

군사과학기술정책연구

Military Science & Technology Annual Report

연구논문

- 인공지능 기반 장교 및 부사관 진급 모델 연구
-제도 검증 및 다면평가 유효성을 중심으로-
: 장용진, 이재홍, 박경욱, 신진수, 양경민, 조윤환, 이현석



국 방 대 학 교
국가안전보장문제연구소

ISSN 1976-5967

제17권

2024년 12월

군사과학기술정책연구

Military Science & Technology Annual Report

국방대학교 국가안전보장문제연구소

발 간 사

국방부에서 추진하고 있는 「국방혁신 4.0」의 중요한 배경은 병력자원감소의 환경 하에서 AI(Artificial Intelligence, 인공지능) 같은 과학기술을 군 전력화와 국방 운영에 활용하자는 것입니다. 특히 병력자원감소는 이러한 국방혁신을 하지 않을 수 없게 하는 핵심적인 요소입니다. 국가 출생률이 낮아져서 병력에 활용한 가용한 자원이 제한되기에 국방혁신을 하지 않을 수 없는 상황입니다.

국방 병력자원 감소와 함께 병력자원 운영면에서 관심을 가져야 할 분야는 군의 허리에 해당되는 중간 간부들의 군 이탈입니다. 근무 환경, 직업 안정성, 병사의 대우에 비한 상대적인 박탈감 등 여러 요인으로 인해서 중간 간부들의 군 이탈이 증가하고 있는 상황입니다. 이러한 문제들에 대한 해결책을 찾고자 다양한 연구가 국방부, 합참, 각 군 그리고 군 관련 기관을 통해서 추진되고 있습니다.

국방대학교 국가안전보장문제연구소도 이러한 국방현안에 대응하기 위해 2024년도 연구주제로 “인공지능 기반 장교 및 부사관 진급 모델 연구”를 선정하여 연구하고 발간하게 되었습니다. 공정하고 합리적인 진급 과정은 장교 및 부사관의 미래에 대한 희망을 주게 되고, 군 중견간부에 대한 직업 안정성 면이나 병력 자원의 안정적 확보에도 기여하게 될 것입니다. 그러한 측면에서 인공지능 기반의 장교 및 부사관 진급 모델 연구가 국방정책상 의미를 가진 시의적절한 연구라 할 수 있습니다.

모쪼록 이번 연구 결과가 국가안보 및 국방정책 관련 부처의 정책개발 및 집행과정에 기여하고, 이 분야를 연구하는 연구자 및 학생들에게 유용한 참고자료로 활용될 수 있기를 기대합니다. 끝으로 제한된 기간에도 불구하고 국가안보를 위한 연구에 열과 성을 다해주신 연구자 여러분께 깊은 감사의 말씀을 전합니다.

2024년 12월 31일

국방대학교 국가안전보장문제연구소장 교수 박 영 준

인공지능 기반 장교 및 부사관 진급 모델 연구

-제도 검증 및 다면평가 유효성을 중심으로-

장용진, 이재홍, 박경욱, 신진수, 양경민, 조윤환, 이현석

2024. 12.



국방대학교 국가안전보장문제연구소

목 차

요약문	11
I. 서론	14
1. 연구 배경 및 필요성	14
1.1. 연구 배경	14
1.2. 진급의 역할	17
1.3. 진급과 성과 평가	18
2. 군조직에서 진급 제도	20
3. 연구 범위와 방법	21
II. 이론적 배경	22
1. 인적자원 관리와 진급	22
1.1. 진급의 역할	22
1.2. 군에서 진급의 역할	23
2. 군 인적자원 관리 연구	24
2.1. 군조직 인사관리 선행연구	24
2.2. 국방혁신 4.0과 AI 인적자원 관리	24
2.3. 진급 모델 연구에서의 머신러닝	25
2.4. 다면 평가 연구	25
3. 주요 머신러닝 알고리즘 특성	27
3.1. 베이지안 네트워크	27
3.2. 나이브 베이즈 모델	29
3.3. Gaussian Process Regression (가우시안 과정 회귀)	30
3.4. Support Vector Machine (서포트 벡터 머신)	31
3.5. Decision Tree (의사결정나무)	33
3.6. Ensemble Methods (앙상블 기법)	34
3.7. Linear Regression (선형 회귀)	35
3.8. Neural Networks (신경망)	36
4. 기댓값 최대화 알고리즘	37
5. 데이터 이산화	39
6. 연구 범위 및 방법	39

III. 부사관 진급 모델 연구	41
1. 부사관 진급제도 현황	41
2. 공군 부사관 기술통계	41
3. 공군 부사관 데이터 상관관계 분석	43
4. 공군 부사관 머신러닝 모델 비교	47
5. 기본 모델링	55
6. 민감도 분석결과	56
7. 다면평가 추가모델 타당성 분석결과	58
7.1. 다면평가 추가모델 타당성 분석	58
7.2. 이산화계획의 효과	60
7.3. 진급 결과에 따른 CPT 교정 결과	67
8. 소결론	73
IV. 장교 진급 모델 연구	75
1. 공군 장교 진급제도	75
2. 공군장교 데이터 기술통계	75
3. 공군장교 데이터 상관관계 분석	78
4. 장교 머신러닝 모델 비교	81
5. 기본 모델 모델링	86
6. 다면평가 추가 모델 모델링	87
7. 민감도 분석결과	87
8. 교정 및 타당성 분석	89
8.1. 다면평가 추가모델 타당성 분석	89
8.2. 이산화 계획의 효과	92
8.3. 진급 결과에 따른 CPT 교정 결과	95
9. 향후 발전 방향	100
9.1. 베이지안 의사결정 개념	100
9.2. 수치 예시(Numerical Example)	102
10. 소결론	104
VII. 결론	105
참고문헌	108
부 록	110

그림목차

〈그림 1-1〉 인적자원 관리 유형	16
〈그림 1-2〉 정부 업무 평가제도	19
〈그림 2-1〉 기본 베이지안 네트워크 모형	27
〈그림 2-2〉 마코프 블랭킷(Markov Blanket) 모델 예시	28
〈그림 2-3〉 가우시안 과정 회귀	30
〈그림 2-4〉 Architecture of Support Vector Machine	32
〈그림 2-5〉 의사결정나무 모형	33
〈그림 2-6〉 신경망 모형	36
〈그림 2-7〉 기댓값 최대화 알고리즘 흐름도	38
〈그림 2-8〉 연구 순서도	40
〈그림 3-1〉 주요 평가항목 상자도표	43
〈그림 3-2〉 산점도 매트릭스	44
〈그림 3-3〉 클러스터링 덴드로그램	46
〈그림 3-4〉 ROC 곡선 문턱값(Threshold) 개념도	48
〈그림 3-5〉 ROC 곡선	48
〈그림 3-6〉 부사관 진급 모델	55
〈그림 3-7〉 부사관 진급 모델 수정(다면평가 추가)	56
〈그림 3-8〉 민감도 분석	56
〈그림 3-9〉 2019년도 ROC	59
〈그림 3-10〉 2020년도 ROC	59
〈그림 3-11〉 2021년도 ROC	59
〈그림 3-12〉 2022년도 ROC	59
〈그림 3-13〉 2023년도 ROC	60
〈그림 3-14〉 통합년도 ROC	60
〈그림 3-15〉 다면평가 노드 진급 모델 Bar chart (3수준)	60
〈그림 3-16〉 이산화(6, 9, 12, 15수준)별 진급 모델 Bar chart	61
〈그림 3-17〉 연도별 다면평가 수준별 분포 (진급자)	62

〈그림 3-18〉 연도별 다면평가 수준별 분포 (미진급자)	62
〈그림 3-19〉 2019년도 ROC	66
〈그림 3-20〉 2020년도 ROC	66
〈그림 3-21〉 2021년도 ROC	66
〈그림 3-22〉 2022년도 ROC	66
〈그림 3-23〉 2023년도 ROC	66
〈그림 3-24〉 통합 ROC	66
〈그림 3-25〉 미진급(Z=0)인 경우의 Bar chart	67
〈그림 3-26〉 진급(Z=1)인 경우의 Bar chart	68
〈그림 3-27〉 통합연도 진급자 bar 그래프 (전체 데이터)	69
〈그림 3-28〉 통합연도 진급자 bar 그래프 (다면실시 데이터)	70
〈그림 3-29〉 2019년 진급자 bar 그래프 (다면실시 데이터)	72
〈그림 3-30〉 2020년 진급자 bar 그래프 (다면실시 데이터)	72
〈그림 3-31〉 2021년 진급자 bar 그래프 (다면실시 데이터)	72
〈그림 3-32〉 2022년 진급자 bar 그래프 (다면실시 데이터)	72
〈그림 3-33〉 2023년 진급자 bar 그래프 (다면실시 데이터)	72
〈그림 3-34〉 통합연도 진급자 bar 그래프 (다면실시 데이터)	72
〈그림 4-1〉 주요 평가항목 상자도표	77
〈그림 4-2〉 주요 평가항목 산점도	78
〈그림 4-3〉 클러스터 덴드로그램	80
〈그림 4-4〉 장교 기본 모델	86
〈그림 4-5〉 장교 다면평가 추가 모델	87
〈그림 4-6〉 민감도 분석결과	88
〈그림 4-7〉 19년도 ROC	90
〈그림 4-8〉 20년도 ROC	90
〈그림 4-9〉 21년도 ROC	91
〈그림 4-10〉 22년도 ROC	91
〈그림 4-11〉 23년도 ROC	91
〈그림 4-12〉 통합 ROC	91

〈그림 4-13〉 19년도 ROC	94
〈그림 4-14〉 20년도 ROC	94
〈그림 4-15〉 21년도 ROC	94
〈그림 4-16〉 22년도 ROC	94
〈그림 4-17〉 23년도 ROC	94
〈그림 4-18〉 통합 ROC	94
〈그림 4-19〉 미진급($Z=0$)인 경우의 Bar chart	95
〈그림 4-20〉 진급($Z=1$)인 경우의 Bar chart	95
〈그림 4-21〉 통합연도 진급자 bar 그래프 (전체 데이터)	96
〈그림 4-22〉 통합연도 진급자 bar 그래프 (다면 실시 데이터)	97
〈그림 4-23〉 2019년 진급자 bar 그래프 (다면 실시 데이터)	99
〈그림 4-24〉 2020 진급자 bar 그래프 (다면 실시 데이터)	99
〈그림 4-25〉 2021년 진급자 bar 그래프 (다면 실시 데이터)	99
〈그림 4-26〉 2022년 진급자 bar 그래프 (다면 실시 데이터)	99
〈그림 4-27〉 2023년 진급자 bar 그래프 (다면 실시 데이터)	99
〈그림 4-28〉 통합연도 진급자 bar 그래프 (다면 실시 데이터)	99
〈그림 4-29〉 베이저안 의사결정 모델 예시	101
〈그림 4-30〉 장교 진급 베이저안 의사결정 모델	102
〈그림 4-31〉 진급확률 vs. 기대 비용	103

표 목 차

〈표 3-1〉 부사관 진급 항목 기술통계	42
〈표 3-2〉 상관관계 분석결과	45
〈표 3-3〉 2019년 모델별 AUC	50
〈표 3-4〉 2019년 모델별 AUC	51
〈표 3-5〉 2021년 모델별 AUC	52
〈표 3-6〉 2022년 모델별 AUC	53
〈표 3-7〉 2023년 모델별 AUC	54
〈표 3-8〉 연도별 민감도 분석결과	57
〈표 3-9〉 부사관 진급 예측도 비교	58
〈표 3-10〉 연도별 정확도 비교	59
〈표 3-11〉 연도별 다면평가 이산화 수준에 따른 모델 정확도	61
〈표 3-12〉 연도별 진급대상자 중 다면평가 현황	63
〈표 3-13〉 연도별 다면평가 실시자 중 진급 현황	63
〈표 3-14〉 다면평가 실시 데이터 학습 후 정확도	64
〈표 3-15〉 전체 및 일부(다면평가 실시 데이터) 반영 후 정확도 비교	65
〈표 3-16〉 통합년도 진급 유무에 따른 CPT (전체 데이터)	68
〈표 3-17〉 통합년도 진급 유무에 따른 CPT (다면평가 실시 데이터)	69
〈표 3-18〉 5개년도 진급 유무에 따른 CPT (다면평가 실시 데이터)	71
〈표 4-1〉 장교 진급 항목 기술통계	76
〈표 4-2〉 상관관계 분석결과	79
〈표 4-3〉 2019년 모델별 AUC	81
〈표 4-4〉 2020년 모델별 AUC	82
〈표 4-5〉 2021년 모델별 AUC	83
〈표 4-6〉 2022년 모델별 AUC	84
〈표 4-7〉 2023년 모델별 AUC	85
〈표 4-8〉 연도별 민감도 분석결과	88
〈표 4-9〉 진급 예측 정확도 비교	89

〈표 4-10〉 연도별 정확도 비교	90
〈표 4-11〉 연도별 다면평가 이산화 수준에 따른 모델 정확도	92
〈표 4-12〉 다면평가 실시 데이터 학습 후 정확도	92
〈표 4-13〉 전체 및 일부(다면평가 실시 데이터) 반영 후 정확도 비교	93
〈표 4-14〉 통합년도 진급 유무에 따른 CPT (전체 데이터)	96
〈표 4-15〉 통합년도 진급 유무에 따른 CPT (다면평가 실시 데이터)	97
〈표 4-16〉 5개년도 진급 유무에 따른 CPT (다면평가 실시 데이터)	98

요 약 문

우크라이나-러시아 전쟁은 그동안의 전쟁과는 명확히 구분되는 다양한 양상을 보여주었다. 인공지능(AI)과 전투용 드론이 본격적으로 활용되었으며, 새로운 형태의 전쟁을 시사하며 전쟁 양상의 변곡점을 보여주었다. 전쟁뿐만 아니라 사회 전반에서 AI 경쟁이 심화되고 있으며, 실생활에 AI가 적용되는 모습을 체감할 수 있다.

새로운 정부 출범에 발맞추어 진행 중인 국방혁신 4.0 또한 이러한 방향성을 잘 보여주고 있다. 첨단 과학 무기체계는 물론 군 전 영역에서 새로운 기술을 접목하려는 노력이 경주되고 있다. 국방혁신 4.0에서도 제시된 유/무인 복합 전투 체계는 새로운 형태의 전쟁 방식을 강요하고 있으며, 이를 운영하는 데 적합한 조직 형태의 변화를 야기하고 있다. 이처럼 한 분야가 아니라 모든 분야에서 변화와 혁신이 필요한 시기이다.

조직과 인사를 다루는 인적자원 관리도 예외는 아니다. 기업의 경우 인재 선발과 배치를 할 때 AI를 활용한 시스템을 활용하는 사례가 확대되는 추세이다. 최종적인 판단은 사람의 몫이지만, 새로운 시스템은 다양한 데이터를 기반으로 의사결정 과정을 지원하는 역할을 하므로 효용도가 높다. 군 조직의 인적관리 체계, 그중에서도 진급에 대한 부분은 많은 사람이 많은 관심을 두고 있는 분야이며, 고위 계급의 선발 과정이 적절하지 않았을 때 조직에 큰 영향을 미치므로 중요한 역할을 한다고 말할 수 있다.

그동안 군의 진급체계에 관한 관심에 비해 충분한 연구들이 이루어지지 않았으며, 기존의 연구에서도 문헌 연구 중심의 질적 연구가 주를 이루고 있으므로 균형을 맞추기 위하여 정량적인 연구를 통하여 그동안의 문제점을 파악하고 바람직한 방향성을 잡기 위한 노력이 필요한 시기이다.

본 연구의 목적은 이처럼 군 간부의 진급에 관하여 정량적인 연구를 통하여 새로운 시사점을 발견하는 것이다. 아울러, 진급 대상자 선발 시 인공지능을 기반으로 하는 접근법을 제시하여, 향후 군 조직에 기여할 수 있는 적합한 인재 선발에 이바지하고자 한다. 이를 위하여 공군 장교, 부사관 진급에 활용되었던 진급평가 데이터를 기반으로, 현재 시스템 점검을 하고 개선점을 제시하고자 한다.

연구 방법에 있어서, 먼저 장교와 부사관을 분리하여 2개의 연구를 진행하였다. 군 간부의 진급관리 제도는 장교와 부사관을 구분하고 있으며, 조직 내에서 역할

이 구분되는 만큼 적합한 인재상에도 차이가 있다. 따라서 장교의 진급 모델과 부사관의 진급 모델을 구분하여 연구하였으며, 두 모델을 비교하여 종합적인 시사점을 제시하였다.

분석에 사용된 도구 선정을 위하여 여러 모델을 비교하고, 베이지안 네트워크를 선택하였다. 일반적으로 데이터를 분석하는 데 가장 널리 사용되는 방법은 선형회귀모델이다. 선형회귀 모델의 경우 전제로 하는 조건을 만족하는 경우, 가장 쉽게 이해할 수 있고 비 편향성이 있다는 장점이 있으므로 데이터 분석에 많이 사용된다. 하지만, 선형회귀 모델은 종속변수의 비선형성이 있는 경우 정확한 추정이 어렵다는 한계점을 갖기 때문에 최근 연구에서는 가우시안 분포를 활용한 확률 모형과 머신러닝 등의 방법론에 대한 관심이 증가하고 있는 추세이다.

그중에서 베이지안 네트워크는 토마스 베이즈(Thomas Bayes)가 제시한 베이지안 정리를 기반으로 한다. 특히 종속변수에 영향을 주는 모든 변수를 확인할 수 없을 때, 불확실성이 존재할 때 널리 활용되는 머신러닝 방법이다. 진급심사의 결과는 합격과 불합격으로만 나누어지므로, 회귀분석에서 가정하는 선형성에 대한 오차가 크게 나는 상황이다. 또한, 진급평가 과정이 자격 점수, 부대 추천 외에서 추천 심사 선발 심사를 거쳐 최종 인사 결정권자의 재가까지 많은 과정을 거치게 되므로 복잡한 과정에서 발생하는 불확실성을 반영하는데 좋은 방법이라고 할 수 있다.

