric Vehicles in Korea Using a Bass Diffusion Model", Korea Environmental Policy and Administration Society, pp. 109-132, 2016. <http://www.dbpia.co.kr/Journal/ArticleDetail/NODE06650762>
- [9] Lee. Jong Sung, "Delphi Method", Kyoyookbook, 2001. http://used.kyobobook.co.kr/product/viewBookDetail.ink?cmdtBrod=7255760563484&orderClick=LIP&Kc=SEB_LBkusedsearch
- [10] Ashton, A. H., Ashton, R. H. & Davis, M. N., "White-collar robotics: Levering managerial decision making", California Management Review, 37, pp. 83-109, 1994. <https://doi.org/10.2307/41165779>
- [11] Armstrong, J. S. & Graefe, A., "Predicting elections from biographical information about candidates: a test of the index method", Journal of Business Research, 64(7), pp. 699-706, 2011. <http://dx.doi.org/10.1016/j.jbusres.2010.08.005>
- [12] James M. Lyneis, "System dynamics for market forecasting and structural analysis", System Dynamics Review, 16, pp. 3-25, 2000. [http://dx.doi.org/10.1002/\(sici\)1099-1727\(200021\)16:1<3::aid-sdr183>3.0.co;2-5](http://dx.doi.org/10.1002/(sici)1099-1727(200021)16:1<3::aid-sdr183>3.0.co;2-5)
- [13] Lam, W. H. K., and Pan L. P., and Eddie C. M, "Forecasts

- and Reliability Analysis of Port Cargo throughput in Hong Kong”, *Journal of Urban Planning and Development*, 2004. [http://dx.doi.org/10.1061/\(asce\)0733-9488\(2004\)130:3\(133\)](http://dx.doi.org/10.1061/(asce)0733-9488(2004)130:3(133))
- [14] Kong, D-S, Kwak, Y-H, and Huh, J-H, “Artificial Neural Network Based Energy Demand Prediction for the Urban District Energy Planning”, *Journal of the Architectural Institute of Korea*, 26(2), pp. 221-230, 2010. <http://www.dbpia.co.kr/Journal/ArticleDetail/NODE01366386>
- [15] Lee, Yeon-Wook, “The importance of operational spare parts”, *Defense Daily*, 2014.
- [16] Sun, Mee-Sun, Woo, Jae-Wong, “A Study on Forecasting of Repair Part Demands of Korean Military: focused on Navy”, *The Quarterly Journal of Defense Policy Studies*, pp. 201-234, 2009. http://kiss.kstudy.com/journal/thesis_name.asp?tname=ki ss2002&key=3108692
- [17] Kim, H-S, Kim, P-S, “Material Requirements Planning for Military Maintenance Depot”, *Journal of society of Korea industrial systems engineering*, 37, pp. 24-34, 2014. <http://scholar.ndsl.kr/schDetail.do?cn=JAKO201406566439139>
- [18] Moon, Seong-Min, “The Impact of Demand Features on the Performance of Hierarchical Forecasting: Case Study for Spare parts in the Navy”, *Korean management science review*, 29(1), pp. 101-114, 2012. <http://scholar.ndsl.kr/schDetail.do?cn=JAKO201215053140715>
- [19] Michael J. A. Berry, Gordon S. Linoff, “Data Mining Techniques: For Marketing, Sales and Customer Relationship Managements”, John Wiley&Sons, pp.10-15, 2004. <http://as.wiley.com/WileyCDA/WileyTitle/productCd-0470650931.html>
- [20] Clifton, Christopher: “Definition of Data Mining”, *Encyclopedia Britannica*, 2010. <https://global.britannica.com/technology/data-mining>
- [21] Richard Roiger, Michael Geatz, “Data Mining A Tutorial-Based Primer”, Pearson, 2003. <https://www.amazon.com/Data-Mining-Tutorial-Based-Primer/dp/0201741288>
- [22] Lee, Kyu-Youg, Seo, Jun-Yong, “A Case Study on the Inventory Management Using the Datamining”, *Journal of the Society of Korea Industrial and Systems Engineering*, 30(3), pp. 20-27, 2007. <http://scholar.ndsl.kr/schDetail.do?cn=JAKO200735836611625>
- [23] Choi, Gee-Seon, Shin, Gang-Wook, Lim, Sang-Heui, Chun, Myung-Geun, “Short-term Water Demand Forecasting Algorithm Based on Kalman Filtering with Data Mining”, *Journal of institute of control robotics and systems*, 15(10), pp. 1056-1061, 2009. <http://scholar.ndsl.kr/schDetail.do?cn=JAKO200935535843994>
- [24] Shin, Gang-Wook, Hong, Sung-Taek, “Data mining analysis for short-term water demand forecasting”, *Electric power systems research*, pp. 14-17, 2009. <http://scholar.ndsl.kr/schDetail.do?cn=NPAP08451473>
- [25] Lee, Sun-a, Chang, Nam-sik, “Detection of Phantom Transaction using Data Mining: The Case of Agricultural Product Wholesale Market”, *Journal of intelligence and information systems*, 16, pp. 3-25, 2000. <http://scholar.ndsl.kr/schDetail.do?cn=JAKO201510534324147>

◎ 저 자 소 개 ◎



김 재 동(Jaedong Kim)

2005년 한양대학교 산업공학과(공학사)
2004년 한양대학교 대학원 산업공학과(공학석사)
2007년~현재 한국국방연구원 국방운영연구센터 선임연구원
관심분야 : 데이터 마이닝, 수요예측, 재고관리, 시뮬레이션 등
E-mail : soobahkin@gmail.com



이 한 준(Hanjun Lee)

2001년 서울대학교 컴퓨터공학과(공학사)
2004년 서울대학교 대학원 컴퓨터공학과(공학석사)
2004년~2007년 육군사관학교 전자정보학과 전임강사
2016년 고려대학교 경영학과(경영학 박사)
2007년~현재 한국국방연구원 국방획득연구센터 선임연구원
관심분야 : 데이터 마이닝, 빅데이터, 소셜미디어, 정보화 평가, 정보화 정책 등
E-mail : hjlee1609@gmail.com