연구 결과 장교와 부사관 모두의 연구에서 배점으로 나타나는 명목상의 중요도와 실제 진급에 영향을 미치는 중요도를 비교하였을 때 부합하지 않는 부분이 있다는 점이 발견되었다. 이는 진급제도를 설계한 의도와 실제 적용 간의 차이가 있다는 것을 의미한다. 제도상 중요도를 높게 설정한 항목 중 실제 영향이 적은 항목들에 대하여 재점검이 필요하다고 할 수 있다. 다시 말하면, 평가하는 방식이 너무 느슨하여 실제 진급 대상자들의 점수 변별력이 떨어진다면 배점이 높더라도 실제 진급에 미치는 영향력은 크게 떨어지기 때문에 평가방법을 재설계하거나, 영향력을 높이기 위해 배점을 더 확대하여 미치는 영향력을 더 강화하는 방향을 고려해야 한다.

다면평가는 평가 인원이 증가할 뿐만 아니라, 평가 방향이 다각화되어, 객관성이 증가한다는 장점이 있는데, 장교의 경우 다면평가를 평가항목에 포함하였을 때 진급 예측도가 높게 나타났다. 부사관의 경우에는 진급 대상자 중 다면평가를 실시하지 않은 인원이 대다수였는데, 다면평가를 실시한 인원들의 데이터를 학습하여 검증한 결과, 매년 진급 예측을 더 잘하는 것으로 나왔다. 하지만 의미 있는 향상을 보여주지는 않았기 때문에 충분한 다면평가가 이루어질 때 점진적으로 도입하

는 것이 더 나은 방법이다. 따라서, 장교의 경우 다면평가를 활용하는 방안을 긍정적으로 검토할 필요가 있다고, 부사관도 점진적인 고려대상이 될 수 있다고 할 수 있다.

I. 서론

1. 연구 배경 및 필요성

1.1. 연구 배경

인적자원은 지식사회에서 더욱 중요해지고 있다. 인적자원 관리란 선발에서부터 해임하는 과정까지 조직에서 사람을 관리하는 모든 과정을 말한다. 인적자원 관리 중에서도 진급은 다음 계급에서 요구되는 우수 인력을 선발한다는 중요한 역할을 한다. 본 연구에서는 군의 진급을 대상으로 새로운 접근을 제시하고자 한다.

어느 조직이나 인재가 부족하다고 한다. 특히 전문화가 중요한 분야일수록 원하는 인재의 수는 더욱 줄어든다. 더욱이 내가 속한 조직에 적합한지 알아내는 것 또한 매우 어려운 일이다.

인적자원 관리 연구에서 조직에 필요한 인재가 부족한 이유는 경험의 뒷, 사람에 대한 무관심, 제도에 대한 집착으로 지목되고 있다.¹⁾ 첫째, 경험의 뒷은 평가자들의 과거의 성공 기준으로 판단하는 것을 말한다. 현재와 과거의 상황이 다름에도 불구하고 현재에도 과거의 기준을 사용하는 것이다. 둘째, 사람에 대한 무관심이다. 인적자원 관리를 실행하는 주체는 관리자이다. 관리자의 부서원에 대한 관심이 조직의 성공에 기초가 된다. 하지만, 인사팀의 업무로만 치부하고 관심을 두지 않거나 인사제도의 취지에 맞지 않는 부서원 관리가 되는 경우이다. 셋째, 제도에 대한 집착은 좋은 인재 우리 조직에 맞는 인재상에 대한 공감대가 부족할 때 발생한다. 인재상이 없으므로, 제도에 상충되는 가치들이 있을 때 규정, 제도에 매달리게 되는 것이다. 모든 제도는 의도가 숨어있다. 인사제도는 우리가 원하는 방향으로 인적자원을 관리하기 위한 것인데, 중심이 되는 사람에 대하여 간과하고 제도 자체에만 집착하면 방향성을 잃는 현상이 발생하고, 좋은 인재가 조직을 떠나는 결과를 초래한다.

조직에 필요한 인재는 조직의 방향성에 따라 달라져야 한다. 1980년대 이후 기업이 경쟁이 심화되면서 전략이 중요한 화두로 떠올랐다. 이로 인해 조직의 전략에 부합하는 인재를 우선으로 하는 전략적 인사관리가 등장하게 된다. 전략적 인사관리에서는 변화하는 환경에 대응하기 위한 전략이 선행되어야 하며 이에 맞는

1) 강성춘. (2020) 인사이드 아웃, 21세기 북스

인사관리를 강조한다.²⁾ 전략은 도달하기 위한 수단과 방법이 설정되어야 한다. 실행을 위한 시스템에 적합한 인재가 선발되고, 시스템적인 관리가 이루어져야 한다. 아울러 적절한 인적자원 관리 시스템은 수직적 및 수평적 적합성이 확보되어야 한다.³⁾

수직적 적합성은 사람 관리를 통해 달성하고자 하는 행동이다. 인사관리 시스템을 통하여 사람이 조직이 원하는 행동을 하도록 할 때 수직적 적합성이 확보되었다고 볼 수 있다. 기업의 경우 원가우위 전략과 차별화 전략을 사용하는 기업의 방향성에 차이가 있다.⁴⁾ 비용을 최소화하려는 기업은 모든 행동의 비용을 따지고 예상 가능하며 결과를 중시하는 행동이 요구된다. 하나하나의 프로세스에 대한 비용 최적화가 중요하기 때문이다. 하지만, 차별화 전략은 다른 경쟁자와 차별화를 위하여 창의성을 중시하고 위험을 감수하는 행동이 더 많이 나타난다. 이러한 전략의 차이는 인사관리에서도 연결이 되며, 비용우위 전략을 추구하는 기업에서 인사정책은 직원들이 경험적이고, 결과를 낼 수 있는 행동을 하도록 장려한다.⁵⁾

수평적 적합성은 가치에 대한 일관적인 방향성이다. 표면적으로는 조직의 단합, 공동체 의식을 강조하면서 인사사고과에서 개별적인 성과를 강조하는 것은 인사관리의 수평적 적합성이 맞지 않는 것이다. 즉, 직원들이 조직에서 중요시하는 일관적인 인재상, 행동을 동일하게 인식하도록 하는 것이 수평적 적합성에서 추구하는 바이다.

군에서도 전략적인 인사관리가 이루어지려면 군의 전략이 무엇인지, 확고한 인재상이 규정될 필요가 있다. 그리고 그를 바탕으로 인사관리 제도들이 정비되어야 한다. 이때 수직적이고 수평적인 적합성 확보가 필요하다. 따라서 본 연구의 대상인 진급제도를 바라볼 때에도 진급 하나만을 볼 것이 아니라, 교육 과정의 설계나, 평가제도, 군에서 지향하는 핵심 역량들과 통합적인 관점에서 접근해야 수평적인 적합성이 확보될 수 있다.

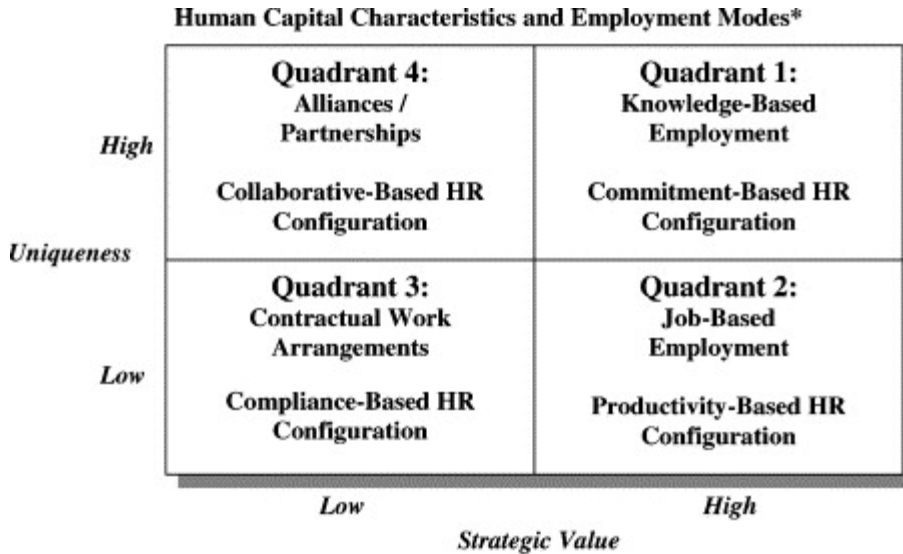
진급 제도가 잘 설계되었는지 수정할 필요가 있는지를 판단할 때에도 진급제도 자체의 완결성을 고민하기보다는 조직의 전략을 현실화하는 데 적합한가를 판단기준으로 삼아야 한다.

2) Tichy, N. M. (1981). Strategic human resource management, Sloan Management Review, 23, 47-61.

3) 강성춘, 전개서.

4) Porter, M. E. (1997). Competitive strategy. Measuring business excellence, 1(2), 12-17.

5) Schuler, R. S. (1989). Strategic human resource management and industrial relations. Human relations, 42(2), 157-184.



〈그림 1-1〉 인적자원 관리 유형⁶⁾

인사관리에도 조직의 특성에 따라 패러다임이 달라진다. 인사조직 아키텍처에서는 크게 시장기반, 협력적, 순응적, 생산적으로 나누고 있다.⁷⁾ 같은 맥락에서 강성춘은 이를 직무 성과주의, 내부 노동 시장형, 스타형, 몰입형 4가지로 분류하였다.⁸⁾ 그중에서 직무성과 주의는 성과를 중심으로 생각하는 관리 방법이며, 단기적인 성과를 중심으로 생각한다. 두 번째, 내부노동 시장형은 조직의 지식을 축적하고 이를 내재화 한 사람을 중심으로 하는데, 제공하는 제품과 서비스의 신뢰성이나 일관성이 중요한 가치이다. 세 번째 스타형은 소수의 고성능자의 역할이 중요한 조직이다. 네 번째 몰입형은 헌신적인 동료들의 소속의식, 연대의식을 중요하게 생각하는 데 자원봉사 단체 등에서 흔히 발견된다. 이러한 패러다임에 따라 인사관리의 방식도 다른 형태를 보인다.

군은 직무 배태성(Job embeddedness)이 큰 조직이다. 직무 배태성이란, 직무의 특수성이 높아 직원의 능력이 다른 곳에 활용되기 어려운 상황을 말한다.⁹⁾ 군

6) David P. Lepak, Scott A. Snell, (2002) Examining the Human Resource Architecture: The Relationships Among Human Capital, Employment, and Human Resource Configurations, Journal of Management, Volume 28, Issue 4, 517-543.

7) Lepak, D. P., & Snell, S. A. (1999). The human resource architecture: Toward a theory of human capital allocation and development. Academy of management review, 24(1), 31-48.

8) 강성춘, 전개서, 아키텍처 모델에서는 시장기반, 협력적, 순응적, 생산적으로 표현하였으나, 위의 저서에서는 성과주의, 내부노동 시장형, 스타형, 몰입형으로 표현을 바꾸었다.

은 배태성이 높아 다른 직장으로 옮기기 어려운 상황이므로 직장 안정성의 욕구가 높다. 아울러 군 조직에서의 인적자원관리는 조직의 신뢰성을 중시하며 충성심을 요구함과 동시에 개인의 전문성을 중요시한다. 이러한 특성을 기준으로 볼 때, 군은 위의 4개의 인사관리 형태 중에서 내부노동 시장형과 직무 성과주의가 혼합된 형태라고 볼 수 있다. 이 두 가지 유형에서 조직원이 바라는 것은 경제적 보상 그리고 고용안정이다. 군에서 인사관리에 여러 방법이 있으나, 특히 진급체제는 언급된 2개의 보상에 직접적인 관련성이 크기 때문에 군 인사관리 전반에 영향력이 매우 크다고 할 수 있다.

1.2. 진급의 역할

인적자원관리에서 진급의 역할은 다양한 시각에서 해석할 수 있다. 첫째 보상의 수단이다. 조직 내에서 진급은 연봉이 증가하는 경향이 있으므로,¹⁰⁾ 그동안 개인의 성과에 대한 보상의 도구로 활용될 수 있다. 둘째, 책임과 권력이 증가한다.¹¹⁾ 보다 높은 위치로 올라갈수록 리더의 역할이 주어지고, 이는 개인뿐만 아니라 조직 전반에 대한 권한과 책임이 증가하는 경향성을 보인다. 이는 대상자가 조직에 핵심 자원에 접근할 수 있는 기회가 더 많아진다는 의미를 갖는다. 셋째, 선발이다. 조직의 신입사원을 뽑는 것과 마찬가지로 진급 또한 적합한 인재를 선택하는 과정이다. 최근 많은 기업에서 경력직에 대한 수요가 높아짐에 따라, 팀장 이상의 중요 보직자를 외부에서 선발하는 경우가 증가하고 있다. 조직의 입장에서 주요 보직에 대한 적임자를 찾는 범위가 조직 내부뿐 아니라 외부로 확장되고 있다는 것을 의미한다. 따라서 조직 내부의 직원을 진급시키는 것은 외부 인사보다 해당 업무를 더 잘할 수 있는 즉, 자리에 적합한 인재로 선택받았다는 의미이다. 특히, 정규직으로 진급하는 경우와 같이 신분 변화가 발생하면 실질적인 선발과 차이가 없다.

진급이 갖는 여러 역할이 있는 만큼 적합한 인재를 진급시키는 것을 중요한 의미를 갖는다. 먼저, 보상의 역할로 보았을 때 선행연구에서는 개인에게 성취에 대

9) Kiazad, K., Holtom, B. C., Hom, P. W., & Newman, A. (2015). Job embeddedness: a multifoci theoretical extension. *Journal of Applied Psychology*, 100(3), 641-699.

10) Gerhart, B., & Milkovich, G. T. (1989). Salaries, salary growth, and promotions of men and women in a large private firm. In R. T. Michael, H. Hartmann, & B. O'Farrell (Eds.), *Pay equity: Empirical inquiries* (pp. 23-43). Washington, DC: National Academy Press.

11) Moore, D. W., & Trout, B. T. (1978). Military Advancement: The Visibility Theory of Promotion. *The American Political Science Review*, 72(2), 452-468.

한 동기를 부여하고, 조직에 대한 긍정적인 인식을 통하여 이직 의도를 낮추는 경향이 있는 것으로 나타났다. 두번째로, 권력이 증가하게 되면 조직의 방향성을 결정하는 데 영향력이 강해진다. 특히 지휘부의 경우 치명적인 영향을 줄 수 있다. 또한 잘못된 리더십은 유능한 부하의 이직의도를 증가시키고, 자원 활용의 비효율을 초래하여 조직의 성과를 더욱 극심하게 떨어뜨린다.

1.3. 진급과 성과 평가

기업 대상으로 한 연구들을 살펴보면, 진급은 개인의 성과를 기준으로 하는 경우가 많다. 성과를 측정하는 방법은 일반적으로 계량적인 성과나 핵심성과지표(Key Performance Indicator, KPI)에 대한 평가하는 경우가 일반적이다. 예를 들어 영업과 관련한 분야에서는 개인별 판매금액으로 개인의 성과를 평가하거나, 생산부서에서 팀의 생산량 증가분을 성과로 측정하기도 한다. 성과측정의 객관성 확보가 중요하기 때문에 위와 같은 항목이 활용된다.

그 밖에 KPI가 성과측정의 지표로 활용된다. 조직이 달성하고자 하는 명확한 목표점을 기준으로 수준별 부서별로 목표를 달성하기 위해 필요한 항목을 선정하고 선정된 항목의 달성률에 따라 성과를 평가하는 방식이다. KPI의 장점은 구체적인 과업의 제시와 명확성에 있다. 전략적으로 조직의 방향성이 정해진다고 하더라도 하부의 조직에서 구체적으로 어떠한 행동을 해야 하는지 가치판단이 어려운 경우들이 발생하며, 때로는 부서 간 충돌을 일으키기도 한다. KPI는 구체적인 과업과 기준을 제시하여 조직이 같은 방향성을 갖도록 운영하는 데 효과적으로 작동한다. 일부 연구에서 KPI의 단순성으로 인하여 업무의 특성에 따라 부적절한 결과를 초래하는 경우들이 보고되기도 하였으나, 부서에 따라 복잡한 성과 계산을 하는 경우 각 분야의 업무 자율성을 해치고, 성과에 대한 평가를 위한 행정업무가 증가하여 조직의 성과를 오히려 해친다는 연구들도 있어서 많은 조직에서 활용되는 방식이다.

공공조직의 성과측정이 어려운 궁극적인 이유는 조직이 제공하는 가치가 공공성을 갖기 때문이다. 기업의 가치는 수치화된 여러 지표를 통하여 확인할 수 있다. 주식시장에 상장된 기업의 시가총액은 주식시장에서 평가받은 기업의 가치이다. 상장되지 않은 기업의 가치는 매출액과 영업이익으로 가늠할 수 있다. 기업의 궁극적인 목적이 재화, 서비스 제공을 통한 이윤 창출이라면, 1분기 혹은 1년 얻은 이익이나 시장의 점유율을 통하여 확인할 수 있다.

공공조직의 가치, 성과는 공공의 이익을 위한 것이므로 해당 조직이 잘 운영되고 있는지 성과를 확인하는 과정이 일반적인 기업에 비해 상대적으로 어렵다.

평가 항목

평가항목	공기업	준정부기관 (위탁집행형)	준정부기관 (기금관리형)
경영관리	• 1. 경영전략(9)	• 1. 경영전략(9)	• 1. 경영전략(9)
	• 2. 사회적 책임(15)	• 2. 사회적 책임(12)	• 2. 사회적 책임(12)
	• 3. 재무성과관리(20)	• 3. 재무성과관리(13)	• 3. 재무성과관리(18)
	• 4. 조직 및 인적자원관리(4)	• 4. 조직 및 인적자원관리(4)	• 4. 조직 및 인적자원관리(4)
	• 5. 보수 및 복리후생관리(7)	• 5. 보수 및 복리후생관리(7)	• 5. 보수 및 복리후생관리(7)
	경영관리 소계 : 총 55점	경영관리 소계 : 총 45점	경영관리 소계 : 총 50점
주요사업	• 주요사업 계획·활동·성과를 종합평가(45)	• 주요사업 계획·활동·성과를 종합평가(55)	• 주요사업 계획·활동·성과를 종합평가(50)
	주요사업 소계 : 총 45점	주요사업 소계 : 총 55점	주요사업 소계 : 총 50점

〈그림 1-2〉 정부 업무 평가제도¹²⁾

위 그림은 정부 업무 평가제도에서 제시한 준정부기관에 대한 업무평가항목이다. 경영전략과 재무성과 관리 인적자원관리 등의 항목이 포함되어 있다. 정부의 업무 평가 제도도 일반적인 기업의 평가 기준과 크게 다르지 않다는 것을 확인할 수 있다. 이를 볼 때 정부의 여러 기관에서도 공공성에 대한 평가 기준을 정립하였다고 보기 어렵다.

군의 임무를 감안할 때 군 또한 성과측정이 쉽지 않다. 군조직의 임무는 국가안보이다. 군인의 지위 및 복무에 관한 기본법 5조에 국군의 강령에서는 국군은 국가를 방위하고 자유 민주주의를 수호하며, 국토를 방위하고, 국민의 생명과 재산을 보호하여야 한다는 내용을 담고 있다. 국가의 안전을 보장한다는 행동은 언제나 외부의 환경에 크게 영향을 받는다. 국가 간의 관계는 정치, 경제, 외교, 국방 등 다양한 요소의 복합적인 관계의 결과이며, 주변국의 위협 정도에 따라 안보의 위협도가 크게 바뀌게 된다. 따라서 군조직이 효율 효과적으로 운용되더라도 주변의 위협이 강화되는 환경에서는 안보 위협이 증가하며, 평화적인 관계와 협력이 강조되는 환경에서는 안보 위협이 줄어들기 때문에 군의 성과를 정량적으로 측정하기가 매우 어렵다고 할 수 있다.

근본적으로 군조직 전체 성과를 평가하기 어렵다는 점은 개인의 성과 평가에 대한 객관성 확보에도 부정적인 영향을 미친다. KPI를 설정하거나 개인의 성과측정

12) 출처 : 국무조정실 홈페이지, https://www.evaluation.go.kr/web/page.do?menu_id=33, 확인일 2024. 7. 18.

에 있어서 주관적인 해석과 평가가 들어갈 수밖에 없는 상황이라고 할 수 있다. 성과에 따라 공정한 배분, 혹은 진급 반영이 어려운 이유가 이러한 조직의 특성에 있다.

2. 군조직에서 진급 제도

진급이 갖는 의미는 개인의 성과에 대한 보상, 권한과 책임의 확장, 개인 역량의 인정이라고 할 수 있다. 군조직은 상하관계가 뚜렷한 대표적인 조직으로 진급에 대한 의미는 더 각별하다. 하지만, 근본적으로 성과를 객관화하기 어려운 구조이며, 평가자의 주관적인 가치판단이 큰 영향을 미치는 구조이다.

군에서는 객관성 확보를 위해 부대 내 평가 외에도 교육 성적, 개인 상훈, 보직 이수 등의 항목을 종합적으로 평가하고 있다. 투명하고 객관적인 판단을 위한 절차이지만, 이처럼 여러 평가항목을 사용할 경우 가중치가 매우 중요한 역할을 하게 된다. 그리고 설정된 가중치는 조직에서 선발하고자 하는 인재상에 부합하도록 설정되어야 한다. 즉 군에서 원하는 인재상이 주어진 업무에 책임감을 갖고 묵묵히 수행하는 인재인지, 혹은 학습 능력이 높고 여러 보직을 두루 거친 인재인지에 따라 각 평가항목별 가중치를 설정해야 한다.

가중치가 여러 차례 검증을 통해 설정되었다면, 각 항목의 평가절차가 공정하게 수행될 수 있도록 시스템을 구성하는 것이 그다음으로 고려해야 할 부분이다. 여기서, 주목해야 할 부분은 최초의 가중치를 설정하는 것이 실제로 반영되는지 검증해야 한다는 점이다. 평가항목에 따라 평가하는 방법 또한 상이하기 때문에 각 점수의 분포가 일정하지 않을 가능성이 크다. 즉, 어떤 항목은 전 진급 대상자가 유사한 점수를 받지만, 어떠한 항목은 대상자별로 차이가 크다면, 처음 설정한 가중치에 따라 영향력을 미치는 것이 아니라, 차이가 큰 항목의 영향력이 향상되는 결과가 나타나게 된다. 따라서 가중치가 왜곡되는 현상이 없는지 확인하기 위하여 정량적인 분석을 하여 이를 검증하는 절차가 필요하다고 할 수 있다. 이러한 검증을 하지 않고 왜곡된 가중치가 지속적으로 적용되는 경우 심사대상자는 비교적 낮은 가중치가 배정된 항목이 당락을 결정하는 경험을 하게 될 것이며, 본인이 집중해야 하는 역량에 대한 우선순위에 혼란을 겪을 수 있으며, 심사 결과에 대한 공정성에 의구심이 생길 수 있다. 다시 말해 진급을 위하여 배점이 높은 항목을 받기 위해 최선을 다하였는데, 해당 항목은 약간 모자라지만 다른 항목의 평가가 더 좋은 경쟁자가 진급 선발이 되는 경험을 하게 된다면 절차적 공정성의 의문을 품

게 될 수 있다. 군은 인사제도에서 분배적 공정성 즉 개인의 성과에 따른 공정한 배분이 어려운 구조이다. 이러한 상황에 절차적 공정성까지 결합된다면 인사제도 자체에 대한 신뢰가 떨어지는 결과를 초래할 수가 있다. 따라서, 본 연구에서는 각 평가항목에 대한 영향력을 분석하여 설정된 가중치와 비교하여 표면적이 중요도에 따라 진급이 이루어지는지 확인하고자 한다.

추가로, 다면평가에 대한 유효성을 검증하고자 한다. 다면평가는 심사대상자를 상급자, 동급자, 하급자가 평가한 결과이다. 현재는 평가항목으로 사용되지는 않으며, 심사위원의 참고 자료로 활용되고 있다. 군의 진급평가가 주관적 평가들이 있어 객관성을 확보하기 쉽지 않다는 점이 지적되고 있는데, 다면평가는 평가자를 다각화하여 객관성을 확보하는 데 도움을 줄 수 있다. 아울러 피평가자가 결과를 수긍하는 것에도 긍정적인 효과를 기대할 수 있다. 하지만 평가항목으로 추가하기 전에 효용성을 검증할 필요가 있으므로, 본 연구에서는 다면평가 결과가 진급자를 예측하는 데 도움이 되는지 추가적인 모델을 통하여 확인하였다. 기존 방식보다 다면평가를 포함한 예측값이 더 정확하다면 진급 대상자 선발에 유용한 도구로 볼 수 있을 것이다.

3. 연구 범위와 방법

진급 모델에 대한 연구를 위하여 5년간의 공군 장교 및 부사관의 진급평가항목과 다면평가에 대한 점수와 진급 결과를 사용하였다. 개인정보보호를 위하여 평가항목에 대한 점수를 제외한 개인정보는 모두 제거된 자료를 사용하였다.

연구 방법은 베이지안 네트워크를 사용하였다. 베이지안 네트워크는 머신러닝 방법 중 널리 활용되는 도구로 기존의 축적된 자료를 바탕으로 주어진 환경에 대한 결과값을 예측하는 방법으로 활용된다. 본 연구에서는 기존의 진급 자료를 학습시켜 각 항목별 영향력을 검증하는 데 활용하였다. 아울러, 기존 평가 방식에서 다면평가를 추가한 모델을 학습시켜 진급에 대한 예측력의 변화를 살펴봄으로써 다면평가의 효용성을 확인하였다.

II. 이론적 배경

1. 인적자원 관리와 진급

1.1. 진급의 역할

인적자원과 관련된 연구에서 진급의 역할은 다양하다. 첫째, 진급을 보상의 수단으로 보는 시각이다. 조직 내에서 진급은 종종 연봉 인상과 같은 경제적인 보상을 동반한다. 이와 같은 역할은 개인별 성과가 뚜렷한 직종과 조직에서 더욱 뚜렷하게 나타나는 현상으로, 주식 분석 전문가나 운동선수를 예로 들 수 있다.

이러한 조직의 특징은 각 개인의 성과가 명확하게 드러나며, 고성과자가 조직에 미치는 영향력이 크며, 개인의 역량이 같은 분야에 다른 조직에서도 충분히 활용될 가능성이 크다는 특징을 보인다. 이른바 스타 플레이어(Star player)가 중요한 역할을 해서 스타 플레이어가 다른 조직으로 가지 않도록 조직이 할 수 있는 최선의 조건과 직위를 부여한다.

둘째, 진급의 의미를 권력의 관점에서 이해할 수 있다. 높은 직책으로 갈수록 더 큰 권한을 갖게 되고, 조직의 전략적인 의사를 결정하는 과정에도 더 많이 참여하게 된다. 리더는 담당하는 조직원의 역할을 결정하고 과업을 수행하는 데 필요한 자원을 할당하며, 조직원의 성과를 평가하는 임무를 수행한다. 개인의 성과측정이 어렵고 팀 전체의 역량으로 성과를 평가받는 조직일수록 권력의 영향력이 강해지므로 진급자 선발에 주의를 기울일 필요가 있다.

셋째, 진급을 통한 역할의 변화가 생기므로, 해당 역할에 걸맞은 인재를 선발하는 의미가 있다. 특히 팀의 성과를 평가받는 조직에서 진급은 과거의 성과보다는 진급 후 역할에 적합성을 볼 필요가 있다. 직접 업무를 하는 실무자에게서는 자신의 업무를 잘 해냈던 직원이 진급하여 관리직으로 업무가 변경되면서 업무 성과가 크게 떨어지는 경우가 발생한다. 이는 주어진 직무 성격이 달라짐에 따라 발생하는 현상이다. 진급되어 상급자가 될수록 적절한 방향성을 제시하고 다양한 과업에 대하여 대응 방법을 결정하고, 조직원과의 대인관계 관리 능력이 중요해진다. 업무의 성격이 바뀔에 따라 요구되는 능력이 바뀔에 적응하지 못하면 조직에 성과를 떨어뜨리는 저성과자로 분류된다.¹³⁾ 따라서, 진급 후에 맡게 될 직책에 걸맞은 능

13) 노명화(2011) B-플레이어를 위한 성공의 재발견, 북코리아

력을 갖추었는가를 기준으로 진급자를 평가하고 선발하며, 필요하다면 조직 외부에서 해당 직책에 맞는 책임자를 찾는 것이 필요할 수 있다.

결론적으로, 진급은 직원들의 동기부여, 조직의 전략적 목표 달성, 적합한 인재 선발 등 다양한 측면에서 중요한 역할을 한다. 진급은 조직의 성과를 증진하고, 장기적인 관점에서 조직의 안정성을 확보하는 데 필수적인 요소라고 할 수 있다. 공정하고 객관적인 진급평가 시스템을 통하여 조직이 필요로 하는 가치를 실현할 수 있는 인재를 적시에 진급시키는 것이 중요하다.

1.2. 군에서 진급의 역할

군은 상하관계가 뚜렷하여 군조직에서 진급은 일반적인 조직보다 더 강한 의미가 있다. 특히 군조직에서 상급자는 전시 명령과 복종을 전제로 하므로 모든 조직원이 진급에 관한 관심이 매우 크다. 최근 직장인을 대상으로 한 설문 결과에서 진급하지 않는 편이 더 좋다는 응답자와 진급을 원한다는 응답자가 비슷한 비율로 제시된 것과 매우 상반된 상황이다. 따라서 진급의 정당성과 공정성에 대한 요구가 일반 기업보다 훨씬 중요한 가치라고 할 수 있다.

조직에 진급의 정당성을 확보하기 위한 기본적인 시각은 성과에 대한 것이다. 하지만, 군조직의 궁극적인 가치는 국가의 안보를 책임지는 것이다. 제공하는 가치 자체가 수치로 환산되거나 개인의 성과가 평가되기 어려운 성격의 것이기 때문에 개인의 성과를 평가하기 어려울 수밖에 없는 환경이다.

인적자원 선발에 있어, 장교와 부사관의 접근 방법의 차이가 필요하다. 장교는 실무자의 관점에서 상급자가 될수록 관리자의 업무를 수행하게 되지만, 부사관은 경험과 전문성이 요구되는 직군에 배치된다. 장교와 부사관의 직무가 명확히 관리자와 전문가로 구분된다고 하기는 어렵지만, 역할의 경향성이 차이가 있으며, 진급을 거듭하여 상급자가 될수록 해당 경향성을 강해진다. 장교의 진급은 관리자로서 책임자를 선발하는 방법에 중점을 두어야 한다면, 부사관은 개인의 전문성에 초점을 두고 그동안의 성과에 초점을 두는 것이 바람직하다.

2. 군 인적자원 관리 연구

2.1. 군조직 인사관리 선행연구

그동안 군에서 진급 관리에 대한 연구는 정성적인 연구들이 주를 이루었다. 최명순, 문영세(2006)의 연구에서는 육군 장교의 인사제도 전반에 관한 연구를 진행하면서 진급 제도 개선 방향을 제시하고 있다.¹⁴⁾ 구용희(2014)는 군 인사관리 변화 과정을 살펴보았는데, 외부 환경에 따라 조직에서의 형평성보다 효용성을 중시하는 발전해 왔다고 평가하였다.¹⁵⁾ 이러한 연구는 기존의 이론적 관점에서 현재의 제도를 평가하거나 비교하는 등 사례연구 중심의 정성적인 연구 방법이 사용되었다.

진급 연구에서도 정량적인 연구를 통하여 연구 방법의 균형을 이루는 것이 중요하다. 구체적인 데이터들을 통하여 해당 관점이 타당성이 있는지 확인하는 과정 없이 주장하는 것에는 한계점이 있다. 이를 개선하기 위해서 실제 데이터를 바탕으로 현상을 분석하는 것부터 출발하는 것이 필요하다.

최근 AI 기술발전이 전 산업 전반에 걸쳐 큰 영향을 줌에 따라 인적자원 관리 분야에서도 AI 활용 방안에 관한 연구가 활발하다. 인원 선발 시 AI 면접을 통한 예비 심사를 거침으로써 시간과 노력을 절약하는 기업이 보고되는가 하면, 다양한 인사 기록과 경력 데이터를 AI로 분석하여 업무 적합도를 고려한 인사 배치 활용 등이 보고되고 있다. 특히 많은 양의 데이터를 처리하고 분석하는 데 AI를 활용하는 방안들이 개발 중이다.

2.2. 국방혁신 4.0과 AI 인적자원 관리

문재인 정권에서 추진하던 국방개혁 2.0을 대신하는 개념으로 2022년 5월 국방혁신 4.0 추진이 발표되었다. 제 2창군 수준으로 국방태세 전반을 재설계하는 것을 목표로 하며, AI 과학기술 강군 육성을 골자로 하고 있다. 세부 내용으로 AI 기반 유/무인 복합 전투 체계로 단계별 전환과 첨단 과학 기술 기반 군 구조 발전 등이 제시되었다. 이러한 기조 속에서 군 전반에 걸쳐 AI를 활용하려는 노력이 추

14) 최명순, 문영세. (2006). 직업군인제 발전을 위한 정책 대안 분석. 한국정책과학학회보, 10(1), 51-78. DBpia.

15) 구용희. (2014). 한국군 인사관리제도의 변천 과정에 관한 연구: 진화론적 특성변화를 중심으로. 목원대학교 박사 학위논문.

진되고 있다. 이에 따라 무기체계 분야뿐만 아니라, AI 참모와 같이 국방 전 분야에 걸쳐 AI를 활용하려는 움직임이 늘어나고 있다.

2.3. 진급 모델 연구에서의 머신러닝

그동안 부족하였던 진급 모델의 정량적인 연구를 할 때 흔히 사용되는 선형 회귀 방법보다는 머신러닝 접근법을 고려할 필요가 있다. 데이터 분석에서 널리 활용되는 방법은 선형 회귀모델이지만 진급 모델을 연구에는 한계점이 있다. 그 이유는 현재 진급 모델은 여러 진급평가항목을 종합적으로 평가하여 점수를 산정하는 방식을 취하고 있는데, 진급 모델과 같이 다양한 변수가 결과에 영향을 주는 경우는 다중선형회귀 모델을 사용할 경우 종속변수에 대한 비선형성이 증가하면 모델의 정확도가 떨어지는 문제가 발생하기 때문이다(Rasmussen, and Williams, 2007)¹⁶. 따라서, 모델의 정확도가 떨어지는 문제를 보완하기 위하여 최근에는 머신러닝을 통해 오차를 줄이는 방법이 사용되고 있다.

2.4. 다면 평가 연구

다면평가는 상급자, 하급자, 동료를 포함한 여러 평가자의 평가결과를 대상자 성장의 기회로 활용하는 제도이다.¹⁷ 미국에서 1990년대 도입되었는데, 경영환경 변화에 따라 조직시민행동(Organizational citizenship behavior)과 같은 소통이 중요해짐에 따라 기존에 평가방법의 한계점을 극복한 방법으로 주목받았다.¹⁸

다면 평가에 관한 주요 연구를 살펴보면, 초기 연구에서 전통적인 계량심리학 관점에서, 복수의 평가자가 평가하면 측정 오류가 줄어들어 정확도가 증가한다는 점에 주목한다. 이 경우 평가자의 편견이 들어갔다고 생각하기 어려워진다.¹⁹ 평가자에 대한 다양한 관점의 평가가 가능해진다는 점도 장점이라고 할 수 있다. 상급자는 과업 중심의 평가를 하지만, 하급자는 관계적 성과와 과업 성과를 모두 고

16) Rasmussen, C. E. and Williams, C. K. I. (2007) Gaussian Processes for Machine Learning, The MIT Press, Cambridge, MA,

17) Borman, W. C. (1997). 360 ratings: An analysis of assumptions and a research agenda for evaluating their validity. *Human Resource Management Review*, 7(3): 299-315.

18) Fletcher, (2001). Performance appraisal and management: The developing research agenda. *Journal of Occupational and Organizational Psychology*, 74(4): 473-487.

19) Atwater L. E., & Brett, J. F. (2006). 360-degree feedback to leaders: Does it relate to changes in employee attitudes?. *Group & organization management*, 31(5): 578-600.

려하여 평가한다. 특히 전통적 평가방법에서 간과되었던 맥락 성과를 동료와 하급자로부터 확인할 수 있다는 점이 장점이라고 할 수 있다.

개인 성과와 다면평가 간의 성과가 관련이 없다는 연구도 있고, 다면평가는 하급자에게 좋은 평가를 받기 위한 인기 위주 행동을 하는 부작용을 낳는다는 우려도 있다. 기존의 연구들이 일관적이지 못한 결과를 보이는 이유가 이론적 기반이 약한 데이터 위주의 연구들이었다는 지적도 있다.²⁰⁾

하지만, 피평가자의 직무 중에 조직원 관리가 중요한 요소라면, 조직원을 관리하는 능력이 직무수행에 중요한 요소로 볼 수 있다. 따라서 직무의 성격에 따라 다면평가 결과를 활용하는 것이 적합할 수 있다.

군의 다면평가에 대하여 적합성을 검토한다면, 직무에 따라 달라지겠지만 장교와 부사관의 업무적 성격이 다르기 때문에 구분하여 생각해 볼 필요가 있다. 장교와 부사관 모두 직책에 따라 관리자로서의 역할과 전문가로서의 역할을 수행하지만, 일반적으로 장교가 지휘관으로서 부대의 관리를 책임지는 역할이 부여되는 경우가 많다는 것을 감안하면, 장교 진급 시 다면평가 결과 활용이 긍정적인 영향을 미칠 수 있다. 부사관의 경우에도 병사를 관리하는 입장에 놓이는 경우가 적지 않으나, 공군의 부사관들은 전문적인 영역을 담당하는 경우가 더 많으므로, 다면평가의 당위성이 다소 줄어들 수 있다고 하겠다.

다면 평가는 긍정적인 역할에 주목할 필요가 있다. 다면평가는 복수의 평가자가 참여한다는 점에서 측정 오류 즉 편견이 끼어들 여지가 줄어들고, 업무 성과적인 특성 말고도 관계적인 특성을 반영할 수 있다는 장점이 있다. 다만, 평가항목으로써 적정성을 판단한다면, 관계적인 특성, 리더십이 얼마나 중요한 직무인가에 따라 개인 평가에 반영하는 것이 필요한지를 판단해야 한다. 장교와 부사관의 직무를 비교한다면 다면평가 결과 반영은 장교에 더 적합할 가능성이 있다.

현재 공군에서는 다면평가를 상급자, 동기, 하급자를 대상으로 하고 있다. 하지만, 다면평가는 위원회에 심사위원에게 참고사항으로 제공될 뿐 평가항목으로는 반영되지 않고 있다. 본 연구에서는 다면평가를 장교와 부사관에 반영하는 것이 적절한가를 판단하기 위하여 다면평가 데이터를 추가하였을 때 모델을 추가로 작성하고 기본 모델과 비교함으로써 장교와 부사관 각각의 적합성을 판단해 보고자 하였다.

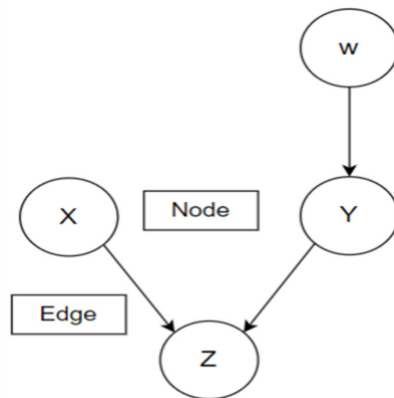
20) 김선형, 강성춘. (2017). 다면평가 연구 리뷰, 시사점 및 향후 연구 방향. 노사관계연구. 28. 123-153

3. 주요 머신러닝 알고리즘 특성

진급 데이터를 분석함에 있어 머신러닝 기법을 적용할 때 각 모델의 특성을 고려하여 본 연구에 적합한 머신러닝 기법을 선정할 필요성이 있다. 따라서 기존에 널리 활용되는 알고리즘을 중심으로 특성을 살펴보면 다음과 같다.

3.1. 베이저안 네트워크

머신러닝 방법 중 하나인 베이저안 네트워크는 토마스 베이즈(Thomas Bayes)가 제안한 베이즈 정리를 기반으로 한다. 진급 모델 연구에서 종속변수는 승진 여부로 더미변수의 형태로 나타나기 때문에 다중선형회귀 모델을 사용할 경우 모델의 정확도가 떨어질 수 있다. 아울러 진급심사는 자력점수 외에도 2차에 걸친 심사 결과가 종합적으로 반영되어 결정된다는 점에서 선형회귀의 잠재변인에 대한 문제에서 자유롭기 어렵다. 그러나 베이저안 네트워크는 변수 간의 상호 의존성을 확률적으로 표현할 수 있어, 모든 변수에 대한 정보가 명확하지 않거나, 변수들 간의 관계가 복잡한 경우에도 불확실한 정보와 변수 간의 관계를 기반으로 추론할 수 있어 유용하다(김예리, 이두열, 2022)²¹⁾. 따라서, 진급심사 과정을 연구하는데 적합도가 높다고 할 수 있다.

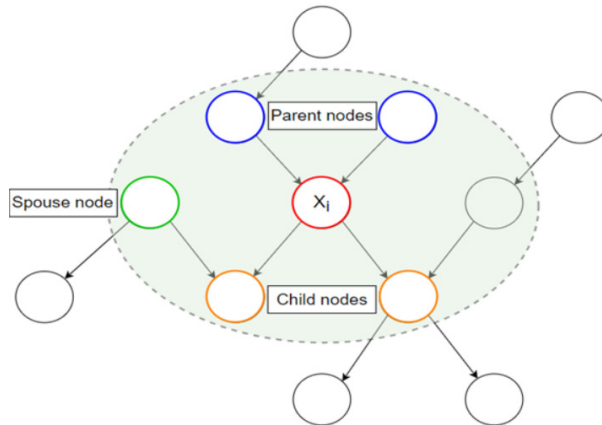


〈그림 2-1〉 기본 베이저안 네트워크 모형

21) 김예리, 이두열. (2022). 베이저안 네트워크 기반 상황인식 및 방책 결정 지원 모델. 한국국방경영분석학회지, 48(1), 131-141.

〈그림 2-1〉은 기본적인 베이지안 네트워크의 예시이다. 그림과 같이 베이지안 네트워크는 노드(Node)와 엣지(Edge)로 구성된 방향성 비순환 그래프(Directed Acyclic Graph, DAG) 모델이다. 여기서 타원은 노드를 나타내고, 화살표는 엣지를 의미한다. 각 노드는 확률변수로 엣지로 연결되는 경우 연관성을 가지며, 연결되지 않은 노드들은 서로 조건부 독립이다. 아래 그림을 보면 화살표가 X노드에서 Z노드로 향하는 것으로 볼 수 있는데 여기서 X는 Z의 부모 노드이고, Z는 X의 자식 노드라고 한다. 즉, 부모 노드는 다른 노드에 영향을 미치는 노드를 의미하며, 자식 노드는 부모 노드로부터 영향을 받는 노드를 의미한다.

베이지안 네트워크는 노드와 엣지를 통하여 노드에 대한 조건부 확률을 도출할 수 있다. 노드들을 조건부 독립을 가정하며, 원인과 결과를 연결시켜서 인과 관계에 따라 네트워크를 구성한다. 이 과정을 통해 조건부 독립을 가정한 적합한 네트워크가 형성되고, 특정 노드에 대한 조건부 확률을 계산할 수 있다.



〈그림 2-2〉 마코프 블랭킷(Markov Blanket) 모델 예시

〈그림 2-2〉는 다양한 요소를 포함하여 베이지안 네트워크를 구성한 예시이다. 조금 더 복잡한 베이지안 네트워크 구조를 분석하기 위해서는 마코프 블랭킷(Markov Blanket)과 조건부 확률의 연쇄법칙(Chain Rule for Conditional Probability)에 대한 이해를 해야 한다. 마코프 블랭킷은 특정 노드의 확률 계산 시 해당 노드에 직접적 영향을 미치는 노드들의 집합이다. 마코프 블랭킷은 부모 및 자식 노드, 자식 노드의 또 다른 부모 노드를 포함하며 V 연결 구조

(V-structure)를 형성한다. 마코프 블랭킷은 복잡한 확률적 의존성을 단순하게 변형시키는 데 중요한 역할을 하며 노드 간 독립, 종속 등 관계를 명확히 하고 불필요한 부분을 제거하여 추론과 확률 계산을 쉽게 할 수 있다.²²⁾

조건부 확률의 연쇄법칙을 통하여 복잡한 확률 구조를 간단한 조건부 확률의 곱으로 분해해 볼 수 있다. 연쇄법칙을 활용하면 계산을 단순화하고 확률적 추론을 쉽게 할 수 있으며, 네트워크 전체의 구조를 이해하기 용이하다. 이러한 조건부 확률의 연쇄법칙은 식 (1)과 같이 나타낼 수 있다.

$$\begin{aligned} P(x_1, x_2, \dots, x_n) & \quad (1) \\ & = P(x_1)P(x_2|x_1)P(x_3|x_1, x_2) \cdots P(x_n|x_1, x_2, \dots, x_{n-1}) \end{aligned}$$

조건부 독립 가정한 베이지 정리, 연쇄법칙과 곱셈규칙을 활용하면 식 (2)와 같이 단순하게 표현할 수 있다²³⁾.

$$P(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{parent}(x_i)) \quad (2)$$

3.2. 나이브 베이지 모델

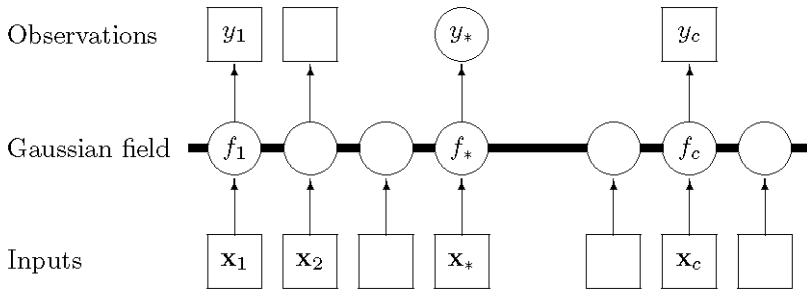
본 연구는 우선적으로 베이지안 네트워크를 통하여 진급 데이터에 대한 정량연구를 진행하였다. 상기한 바와 같이 베이지안 네트워크는 노드와 엣지로 구성된다. 진급 모델의 경우 선발 과정에서 평가항목들이 각각 독립적으로 진급 결과에 영향을 미치는 구조이다. 따라서 본 연구에서는 나이브 베이지(Naive Bayes) 모델을 사용하였다. 나이브 베이지 모델은 1개의 단으로 구성되는 모델로, 진급결과에 각각 평가항목이 독립적으로 영향을 미치는 형태를 반영하여 구성하였다. 실제로 진급평가항목들이 직접 평가 점수에 독립적으로 반영되고 최종 결과에 영향을 주므로, 나이브 베이지 모델이 가장 현실을 잘 반영한 모델이라고 할 수 있다. 본 연구에서는 나이브 베이지를 기본 형태로 하며, 연구 목적에 따라 항목을 추가로 구성하는 등의 변형을 주고 이때 나타나는 변화에 대하여 살펴보았다.

22) D. Margaritis. Learning Bayesian Network Model Structure from Data, (2003). Doctoral dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA,

23) Sucar, L. E., Bielza, C., Morales, E. F., Hernandez-Leal, P., Zaragoza, J. H., & Larranaga, P., "Multi-label classification with Bayesian network-based chain classifiers", Pattern Recognition Letters, Vol.41, pp.14-22, 2014.

3.3. Gaussian Process Regression (가우시안 과정 회귀)²⁴⁾

Gaussian Process Regression (GPR)은 비선형 회귀 문제를 해결하기 위한 베이저안 접근법에 기반한 강력한 모델이다. GPR은 특정한 함수 형태를 가정하지 않고, 주어진 데이터에 대해 함수 분포를 직접 모델링함으로써 매우 유연하게 비선형 관계를 학습할 수 있다. GPR은 확률적 과정(Probabilistic process)으로, 각 데이터 포인트 간의 상관관계를 기반으로 미래 값을 예측하는 데 사용된다. 이 모델은 특히 데이터의 불확실성을 잘 반영하며, 그 예측값에 대한 신뢰도(분산)까지 함께 제공하는 강점을 갖고 있다.



〈그림 2-3〉 가우시안 과정 회귀

GPR의 핵심은 커널 함수(Kernel function)로, 이는 두 데이터 포인트 간의 유사도를 측정하는 역할을 한다. GPR에서는 다양한 커널 함수를 사용할 수 있으며, 커널의 선택은 데이터의 특성에 맞는 비선형 패턴을 반영하는 데 매우 중요하다.

3.3.1. Gaussian Process의 개념

Gaussian Process는 함수 $f(x)$ 가 다변량 가우시안 분포를 따른다고 가정한다. 이는 주어진 관측값을 기반으로 미래의 값을 예측하는 데 유용하며, 특히 데이터 간의 상관관계를 반영하는 커널 함수를 통해 비선형적 패턴을 학습할 수 있다.

3.3.2. 커널 함수의 종류

GPR에서 성능을 결정하는 중요한 요소는 커널 함수의 선택이다. 주요 커널 함

24) Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian Processes for Machine Learning. MIT Press.

수는 다음과 같다. Exponential Kernel은 두 데이터 포인트 사이의 거리가 멀어질수록 유사도가 지수적으로 감소하는 함수로, 급격한 변동이 있는 데이터에 적합하다. Matern 5/2 Kernel은 다양한 수준의 부드러움을 조절할 수 있어, 데이터가 부드럽게 변하거나 급격하게 변할 때 모두 적용 가능하다. Squared Exponential Kernel은 매우 부드러운 함수 관계를 모델링하는 데 사용되며, 연속적이고 부드러운 데이터에 적합하다. Rational Quadratic Kernel은 여러 스케일에서 변동성이 큰 데이터를 모델링하는 데 유리하며, 다양한 크기의 변화를 포착하는 데 적합하다.

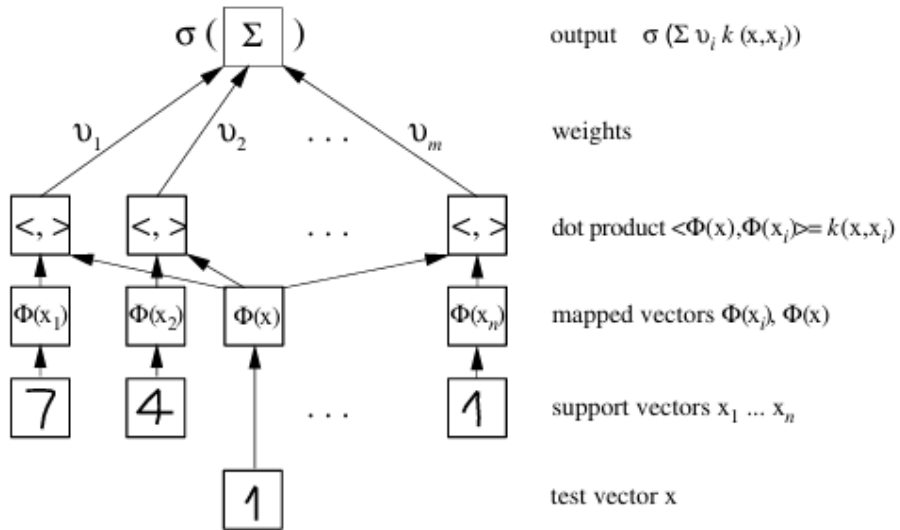
3.3.3. GPR의 장점과 한계

GPR은 비선형 데이터에서도 유연하게 모델링할 수 있으며, 예측과 함께 그 예측값에 대한 신뢰도(분산)를 제공하는 강점을 가지고 있다. 그러나 계산 비용이 크기 때문에 데이터 세트가 커질수록 학습과 예측 시간이 증가하며, 커널 함수 선택에 따라 모델 성능이 크게 달라질 수 있다는 한계를 지닌다.

3.4. Support Vector Machine (서포트 벡터 머신)²⁵⁾

Support Vector Machine (SVM)은 분류 및 회귀 문제를 해결하기 위한 강력한 지도 학습 알고리즘이다. SVM은 서포트 벡터(Support vectors)를 사용하여 클래스 간의 경계를 정의하며, 다양한 커널 함수를 사용해 비선형 데이터를 처리할 수 있다. SVM은 고차원 공간에서 데이터 간의 최적의 경계(결정 경계)를 찾는다. 마진(Margin)은 서포트 벡터와 결정 경계 사이의 거리를 의미하며, 이 마진을 최대화하는 것이 SVM의 주목적이다.

25) Scholkopf, B., & Smola, A. J. (2001). Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press.



〈그림 2-4〉 Architecture of Support Vector Machine

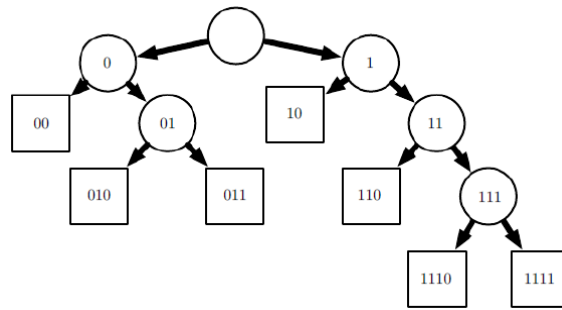
3.4.1. SVM 모델의 종류

SVM Coarse Gaussian은 가우시안 커널을 사용하며, 커널 폭이 큰 경우에 해당하여 전체적인 큰 패턴을 반영하는 데 적합하지만 세밀한 경계를 형성하지 못할 수 있다. SVM Medium Gaussian은 중간 크기의 가우시안 커널로, 데이터의 전체적인 패턴과 세부적인 특성을 균형 있게 반영하며 비선형적 패턴을 처리할 수 있다. SVM Fine Gaussian은 매우 작은 가우시안 커널을 사용해 가까운 데이터 포인트들 간의 차이를 세밀하게 구분할 수 있지만 과적합의 위험이 존재한다. SVM Linear는 선형 데이터를 빠르게 학습할 수 있지만 비선형 데이터를 처리하는 능력은 제한적이다. SVM Quadratic은 2차 다항식 커널을 사용해 이차적인 경계를 학습하고, SVM Cubic은 3차 다항식 커널을 통해 더 복잡한 비선형 경계를 모델링하지만 복잡도가 높아질수록 계산 비용과 과적합 가능성이 증가한다. SVM Kernel은 다양한 커널을 적용해 비선형 데이터를 처리할 수 있어 유연하게 다양한 데이터에 적용할 수 있다.

3.4.2. SVM의 장점과 한계

SVM의 장점은 고차원 공간에서도 잘 작동하며, 적은 수의 서포트 벡터로 데이터를 효과적으로 분류할 수 있다는 점이다. 또한, 정규화 파라미터를 사용해 과적합을 방지할 수 있다. 하지만 데이터 세트가 매우 크거나 노이즈가 많은 경우 성능이 저하될 수 있으며, 커널 함수와 하이퍼파라미터를 신중하게 선택해야 한다는 한계가 있다.

3.5. Decision Tree (의사결정나무)²⁶⁾



〈그림 2-5〉 의사결정나무 모형

데이터를 기반으로 의사결정을 내리기 위한 규칙 기반 알고리즘이다. 트리 구조는 데이터를 반복적으로 분할하여 작동하며, 각 분기점(Node)에서 특정한 데이터 특징을 기준으로 데이터를 나눈다.

〈그림 2-5〉와 같은 과정을 반복하여 최종적으로 리프 노드(Leaf Node)에서 예측값을 산출한다. 결정 트리는 분류 문제와 회귀 문제 모두에 적용할 수 있으며, 직관적이고 해석이 용이한 모델로 널리 사용된다.

3.5.1. 의사결정나무 모형의 종류

의사결정나무는 분할의 세밀함과 모델 복잡도에 따라 여러 형태로 나뉜다. Fine Tree는 데이터를 매우 세밀하게 분할하여 높은 정확도를 제공하지만 과적합의 위험이 있다. Medium Tree는 덜 세밀한 분할을 통해 가지치기를 적절히 수행하여

26) Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. Wadsworth.

과적합을 방지하면서도 적당한 예측 성능을 유지할 수 있다. Coarse Tree는 매우 단순한 구조로 데이터를 크게 분류하여 계산 속도가 빠르고 과적합의 위험이 적지만, 데이터의 세부적인 패턴을 충분히 반영하지 못할 수 있다.

3.5.2. 의사결정나무의 장점과 한계

의사결정나무는 이해하고 해석하기 쉬운 장점을 가지고 있으며, 트리 구조를 통해 데이터의 분할 과정을 직관적으로 시각화할 수 있다. 또한, 범주형 변수와 연속형 변수를 모두 다룰 수 있고, 빠르게 학습할 수 있으며 데이터 전처리가 비교적 단순하다. 그러나 트리가 깊어질수록 과적합의 위험이 커지고, 트리 구조는 작은 변화에도 크게 변할 수 있어 모델의 안정성이 낮아질 수 있다는 한계가 있다. 또한, 결정 트리 단독으로는 예측 성능이 제한적일 수 있어, 보통 앙상블 기법과 결합해 사용되는 경우가 많다.

3.6. Ensemble Methods (앙상블 기법)

앙상블 기법은 여러 개의 학습 모델을 결합하여 성능을 향상시키는 방법이다. 단일 모델이 충분한 예측 성능을 내지 못할 때, 여러 개의 약한 학습기를 결합하여 더 강력한 예측 성능을 얻을 수 있다. 대표적인 앙상블 기법에는 Bagging과 Boosting이 있다. 이 두 방법 모두 결정 트리(Decision Tree)를 기반으로 하지만, 데이터를 처리하는 방식이 다르다.

3.6.1. Bagged Trees (배깅 트리)²⁷⁾

Bagging은 원본 데이터를 여러 개의 부트스트랩 샘플링(데이터의 일부를 중복 허용하여 랜덤하게 추출)하여 각각의 샘플에 대해 결정 트리를 학습하고, 그 결과를 결합하는 방식이다. 여러 샘플 데이터로 학습된 결정 트리의 예측 결과는 회귀의 경우 평균을, 분류의 경우 다수결 투표로 최종 예측을 내린다. Bagging의 장점은 각 트리가 독립적으로 학습되어 과적합을 방지하고, 데이터의 다양한 패턴을 반영할 수 있다는 점이다. 그러나 개별 트리들이 독립적으로 작동하기 때문에 데이터의 특정 패턴을 충분히 반영하지 못할 수 있다는 한계가 있다.

27) Breiman, L. (1996). Bagging Predictors. Machine Learning.

3.6.2. Boosted Trees (부스팅 트리)²⁸⁾

Boosting은 약한 학습기를 순차적으로 학습시키며, 이전 학습기에서 잘못 예측된 데이터에 더 많은 가중치를 부여하여 성능을 향상시키는 기법이다. Boosted Trees는 이러한 방식으로 여러 결정 트리를 학습해 결합한 모델로, 첫 번째 트리가 학습된 후, 그 트리에서 잘못 예측된 데이터에 가중치를 부여하여 다음 트리를 학습시키는 과정을 반복한다. Boosting의 장점은 매우 높은 예측 성능을 제공하며, 소수의 트리로도 높은 정확도를 달성할 수 있다는 것이다. 그러나 순차적인 학습 방식으로 인해 병렬 처리가 어렵고, 과적합이 발생할 수 있다는 한계가 있다.

3.7. Linear Regression (선형 회귀)²⁹⁾

Linear Regression (선형 회귀)는 독립 변수와 종속변수 간의 선형 관계를 모델링하는 기법이다. 주어진 데이터에 가장 잘 맞는 직선을 그려, 입력 변수들이 출력 변수에 어떻게 영향을 미치는지 예측할 수 있다. 선형 회귀는 가장 기본적인 회귀 기법 중 하나로, 최소 제곱법 (Least Squares)을 사용하여 잔차(Residuals)를 최소화한다.

3.7.1. Linear Regression의 종류

기본 선형 회귀(Basic Linear Regression)는 단순 선형 모델을 사용하여 독립 변수와 종속변수 간의 선형 관계를 모델링한다. 상호작용 선형 회귀(Interaction Linear Regression)는 독립 변수 간의 상호작용을 고려하여 다중 회귀 모델을 생성하는 방식이다. Stepwise Linear Regression은 변수 선택 기법을 사용해 중요한 변수들만 선택함으로써 모델의 복잡도를 줄이는 기법이다. 강건 선형 회귀(Robust Linear Regression)는 노이즈와 이상치가 있는 데이터에서도 견고하게 작동하며, 일반 선형 회귀가 이상치에 민감한 반면, M-추정법(M-estimator) 등의 기법을 통해 이상치의 영향을 줄이는 방식으로 작동한다.

3.7.2. Linear Regression의 장점과 한계

선형 회귀는 이해하고 해석하기 쉬우며, 계산 비용이 적고 매우 효율적이라는

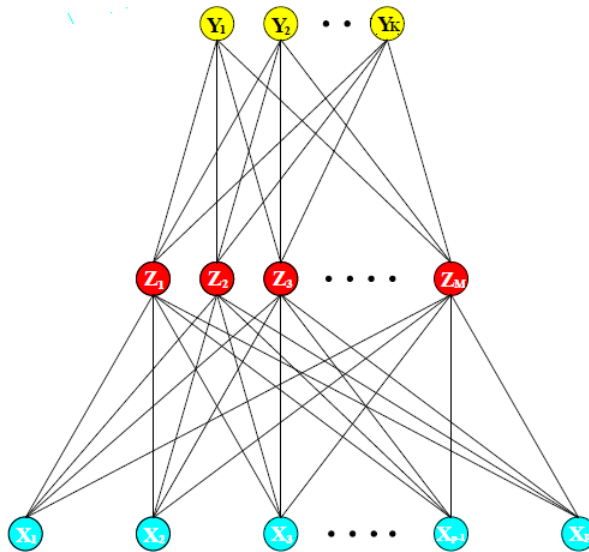
28) Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*.

29) Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. Wiley.

장점을 가지고 있다. 특히, 선형적으로 분포된 데이터에서는 강력한 예측 성능을 발휘한다. 그러나 데이터가 비선형적이거나 독립 변수들 간의 상호작용이 복잡할 경우 성능이 크게 떨어질 수 있으며, 이상치에 매우 민감해 이상치가 많을 경우 예측 성능이 저하될 수 있다는 한계가 있다.

3.8. Neural Networks (신경망)³⁰⁾

Neural Networks (신경망)은 인간 뇌의 신경 구조를 모방한 기계 학습 알고리즘으로, 여러 계층(Layer)으로 구성된 뉴런(Neuron)들이 상호작용하면서 입력 데이터를 처리하고 예측을 수행한다. 신경망은 복잡한 비선형 패턴을 학습할 수 있어, 다양한 예측 문제에서 우수한 성능을 발휘한다.



〈그림 2-6〉 신경망 모형

3.8.1. 신경망 모델의 종류

신경망의 구조와 은닉층의 크기에 따라 다양한 종류로 나뉜다. Narrow Neural Network는 적은 수의 뉴런과 은닉층을 가진 신경망으로, 간단한 데이터 패턴을 처리하는 데 적합하다. 이 모델은 구조가 단순하여 학습 속도가 빠르지만, 복잡한

30) Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

문제를 처리하는 데 한계가 있을 수 있다. Medium Neural Network는 중간 크기의 은닉층과 뉴런 수를 가진 모델로, 적절한 복잡성과 성능을 제공하며 비선형 패턴을 처리할 수 있고, 과적합과 과소적합 간의 균형을 유지할 수 있다. Wide Neural Network는 많은 뉴런을 가진 은닉층을 포함한 모델로, 복잡한 데이터 패턴을 세밀하게 처리할 수 있고 큰 데이터셋에서 효과적이지만, 계산 비용이 높고 과적합의 위험이 있다.

Bi-layered Neural Network는 두 개의 은닉층을 가진 신경망으로, 단층 구조보다 복잡한 비선형 패턴을 학습할 수 있으며, 적절한 성능을 제공하지만 학습 속도가 느려질 수 있다. Tri-layered Neural Network는 3개의 은닉층을 가진 신경망으로 매우 복잡한 데이터 패턴을 학습할 수 있으며, 높은 복잡도의 데이터를 처리할 수 있지만, 계산 비용이 높고 학습 시간이 오래 걸릴 수 있다.

3.8.2. 신경망의 장점과 한계

신경망은 복잡한 비선형 패턴을 학습할 수 있고, 다양한 데이터 유형에 적용 가능하다는 장점이 있다. 특히 큰 데이터 셋을 처리할 때 매우 강력한 성능을 발휘한다. 그러나 계산 자원이 많이 필요하고, 학습 시간이 길어질 수 있으며, 과적합의 위험이 있어 정규화 기법 등을 사용해 이를 방지해야 한다는 한계가 있다.

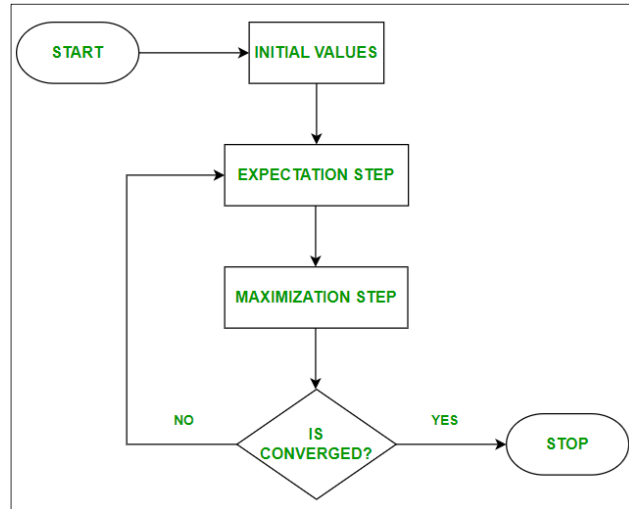
4. 기댓값 최대화 알고리즘

기댓값 최대화(Expectation-Maximization, EM) 알고리즘이란 주어진 데이터가 관찰 가능하고, 이와 밀접하게 관련된 어떤 정보가 관찰되지 않아 잠재적인 상황에서 확률 모델의 파라미터를 효과적으로 추정할 때 사용하는 반복 알고리즘이다.³¹⁾ EM 알고리즘은 히든 마르코프 모델(Hidden Markov Models, HMMs)이나 베이저안 네트워크(Bayesian Networks)와 같은 확률 모델 사용 시 데이터가 불완전할 경우 관찰된 데이터에서 학습된 패턴을 기반으로 잠재 변수를 추정하고 이를 통해 불완전한 데이터에서의 확률 모델링을 가능하게 한다.

EM 알고리즘은 누락된 확률분포를 추정하는 단계(E-step)와 이렇게 완성된 데이터 및 확률분포를 사용하여 모델 파라미터를 재추정하는 단계(M-step)를 번갈아

31) G.-T. Park, C. Tran, and W.-Y. Shin, 2019, "Network-Embedding-Based Link Prediction Using the Expectation Maximization Algorithm," J. Korean Inst. Commun. Inf. Sci., Vol. 44, No. 11, pp. 2123-2126.

수행하며 데이터 값을 예측하는데, 이를 도식화하면 <그림 2-7>과 같이 나타낼 수 있다.



<그림 2-7> 기댓값 최대화 알고리즘 흐름도

EM 알고리즘에서는 최초 값이 입력되면 기대(E) 단계에서는 관찰된 데이터와 현재 매개변수를 사용하여 추정 가능한 확률분포를 계산하며, 최대화(M) 단계에서는 기대 단계에서 추정한 데이터를 바탕으로 새 매개변수를 결정한다. 이러한 과정은 일반적으로 로그 가능도 또는 매개변수 추정치의 변화가 사전 정의된 임계값 아래로 떨어질 때까지 계속되며 해당 기준을 만족할 경우 결과값을 산출한다.

EM 알고리즘의 적용을 위한 E-step, M-step을 수식으로 나타내면 식 (3), (4)와 같이 나타낼 수 있다.

$$Q(\theta|\theta^{(t)}) = E_{Z|X,\theta^{(t)}}[\log P(X,Z;\theta)] \quad (3)$$

$$\theta^{(t+1)} = \operatorname{argmax}_{\theta} Q(\theta|\theta^{(t)}) \quad (4)$$

위 식에서 X 는 관찰된 데이터를 의미하며, Z 는 잠재 변수를 의미한다. 또한, θ 는 모델의 파라미터를 $\theta^{(t)}$ 는 현재 단계에서의 파라미터 추정치를 나타내며, Q 는 예상되는 기댓값을 의미한다.

5. 데이터 이산화

진급 모델에 사용된 데이터는 점수화된 연속형 데이터(Continuous data)이다. 진급심사에 사용되는 데이터는 항목이 다양하고 범위가 넓기 때문에, 모델의 복잡성을 줄이기 위해 데이터를 이산화할 필요가 있다. 이산화는 연속적인 값을 특정 구간으로 나누어 범주형(Categorical data) 데이터로 변환하는 과정을 의미한다.³²⁾ 이를 통해 데이터의 복잡성을 줄이고, 분석 및 모델링에서 특정 구간별 통계를 쉽게 파악할 수 있다. 이번 연구에서는 Equal Width Binning(EWB) 방법을 사용하였다. EWB는 연속형 데이터를 일정한 폭으로 나누어 구간을 설정하는 방식이다. 이 방법은 데이터가 고려되지 못하는 한계가 있지만 간단하고 빠른 계산이 가능하며 구현이 간단하다. 또한, 나이브 베이즈 분류기에서 예상보다 성능이 잘 나오는 것으로 자주 사용된다.³³⁾

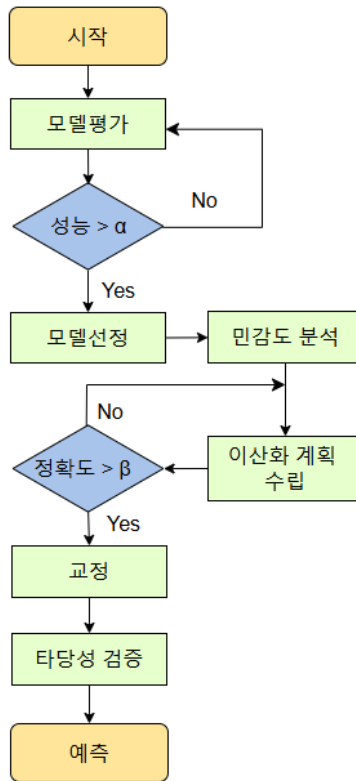
예를 들어, 0점에서 100점까지 분포된 데이터를 EWB 방식으로 3개의 그룹(1그룹: 0~33점, 2그룹: 34~66점, 3그룹: 67~100점)으로 나누면, 다양한 점수대를 간단하게 요약할 수 있다. 이러한 방식은 베이지안 네트워크 모델에서 학습 데이터를 구성할 때도 효과적으로 활용될 수 있다.

6. 연구 범위 및 방법

본 연구에서 부사관의 경우 모든 계급의 진급 데이터가 분석에 사용되었다. 장교의 경우 소령에서 중령으로의 진급 데이터만 사용하였다. 왜냐하면 공군의 경우 대위에서 소령으로의 진급은 특별한 문제가 없는 한 대부분의 인원이 진급하기 때문이다. 중령에서 대령으로의 진급은 데이터 수가 제한되어 제외하였다.

32) James Dougherty, Ron Kohavi, Mehran Sahami. (1995). "Supervised and Unsupervised Discretization of Continuous Features." In Armand Prieditis & Stuart Russell (Eds.), Machine Learning, Proceedings of the Twelfth International Conference, Morgan Kaufmann Publishers, San Francisco, CA.

33) Yang, Y. & Webb, G. I. (2002). A comparative study of discretization methods for naive-Bayes classifiers. Proceedings of the Pacific Rim Knowledge Acquisition Workshop, 159-173.



〈그림 2-8〉 연구 순서도

〈그림 2-8〉은 본 연구의 순서도를 나타낸다. 먼저 기술통계를 통해 데이터의 정성적인 특성을 확인하였다. 다음으로 다양한 인공지능모델의 학습과 타당성 검증을 통해 우수한 신뢰성을 나타내는 모델을 선정하였다. 이후 민감도분석을 통해 진급에 미치는 중요한 변수들을 식별하였다. 일부 모델의 경우 이산화 과정이 수반되며, 이 경우 이산화계획이 큰 영향을 미칠 수 있으므로 이에 대한 민감도 또한 분석하였다. 또한, 데이터 조정하여 재비교한 후 타당성 검증을 실시하였다.

데이터의 경우 복잡도가 상대적으로 높고, 정량적인 지표가 진급에 큰 영향을 미친다고 생각되는 부사관 진급 데이터를 우선적으로 분석하였다. 여기서 결정된 모델과 동일한 모델을 장교 진급 데이터 분석에 사용하였다.

Ⅲ. 부사관 진급 모델 연구

1. 부사관 진급제도 현황

본 연구에서는 공군 장교와 부사관에 대한 자료를 활용하여 연구하였다. 공군 부사관 진급심사는 장병 진급관리 규정을 따르고 있다.³⁴⁾ 부사관 진급평가는 전반적인 과정에는 큰 차이가 없으나 평가항목과 심사 대상에 차이가 있다. 특히 장교의 경우보다 각 항목이 세분화되고 구체적인 평가가 이루어진다는 점이 차이점이라고 할 수 있다.

세부 평가항목을 보면, 근무평정, 군사교육, 보직관리, 자격증, 근속기간, 포상, 지휘추천, 분야심사 위원 점수, 추천심사위원 점수, 선발심사위원 점수, 감점으로 항목이 구성되어 있다. 장교와 차이점을 보자면 주요 경력에 격오지 점수 등 각 점수에 관한 내용이 구체적으로 구분되어 있어 근무 노력에 대한 보상 측면이 반영되어 있으며, 자격증 항목이 추가되어 있어 직무 수행에 필요한 전문성이 중요한 덕목으로 평가된다는 점을 알 수 있다. 이는 장교와 부사관의 역할 구분에 따른 것으로 해석할 수 있는데, 장교의 경우 상대적으로 리더십과 부대 관리 능력이 강조된다면, 부사관은 장교 리더십을 적절히 발휘하기 위한 중간 역할과 전문성 발휘가 중요하기 때문이다. 특히 진급에 따라 역할의 변화가 크지 않기 때문에, 진급의 역할 중 보상의 개념이 강하여 격오지 등 고된 임무를 수행한 사람에 대하여 점수를 더 부여하는 점을 볼 수 있으며, 근속에 대한 부분도 같은 의미에서 높은 점수를 부여한다고 이해할 수 있다.

다면평가는 대상자에 대한 상급자, 동기, 하급자들이 평가한 점수로 진급심사에 평가항목으로 반영되지는 않으며, 심사위원들의 참고 자료로만 활용되는 점수이다.

2. 공군 부사관 기술통계

공군 부사관 데이터에 대한 특성 파악을 위하여 기술통계를 분석하였다. 2019년부터 2023년 사이 공군 부사관 진급 대상자 45,045명에 대한 데이터를 사용하였으며, 부사관 진급 항목에 대한 기술통계는 <표 3-1>과 같다. <그림 3-1>은 배점

34) 공군규정 2-23, (2023) 장병 진급관리

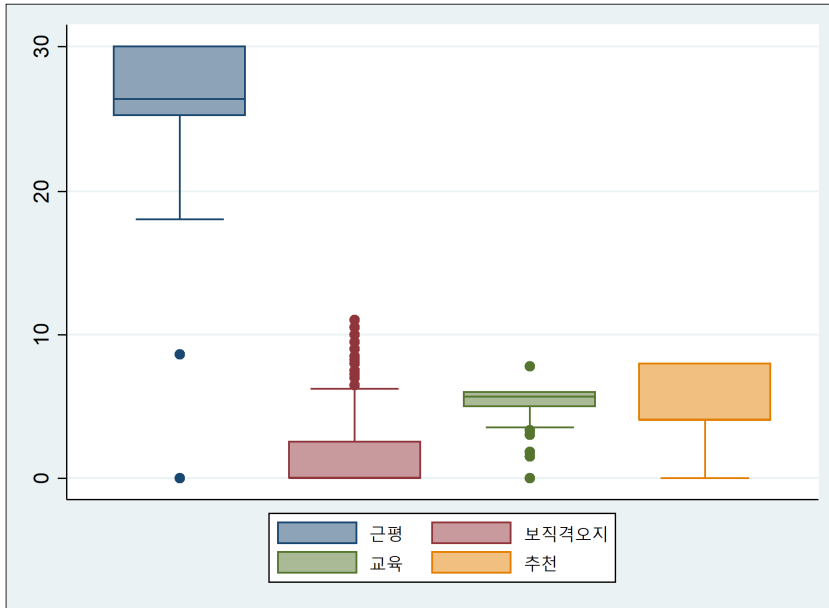
〈표 3-1〉 부사관 진급 항목 기술통계

변수	평균	표준편차	최소값	최대값
근속점수	1.44	2.55	0.00	6.00
상훈	4.21	1.45	0.00	5.00
격오지	1.48	2.11	0.00	11.00
자격증	2.75	1.83	0.00	5.00
웰던창안	0.02	0.18	0.00	4.00
교육	5.42	1.20	0.00	7.80
학위	0.25	0.40	0.00	2.00
해외교육	0.02	0.20	0.00	2.00
근무평정	26.09	5.79	0.00	30.00
지휘추천	4.65	3.12	0.00	8.00
입원	0.00	0.00	-1.00	0.00
처벌	-0.05	0.45	-12.00	0.00
체력검정	4.43	1.03	0.00	5.00
성폭력 별점	0.00	0.01	-2.00	0.00
보안 별점	0.00	0.04	-6.00	0.00
분야점수	0.73	1.33	0.00	5.00
핵심역량 갑	0.73	1.33	0.00	5.00
핵심역량 을	0.73	1.33	0.00	5.00
선발점수	0.84	1.92	0.00	6.00
진급결과	0.16	0.37	0(탈락)	1(합격)
상급자 다면평가*	83.61	96.48	1.00	451.00
동급자 다면평가*	40.86	46.81	1.00	201.00
하급자 다면평가*	48.75	72.39	1.00	322.00

* 다면평가 점수는 참고사항으로 평가점수에 포함되지 않음

이 높은 주요 항목에 대한 평가항목을 상자도표로 나타낸 것이다. 부사관 진급은 장교에 비하여 평가항목이 세분화되어 있어 여러 항목으로 구성되어 있다. 배점이 가장 높은 항목은 근무평정이며, 격오지 근무, 부대 추천, 교육 등이 높은 항목으

로 나타났다. 표준 편차의 경우에는 근무평정, 근속, 부대 추천순으로 큰 것으로 나타나 배점이 클수록 편차가 큰 것으로 나타났다.



〈그림 3-1〉 주요 평가항목 상자도표

〈그림 3-1〉은 배점이 큰 주요 항목에 대한 박스 도표로 사각형으로 표시된 부분은 사분위 범위로, 1분위에서 3분위 사이를 나타낸다. 이는 상위와 하위 25%를 제외한 중간 50%에 대한 점수의 범위를 나타낸다. 장교의 경우보다 박스의 형태가 더 크게 나타나 점수 변별력이 부사관이 더 많다는 것을 알 수 있다. 항목 간 비교를 해보면 배점이 가장 큰 근무평정의 사분위 범위가 크지만, 배점에 비하면 다른 2개의 범위와 큰 차이가 나지 않으며, 교육은 사분위 범위가 다소 좁게 나타나 중간 순위의 점수 차이가 매우 작다는 것을 보여준다.

3. 공군 부사관 데이터 상관관계 분석

평가항목 간의 연관성을 분석하기 위해 상관관계를 분석하였다. 분석도구는 STATA 14를 사용하였다. 상관관계 분석결과는 〈표 3-2〉와 같다. 상관계수가 0.7 이상으로 상관관계가 강한 관계는 분야점수, 핵심역량 갑, 핵심역량 을, 선발점수 간의 관계이다. 이들은 모두 평가위원이 내리는 점수로, 서로 다른 평가위원이라도

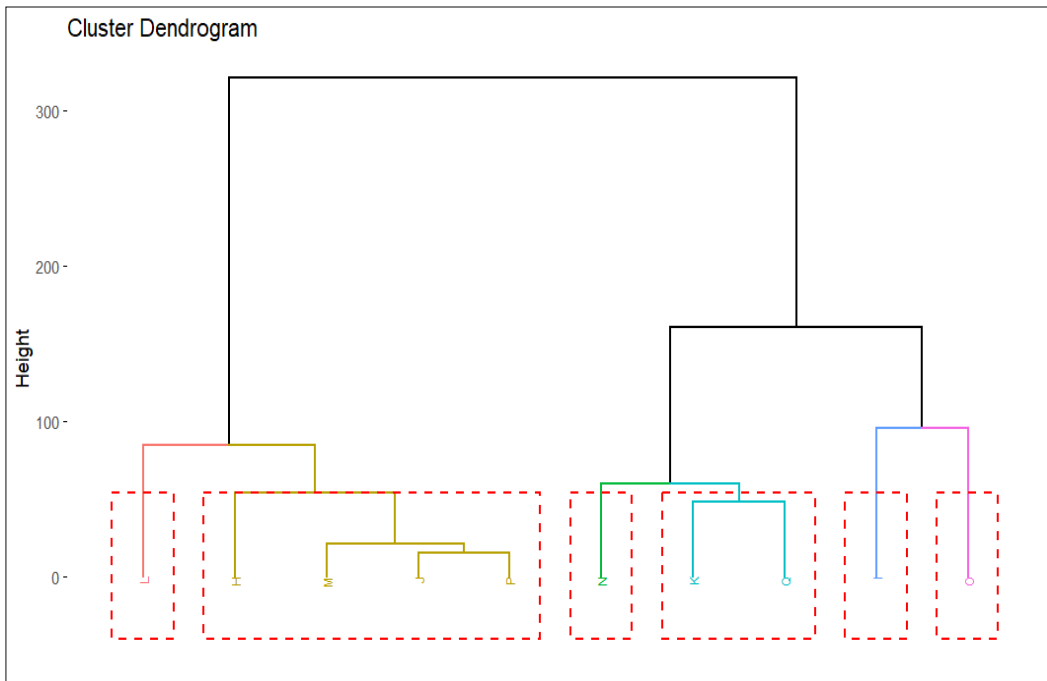
〈표 3-2〉 상관관계 분석결과

변수	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) 근무평정	1													
(2) 보직적응지	0.015*	1												
(3) 지휘추진	0.191*	0.177*	1											
(4) 교육	0.115*	0.082*	0.104*	1										
(5) 근속	-0.051*	0.362*	0.026*	0.066*	1									
(6) 선발점수	0.123*	0.101*	0.444*	0.061*	-0.067*	1								
(7) 상훈	0.312*	0.314*	0.277*	0.162*	0.290*	0.166*	1							
(8) 자격증	0.118*	0.060*	0.277*	0.211*	0.176*	0.197*	0.371*	1						
(9) 체력검정	0.544*	0.148*	0.186*	0.119*	0.111*	0.125*	0.459*	0.268*	1					
(10) 분야점수	0.162*	0.129*	0.490*	0.067*	-0.048*	0.724*	0.189*	0.204*	0.158*	1				
(11) 핵심역량 감	0.158*	0.126*	0.488*	0.063*	-0.048*	0.722*	0.187*	0.200*	0.152*	0.893*	1			
(12) 핵심역량 을	0.157*	0.127*	0.485*	0.064*	-0.048*	0.714*	0.184*	0.200*	0.153*	0.879*	0.897*	1		
(13) 선발점수	0.123*	0.101*	0.444*	0.061*	-0.067*	1.000*	0.166*	0.197*	0.125*	0.724*	0.722*	0.714*	1	
(134) 선발여부	0.123*	0.104*	0.446*	0.061*	-0.060*	0.996*	0.168*	0.199*	0.126*	0.724*	0.722*	0.714*	0.996*	1

* p<0.05

〈그림 3-2〉는 부사관의 산점도 매트릭스이다. 상자도표에서 본 것처럼 근무평정은 대부분 높은 점수대를 형성하고 있으며, 보직 중 격오지 근무는 전반적으로 고르게 분포되어 있지만, 저점대에 인원이 많은 것으로 나타난다. 산점도에서 선형의 형태가 보이는 경우 상관관계가 있다고 할 수 있는데, 제시된 그래프에서 선형 형태를 찾아볼 수 없으므로 자력점수 간 상관관계가 있다고 보기는 어렵다.

추가로 R을 사용하여 그룹화(Clustering)을 통하여 패턴 간의 유사성이 있는지를 분석하였다. 전체 항목 중에서 둘의 패턴의 유사성을 기준으로 묶는 것이기에, 상관관계 분석처럼 직접적인 관계가 있다고 말하기는 어렵다.



〈그림 3-3〉 클러스터링 덴드로그램

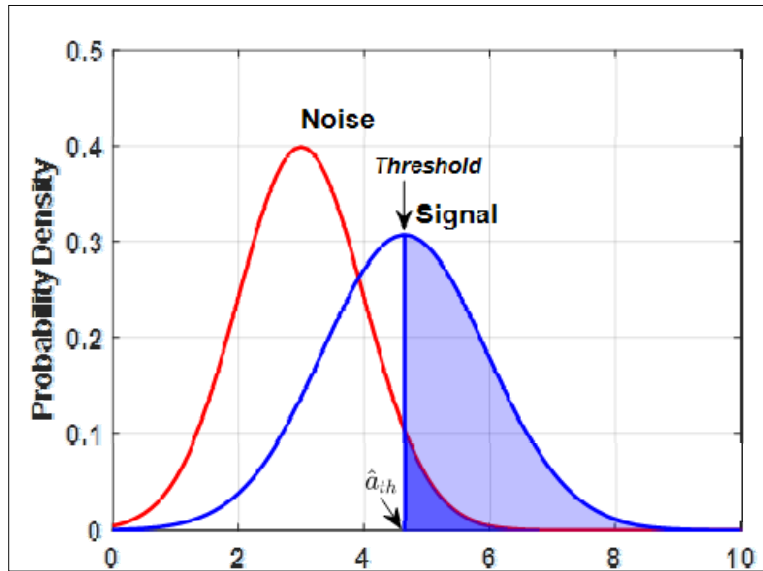
〈그림 3-3〉은 클러스터링 결과를 계층 덴드로그램으로 표현한 그래프이다. 아래 계층을 보면 크게 2개의 묶음이 있고, 나머지는 개별 항목으로 표시가 되었다. 첫 번째 묶음은 근무평정, 처벌, 상훈, 체력검정 점수이다. 두 번째는 가점과 관련된 성과점수와 교육점수이다. 따라서 근무평정, 처벌, 상훈, 체력검정 간에 유사한 패턴이 있고, 군사교육 성적이 좋은 사람은 추가 성과 점수도 올라가는 패턴이 있다고 볼 수 있다. 하지만, 패턴이 유사한 관계 간에 상관관계가 있는지 재확인할 필

요가 있다. <표 3-2>에서 해당항목 간의 상관관계를 보면 다른 항목보다는 약간 높은 수치를 보여주지만 일반적으로 상관관계가 있다고 볼 수 있는 0.6을 넘는 항목은 없다. 상관관계가 높은 경우 클러스터링이 시사하는 바가 크지만, 부사관 진급에서는 상관관계를 재 확인 하였을 때에도 낮게 나왔으므로, 큰 의미를 부여하기는 어렵다. 따라서 부사관 평가항목 간 상관관계 수준이 낮으나, 그중에서 조금 이나마 유사성을 찾자면 언급된 2개의 그룹 내의 항목이 비슷하다고 할 수 있다.

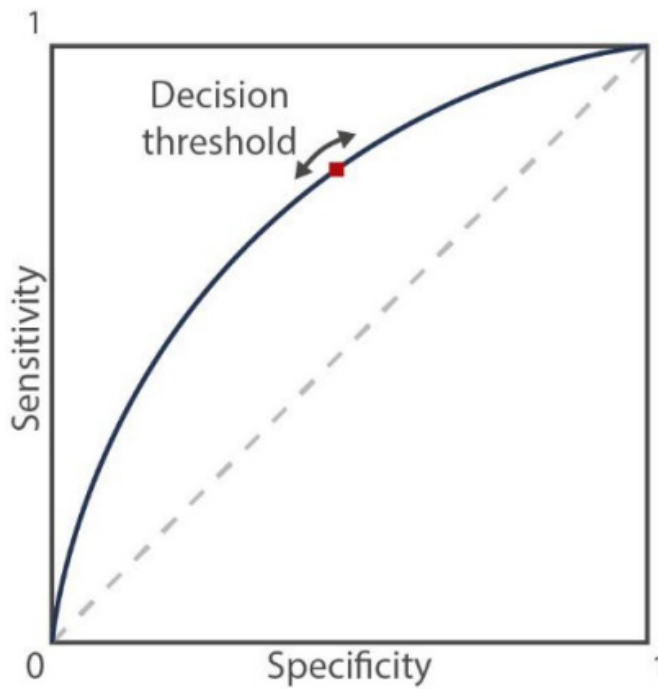
4. 공군 부사관 머신러닝 모델 비교

이번 장에서는 다면평가를 추가한 모델을 기준으로 다양한 머신러닝 알고리즘의 성능을 AUC(Area Under the Curve)로 비교하여, 베이지안 네트워크 사용의 적절성을 평가하였다. AUC는 ROC 곡선 아래의 면적을 나타내며, 모델이 이진 분류 문제에서 얼마나 정확하게 작동하는지를 측정하는 중요한 지표이다.

ROC 곡선이란 문턱값에 따른 모델의 신뢰도를 말한다. <그림 3-4>는 ROC 곡선 개념을 나타내고 있다. ROC 곡선은 성공과 실패의 판정(Threshold) 유의수준에 따른 정확도를 보여주는 그래프이다. 시스템이 연속적인 출력값을 도출할 때, 신호와 잡음을 제대로 구분하거나 그렇지 않은 경우는 4가지이다. 이 중 참은 신호가 문턱값을 넘을 때 신호로 판정한 경우와 문턱값 미만인 잡음을 잡음으로 판정한 경우이다. 오류는 문턱값을 넘지 못하여 잡음으로 판정해야 하는데, 신호로 잘못 판정하는 경우와 문턱값을 넘는 신호를 잡음으로 잘못 판정하는 경우이다. 4개의 비율은 문턱값 설정값에 따라 달라진다. <그림 3-4>에서 볼 수 있듯이 오류를 줄이려고 문턱값을 높이면 신호를 탐지할 확률이 감소한다. 반면 문턱값을 낮추면 오류의 확률이 증가한다. 특히, 거짓을 참으로 판정하는 경우를 지칭하는 1종 오류(False Positive)는 분류기(Classifier)의 신뢰성을 크게 떨어뜨린다.



〈그림 3-4〉 ROC 곡선 문턱값(Threshold) 개념도



〈그림 3-5〉 ROC 곡선

〈그림 3-5〉는 문턱값에 따른 1종오류에 대한 신뢰성을 나타낸 ROC 곡선이다. 곡선이 좌측 상단에 가까울수록 신뢰도가 높은 우수한 모형이라고 할 수 있다. AUC는 곡선의 하단 부분의 크기를 말한다. AUC 값이 1에 가까울수록 모델의 성능이 우수함을 의미하고, 그림에 점선과 같이 0.5인 경우는 무작위 예측과 동일한 성능을 나타낸다. 0.5 미만이라면 무작위 예측보다 낮으므로 사용하지 않는 것이 더 적절하다.

사용된 머신러닝 알고리즘은 다음과 같다. 먼저, 결정 트리(Decision Tree)는 데이터를 기준에 따라 분할하여 예측을 수행하는 트리 구조의 모델이다. 그다음, 앙상블 기법(Ensemble Methods)은 여러 모델을 결합해 예측 성능을 높이는 기법으로, 대표적으로 Bagged Trees와 Boosted Trees가 있다. SVM(Support Vector Machine, 서포트벡터머신) 서포트벡터와 마진을 통해 데이터를 분류하며, 다양한 커널 함수가 적용될 수 있는 강력한 분류 모델이다. 신경망(Neural Networks)은 여러 층의 뉴런을 통해 데이터를 학습하며, 비선형 패턴을 처리하는데 매우 뛰어난 성능을 발휘한다. 마지막으로, 선형 회귀(Linear Regression)는 입력과 출력 간의 선형 관계를 모델링하는 기법으로, 가장 기본적인 회귀분석 방법이다.

보다 세부적인 모델 적합도를 측정하기 위하여 각 연도별 데이터를 구분하여 데이터 머신러닝 AUC 분석을 실시하고 예측 성능을 비교하였다.

4.1.1. 2019년 데이터 머신러닝 AUC 분석결과

Bayesian Network(0.87)는 가장 우수한 예측 성능을 보였으며, Neural Networks나 GPR과 같은 복잡한 모델보다 나은 성능을 발휘했다. 이는 변수 간의 인과관계를 명시적으로 모델링함으로써 데이터의 구조적 특성을 잘 반영했기 때문으로 분석된다.

베이지안 네트워크는 복잡한 비선형 패턴을 학습하는 신경망에 비해 더 간결한 구조로도 우수한 예측 성능을 보였다.

〈표 3-3〉 2019년 모델별 AUC

Model	AUC	Model	AUC
Bayesian Network	0.87	Narrow Neural Network	0.81
Medium Neural Network	0.86	Boosted Trees	0.70
Rational Quadratic GPR	0.86	Bagged Trees	0.68
Exponential GPR	0.86	SVM Quadratic	0.68
Matern 5/2 GPR	0.85	SVM Cubic	0.68
Wide Neural Network	0.85	Fine Tree	0.67
Squared Exponential GPR	0.85	SVM Coarse Gaussian	0.62
Robust Linear	0.84	SVM Medium Gaussian	0.62
Stepwise Linear Regression	0.84	Kernel SVM	0.62
Linear Regression	0.84	Fine Gaussian SVM	0.62
Interactions Linear Regression	0.84	SVM Linear	0.60
Bi-layered Neural Network	0.83	Medium Tree	0.58
Tri-layered Neural Network	0.81	Coarse Tree	0.50

4.1.2. 2020년 데이터 머신러닝 AUC 분석결과

2019년과 마찬가지로 Bayesian Network(0.88)는 모델 중 최고의 성능을 보였다. Linear Regression(0.86) 및 Bi-layered Neural Network(0.85) 모델도 높은 수준의 예측 성능을 발휘했다.

〈표 3-4〉 2019년 모델별 AUC

Model	AUC	Model	AUC
Bayesian Network	0.88	Fine Tree	0.68
Linear Regression	0.86	SVM Linear	0.66
Interactions Linear Regression	0.86	Boosted Trees	0.63
Bi-layered Neural Network	0.85	Bagged Trees	0.62
Medium Neural Network	0.85	SVM Quadratic	0.62
Narrow Neural Network	0.84	SVM Cubic	0.61
Exponential GPR	0.84	Coarse Tree	0.61
Matern 5/2 GPR	0.83	Medium Tree	0.61
Rational Quadratic GPR	0.83	Kernel SVM	0.59
Squared Exponential GPR	0.83	SVM Medium Gaussian	0.59
Wide Neural Network	0.83	Fine Gaussian SVM	0.59
Tri-layered Neural Network	0.83	SVM Coarse Gaussian	0.58
Stepwise Linear Regression	0.82	Robust Linear	0.50

4.1.3. 2021년 데이터 머신러닝 AUC 분석결과

Bayesian Network(0.88)는 2021년 데이터에서도 가장 높은 AUC를 나타내었으며, Linear Regression, Neural Networks 및 GPR 모델보다 나은 예측 능력을 보였다.

〈표 3-5〉 2021년 모델별 AUC

Model	AUC	Model	AUC
Bayesian Network	0.88	Medium Tree	0.65
Linear Regression	0.85	Fine Tree	0.64
Interactions Linear Regression	0.85	SVM Cubic	0.64
Narrow Neural Network	0.84	Boosted Trees	0.62
Exponential GPR	0.84	Coarse Tree	0.61
Matern 5/2 GPR	0.84	SVM Quadratic	0.60
Tri-layered Neural Network	0.84	SVM Linear	0.57
Rational Quadratic GPR	0.84	Bagged Trees	0.56
Squared Exponential GPR	0.84	SVM Medium Gaussian	0.53
Stepwise Linear Regression	0.83	Kernel SVM	0.53
Medium Neural Network	0.82	SVM Coarse Gaussian	0.53
Wide Neural Network	0.81	Fine Gaussian SVM	0.53
Bi-layered Neural Network	0.79	Robust Linear	0.28

4.1.4. 2022년 데이터 머신러닝 AUC 분석결과

Bayesian Network(0.89)는 GPR, Neural Networks 모델에 비해 소폭 떨어지지만, 여전히 매우 강력한 예측 성능을 보였다. 이 데이터에서는 Rational Quadratic GPR(0.91), Exponential GPR(0.91)과 같은 Gaussian Process Regression 모델들이 특히 강력한 성능을 보였지만, BN은 여전히 데이터의 구조적 관계를 모델링하여 안정적인 성능을 유지하였다.

〈표 3-6〉 2022년 모델별 AUC

Model	AUC	Model	AUC
Rational Quadratic GPR	0.91	Boosted Trees	0.76
Exponential GPR	0.91	Fine Tree	0.73
Bi-layered Neural Network	0.91	SVM Cubic	0.73
Wide Neural Network	0.91	Bagged Trees	0.70
Linear Regression	0.91	SVM Quadratic	0.69
Interactions Linear Regression	0.91	SVM Medium Gaussian	0.66
Narrow Neural Network	0.90	Kernel SVM	0.65
Tri-layered Neural Network	0.90	Fine Gaussian SVM	0.65
Stepwise Linear Regression	0.90	SVM Coarse Gaussian	0.65
Medium Neural Network	0.90	Medium Tree	0.59
Matern 5/2 GPR	0.90	SVM Linear	0.53
Bayesian Network	0.89	Coarse Tree	0.50
Squared Exponential GPR	0.88	Robust Linear	0.50

4.1.5. 2023년 데이터 머신러닝 AUC 분석결과

Bayesian Network(0.88)는 2023년에도 우수한 성능을 보였으나, Neural Networks(0.90)나 Linear Regression(0.90) 모델들에 비해 예측 성능이 소폭 감소했다. 이는 데이터 복잡성이 증가함에 따라 Bayesian Network가 복잡한 비선형 패턴을 처리하는 데 한계가 있을 수 있음을 시사한다. 그러나 Bayesian Network는 여전히 높은 수준의 예측 성능을 보이고 있으며, 변수 간의 관계를 명확히 설명할 수 있는 능력이 중요한 상황에서는 여전히 경쟁력 있는 모델이라 할 수 있다.

〈표 3-7〉 2023년 모델별 AUC

Model	AUC	Model	AUC
Medium Neural Network	0.90	Fine Tree	0.67
Linear Regression	0.90	SVM Cubic	0.65
Interactions Linear Regression	0.90	SVM Quadratic	0.65
Tri-layered Neural Network	0.90	Boosted Trees	0.64
Wide Neural Network	0.90	Bagged Trees	0.62
Bi-layered Neural Network	0.90	Medium Tree	0.62
Exponential GPR	0.90	SVM Coarse Gaussian	0.59
Rational Quadratic GPR	0.90	SVM Medium Gaussian	0.59
Matern 5/2 GPR	0.90	Fine Gaussian SVM	0.59
Narrow Neural Network	0.89	Kernel SVM	0.59
Squared Exponential GPR	0.89	Coarse Tree	0.50
Stepwise Linear Regression	0.89	SVM Linear	0.50
Bayesian Network	0.88	Robust Linear	0.50

4.1.5. 머신러닝 AUC 분석결과 고찰

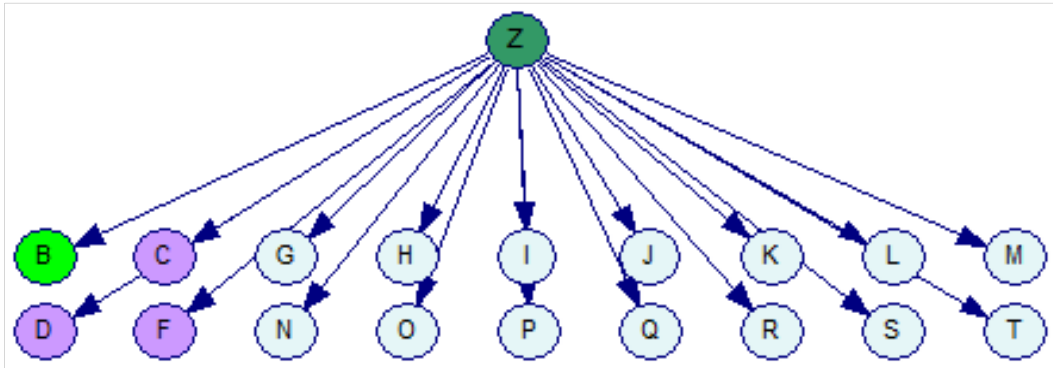
각 연도별 데이터를 사용하여 다양한 머신러닝 모델을 학습한 후 ROC 곡선의 AUC를 계산한 결과 베이지안 네트워크 모델을 포함한 여러 모델이 0.90에 가까운 값을 나타내었다. GPR 모델 또한 우수한 신뢰도를 나타내나 고차원 데이터를 GPR로 표현하는 데 있어 발생하는 여러 문제가 있으므로 본 연구에는 적합하지 않은 것으로 판단된다.³⁵⁾ 인공신경망 모델이 경우 인과 관계를 설명하는 데 어려움이 있으므로 고려하지 않았다. 베이지안 네트워크 모델의 경우 인자 사이의 연관관계를 잘 나타내고 있으며, 복잡한 조건부 확률을 계산하는 데 큰 장점이 있으므로 본 연구의 모델로 채택하였다. 또한, 나이브 베이스 모델의 경우는 각 노드의

35) Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). Global Sensitivity Analysis: The Primer. John Wiley & Sons.

부모 노드가 하나로 제한되므로 추론해야 할 파라미터의 숫자가 선형적으로 증가하므로 계산이 다항시간(Polynomial Time) 안에 수행된다.

5. 기본 모델링

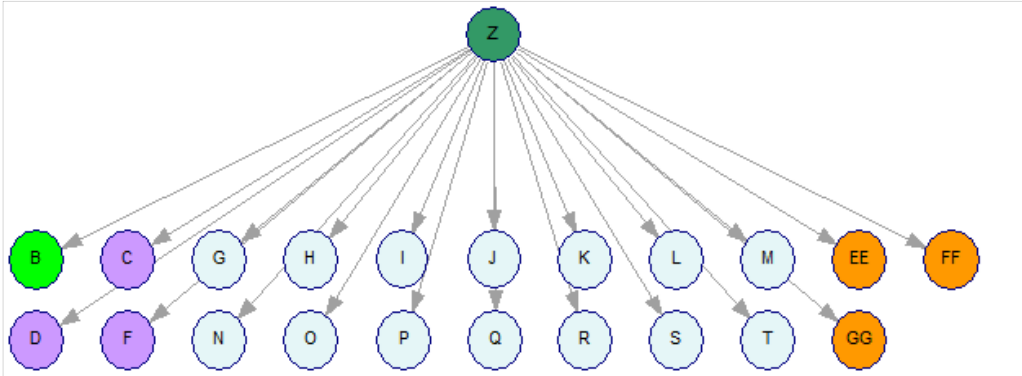
공군 부사관 진급 대상자 데이터를 사용하여 나이브 베이즈 모델 형태로 공군 부사관 진급 모델을 작성하였다. 나이브 베이즈 모델은 베이지안 확률론을 바탕으로 각 노드 간의 관계는 독립이라는 가정하에 주어진 노드에 대한 조건부 확률을 계산할 수 있다. 진급평가가 각각의 항목이 직접 평가점수에 독립적으로 반영되므로, 나이브 베이즈 모델로 작성하였으며, 부사관의 각 평가항목에 맞게 구성하였다.



〈그림 3-6〉 부사관 진급 모델

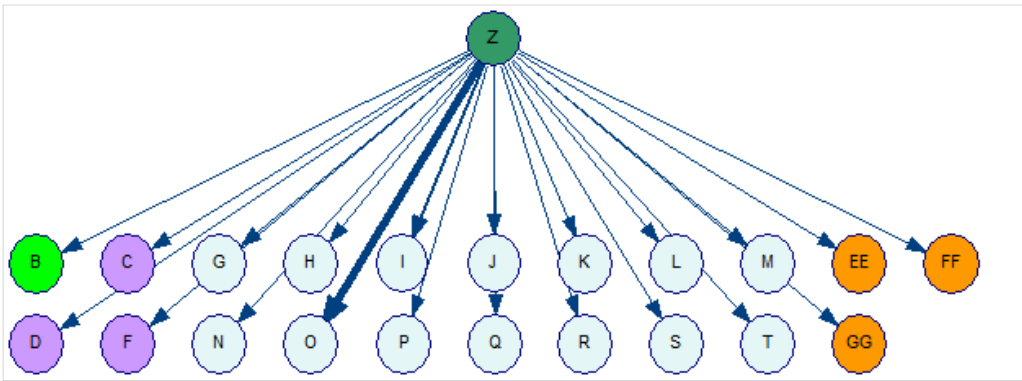
〈그림 3-6〉은 부사관 진급 모델을 나이브 진급 모델로 작성한 결과이다. 부사관의 평가에 맞는 항목으로 구성되었다. Z는 진급 결과로 1은 선발, 0은 비선발을 의미한다. 하부에 여러 자식 노드가 있는데, B부터 T로 나타난 자식 노드들은 진급평가항목과 진급에 영향을 줄 수 있는 진급연도, 특기 등의 요소로 구성되었다.

부사관 심사에서 다면평가를 수행하고 있지만 평가항목으로 활용하고 있지는 않고 있다. 부사관 심사에서 다면평가의 영향력을 살펴보기 위하여 〈그림 3-7〉과 같이 다면평가 추가한 항목을 추가하였다.



〈그림 3-7〉 부서관 진급 모델 수정(다면평가 추가)

6. 민감도 분석결과



〈그림 3-8〉 민감도 분석

〈그림 3-8〉에서 굵게 표시된 항목은 진급추천, 자격증, 상훈, 학위이다. 해당 항목이 진급에 큰 영향을 미치는 것으로 나타났다.

먼저 진급추천을 살펴보면, 추천 대상자를 3그룹으로 나뉘며 첫 번째 그룹의 배점이 가장 크며, 진급에 가장 큰 영향을 미치는 것으로 확인하였다. 두 번째, 자격증은 필수 자격증, 특수자격 요건에 따라 점수가 부여되며 영향력이 강하게 나타났다. 세 번째, 상훈은 해계급 내 최고 훈격의 배점이 적용되고, 진급 예측도를 높이는데 기여도가 큰 것으로 확인되었다. 네 번째, 학위 점수는 전문학사부터 박사까지 최상위급 취득 학위 1개만을 인정하고, 이 또한 진급 예측도를 높인다. 이러한 항목이 진급 예측률에 영향이 크다.

이러한 결과는 배점의 기준과 실제 진급률에 주는 영향에 차이가 있다는 점을

알려준다. 배점은 근무평정이 가장 큰 비중을 차지하지만, 비중이 낮은 자격증, 학위 등이 오히려 영향력이 큰 것으로 나타났다. 따라서 배점 기준에 따른 중요도가 잘 반영되고 있지 않다는 점을 알 수 있다.

〈표 3-8〉 연도별 민감도 분석결과

민감도 순위	연도					
	19	20	21	22	23	통합
1	추천	추천	추천	추천	추천	추천
2	자격증 어학	자격증 어학	자격증 어학	상훈	자격증 어학	자격증 어학
3	상훈	상훈	상훈	자격증 어학	학위	상훈
4	다면하	학위	신검	근속	상훈	학위
5	다면중	진급년도	근속	신검	근평	근속
6	다면상	근속	학위	진급년도	진급년도	진급년도
7	보직 (격오지)	다면중	진급년도	보직격오지	보직격오지	신검
8	학위	다면상	보직 (격오지)	근평	다면중	보직 (격오지)
9	진급년도	신검	근평	다면상	교육	근평
10	근속	다면하	다면중	학위	다면상	심사년도

〈표 3-8〉은 민감도 순위를 1위부터 10위까지 나타낸 것이다. 매년 진급추천이 가장 민감한 요소로 확인되었으며, 자격증과 상훈이 각각 2위와 3위로 높은 영향도를 보였다. 반면, 가장 높은 배점을 가진 근무평정은 비교적 낮은 순위에 위치해, 배점에 비해 실제 영향력은 다소 낮다고 해석할 수 있다. 다면평가 또한 상대적으로 낮은 영향을 미친 것으로 나타났다.

7. 다면평가 추가모델 타당성 분석결과

7.1. 다면평가 추가모델 타당성 분석

진급 대상자의 데이터를 활용하여 K-fold Cross Validation을 통해 모델의 타당성을 분석하였다. <표 3-9>는 진급확률 0.5를 기준으로 하였을 때 결과이다. 즉, 진급확률이 0.5 이상이면 진급했다고 판단할 경우 실제 진급 결과와 비교하여 일치하는가를 판단하였다. 기본 모델은 현재 평가항목만을 기준으로 한 정확도이고, 다면평가 추가모델은 다면평가를 평가항목으로 추가한 정확도이다.

<표 3-9> 부서관 진급 예측도 비교

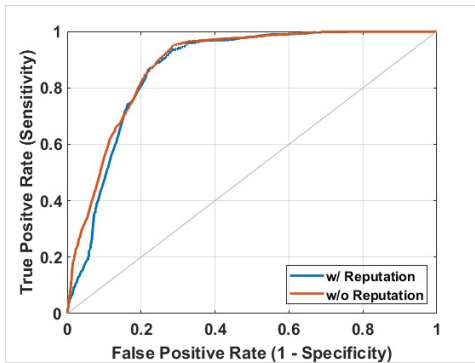
기본 모델		예측	
		비선발	선발
실제	비선발	33,227	4,313
	선발	2,833	4,672
다면평가 추가모델		예측	
		선발	비선발
실제	비선발	32,727	4,813
	선발	2,628	4,877

5개년간의 진급 대상자 데이터를 분석한 결과, 기본 모델의 예측 정확도는 84.1%(37,899/45,045)였으나, 다면평가를 추가한 모델의 예측 정확도는 83.4%(37,604/45,045)로 소폭 감소하였다. 이는 다면평가를 추가했을 때 오히려 진급 예측 능력이 감소했음을 의미한다. <표 3-10>은 연도별 정확도를 비교한 것으로 다면평가 연도에 따라 기본모델이 더 정확한 경우도 있고 그렇지 않은 경우도 있다. 연도별로 분석한 결과를 보면 비교적 최근 연도로 갈수록 다면평가가 더 우세한 경향이 보인다.

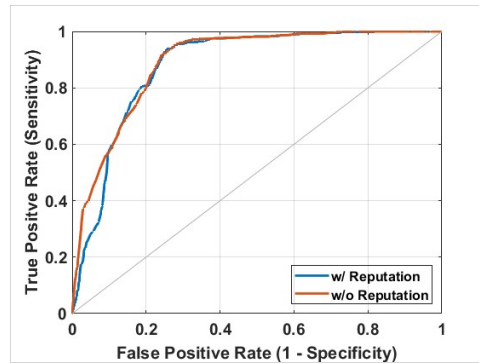
〈표 3-10〉 연도별 정확도 비교

진급연도(년)	기본모델	다면평가 추가
19	85.3%	83.7%
20	85.3%	84.5%
21	85.3%	83.4%
22	82.9%	84.5%
23	81.7%	82.2%

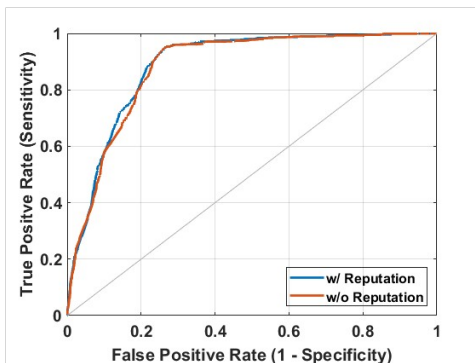
2020년부터 2023년까지 다면평가 유무에 따른 ROC 곡선을 그려보았다. 결과는 아래 그림과 같으며, 붉은 색이 기본모델(다면평가 제외), 푸른색이 다면평가 추가모델의 ROC 커브이다. 커브의 아래 면적이 클수록 우수한 모델이다.



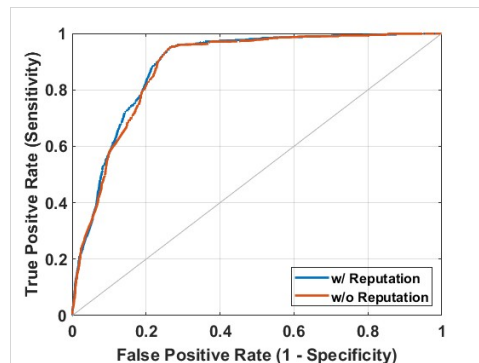
〈그림 3-9〉 2019년도 ROC



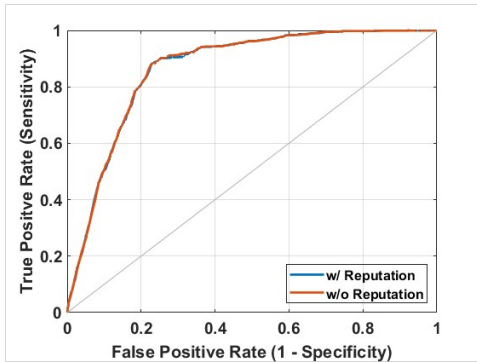
〈그림 3-10〉 2020년도 ROC



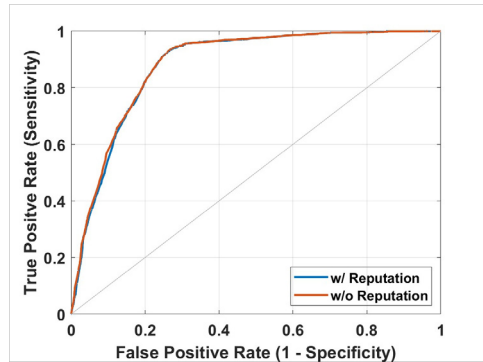
〈그림 3-11〉 2021년도 ROC



〈그림 3-12〉 2022년도 ROC



〈그림 3-13〉 2023년도 ROC

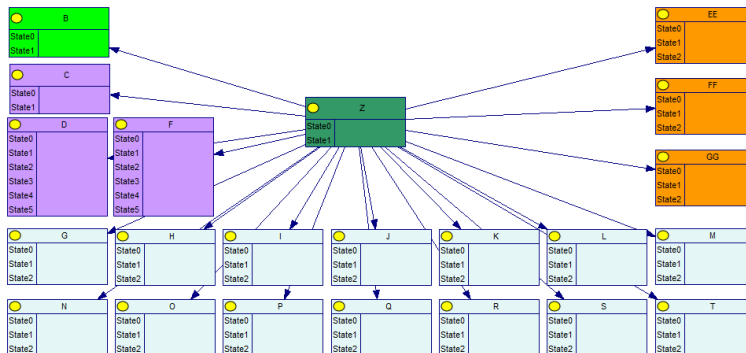


〈그림 3-14〉 통합년도 ROC

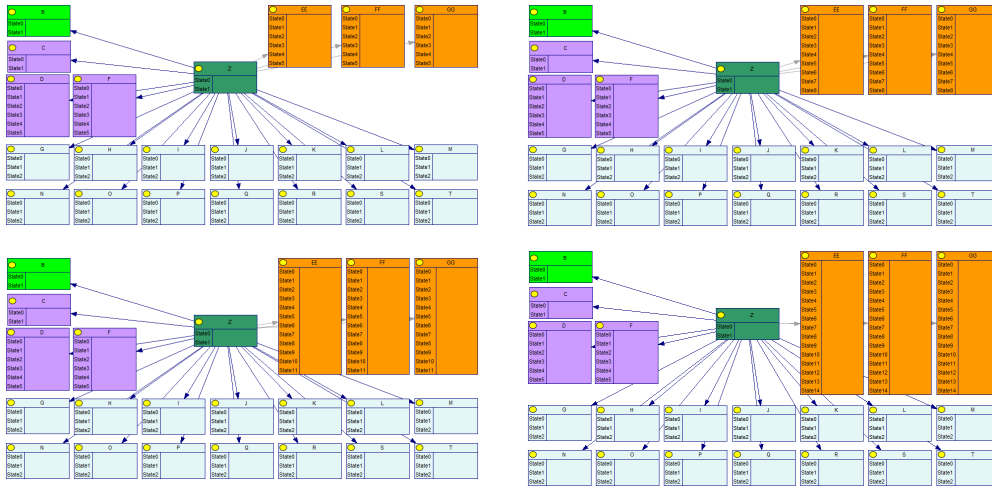
다면평가 적용 여부에 따른 ROC 곡선을 분석한 결과, 다면평가가 포함된 모델은 정확도가 약간 낮아지거나 거의 동일한 성능을 보였다. 이는 다면평가 요소를 추가하는 것이 진급 예측 모델의 성능을 향상시키기에는 한계가 있다는 것을 시사한다.

7.2 이산화계획의 효과

다면평가의 정확도가 떨어지는 이유가 이산화계획에 따른 것인지 확인하기 위하여 다면평가의 이산화계획을 다각화 하였다. 기존 모델을 막대 차트로 표현한 것이 〈그림 3-15〉이다. 알파벳 B부터 T는 진급심사에 포함되는 다양한 요소를 나타내며, EE, FF, GG는 참고 요소인 다면평가항목이다. 추가적인 이산화 과정을 통해 다면평가의 이산화 수준을 6, 9, 12, 15 단계로 확장한 결과는 〈그림 3-16〉에 제시되어 있다.



〈그림 3-15〉 다면평가 노드 진급 모델 Bar chart (3수준)



〈그림 3-16〉 이산화(6, 9, 12, 15수준)별 진급 모델 Bar chart

2019년부터 2023년도까지 연도별 진급 대상자의 데이터를 각각 학습한 후 정확도를 계산한 결과는 〈표 3-11〉과 같다. 이산화 후 각 모델에 다면평가 데이터를 포함하여 학습한 후 정확도를 확인하였다. 이산화 후 대부분의 경우, 2022년도를 제외하고 정확도가 향상된 것을 확인할 수 있다.

〈표 3-11〉 연도별 다면평가 이산화 수준에 따른 모델 정확도

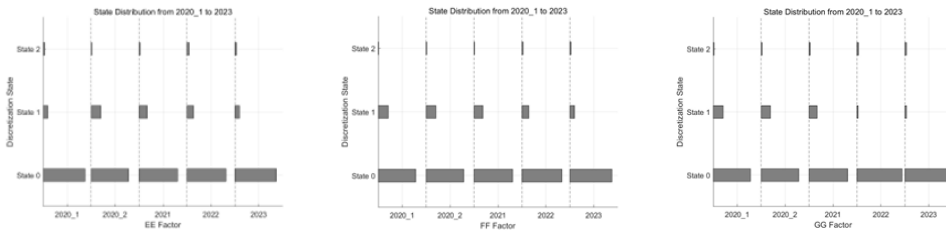
진급연도(년)	이산화 수준(States) 및 확률(%)				
	3	6	9	12	15
19	83.7	84.02	83.94	84.06	84.01
20	84.51	84.74	85.00	85.01	85.00
21	83.26	84.30	84.50	84.18	84.23
22	84.71	84.47	84.39	84.38	84.61
23	82.23	83.53	84.29	84.20	84.45
통합	83.59	83.92	84.11	84.12	84.19

2019년과 2020년에는 12수준일 경우 각각 84.06%, 85.01%, 2021년도는 9 수준일 경우 84.5%, 23년도는 15수준일 경우 84.45%로 가장 높았다. 2022년도를 제외하곤 9에서 15수준까지의 범위에서 가장 높은 정확도를 확인할 수 있다.

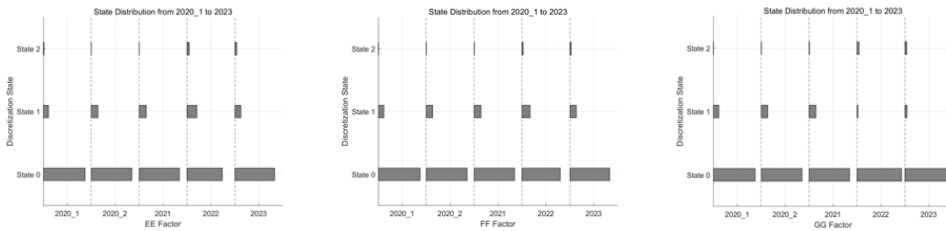
이산화계획의 변화에 따른 정확도의 최대, 최소값의 차이는 2019년 0.36%, 2020년 0.5%, 2021년 1.24%, 2022년 0.33%, 2023년 2.22%로, 2022년을 제외하고는 그 차이가 매년 커지고 있지만 큰 차이는 아닌 것으로 판단된다.

이산화계획의 변화에 따라 정확도 차이가 큰 변화를 보이지 않는 이유는 다면평가 실시율이 저조하기 때문일 수 있다. 실재 2019년부터 2023년까지 다면평가를 실시한 비율은 진급대상자의 24% 수준이다. 분석과정에서 다면평가를 미실시한 경우 데이터를 '없음(NaN)'으로 반영하고, 이는 학습 시 가장 낮은 수준인 State = 0으로 처리되었다.

3수준 기준으로 다면평가 이산화계획은 최고점과 최저점을 기준으로 3개의 영역으로 나누어, 가장 낮은 점수부터 1/3 지점까지는 State 0, 1/3 지점과 2/3 지점 사이는 State 1, 2/3 지점부터 최고득점까지 State 2로 분류된다.



〈그림 3-17〉 연도별 다면평가 수준별 분포 (진급자)



〈그림 3-18〉 연도별 다면평가 수준별 분포 (미진급자)

〈그림 3-17〉 및 〈그림 3-18〉은 연도별 진급 여부에 따른 다면평가 점수를 3개의 수준으로 분류한 그래프이다. 2019년 진급대상자 8,660명 중 다면평가 점수가

없는 인원은 7,429명으로, 이는 전체 대상자의 85.8%에 해당한다. 진급 여부에 관계없이 다면평가 점수가 없는 인원이 압도적으로 많다.

〈표 3-12〉 연도별 진급대상자 중 다면평가 현황

진급연도	진급대상(명)	미실시자(명)	실시자(명)	미실시율(%)	실시율(%)
19	8,660	7,429	1,231	85.8	14.2
20	8,719	7,322	1,397	83.9	16.1
21	8,947	7,484	1,463	83.6	16.3
22	9,222	6,342	2,880	68.8	31.2
23	9,497	6,895	2,602	72.6	27.4

〈표 3-12〉에는 연도별 진급대상자, 다면평가 미실시자, 다면평가 실시자의 데이터가 있으며, 4열과 5열은 연도별 진급대상자 대비 다면평가 미실시자 및 실시자 비율을 나타낸다. 연도별 다면평가 미실시 인원은 2019년 7,429명(85.8%), 2020년 7,322명(83.9%), 2021년 7,484명(83.6%), 2022년 6,342명(68.8%), 2023년 6,895명(72.6%)이며 매년 미실시자 비율은 줄어드는 경향을 보인다. 다면평가 실시자는 2019년 1,231명(14.2%), 2020년 1,397명(16.1%), 2021년 1,463명(16.3%), 2022년 2,880명(31.2%), 2023년 2,602명(27.4%)으로 증가하였다.

다면평가 실시 비율은 매년 증가하고 있음을 확인할 수 있다. 이는 다면평가의 중요성이 강조되면서 평가 실시 비율도 높아지고 있는 것으로 생각된다. 다면평가 실시자 중 미진급자와 진급자의 비율은 〈표 3-13〉과 같다.

〈표 3-13〉 연도별 다면평가 실시자 중 진급 현황

진급연도	다면평가 실시(명)	미진급자(명)	진급자(명)	미진급 비율(%)	진급비율(%)
19	1,231	934	297	75.8	24
20	1,397	1,088	309	77.8	22
21	1,463	1,185	278	81	19
22	2,880	2,542	338	88	11
23	2,602	2,257	345	86	13

다면평가 실시자 중 미진급자 비율은 2019년에 75.8%로 가장 낮았으나, 2021년부터는 80% 이상으로 상승하였다. 진급자는 2019년에 24%로 가장 높았지만, 2023년까지 지속적으로 감소하는 경향을 보인다.

이와 같이 다면평가 자의 미실시율이 기본모델과 다면평가 추가 모델의 비교하는데 걸림돌이 될 수 있다고 판단하였기에, 연도별 다면평가 점수가 반영된 데이터만을 추출하여 이산화 수준에 따른 정확도를 재측정하였다.

〈표 3-14〉 다면평가 실시 데이터 학습 후 정확도

진급연도	이산화 수준(States)에 따른 정확도(%)				
	3	6	9	12	15
19	77.1	77.3	77.3	76	77
20	77	77.4	77.2	77.4	77.1
21	81.7	81.1	81.2	80.5	80.7
22	88.88	88.05	88.2	87	88
23	87.5	87.2	87.7	87	87.6
통합	82.77	83	88.75	82.65	82.71

다면평가 이산화 계획에 따라 3수준일 때 77.1%, 6 및 9수준에서는 77.3%, 12수준은 76%, 15수준에서는 77%였다. 이산화 계획의 변화에 따라 정확도는 대략 76~77% 수준으로 큰 변화는 없었다. 2020년은 77%, 21년에는 81.7%, 2022년에는 88.88%, 2023년에는 87.5%를 정확도를 보인다. 다면평가 점수를 3에서 15수준까지 이산화한 후에도 정확도의 상승 폭은 크지 않았다. 5개년 통합 시에도 이산화에 따른 변화는 거의 없었다. 따라서 이산화 계획은 모델의 정확도에 큰 영향을 미치지 않는 것으로 판단된다.

다만 다면평가를 실시한 인원의 데이터를 학습했을 경우 정확도는 크게 상승하는 것을 볼 수 있다. 기존 전체 인원의 데이터가 반영되었던 결과와 비교하면 큰 차이가 있다. 그 결과는 〈표 3-15〉와 같다.

〈표 3-15〉 전체 및 일부(다면평가 실시 데이터) 반영 후 정확도 비교

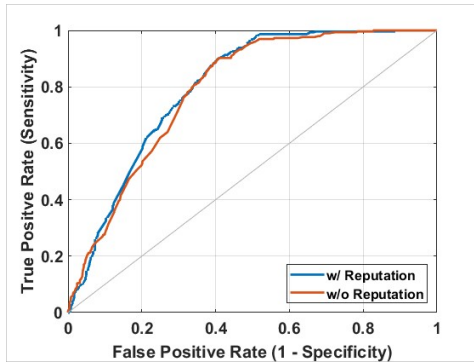
진급연도	이산화 수준(States)에 따른 정확도(%)									
	3		6		9		12		15	
	전체	일부	전체	일부	전체	일부	전체	일부	전체	일부
19	83.7	77.1	84.02	77.3	84.02	77.3	84.06	76	84.01	77
20	84.51	77	84.74	77.4	84.74	77.2	85.01	77.4	85.00	77.1
21	83.26	81.7	84.30	81.1	84.30	81.2	84.18	80.5	84.23	80.7
22	84.71	88.88	84.47	88.05	84.47	88.2	84.38	87	84.61	88
23	82.23	87.5	83.53	87.2	83.53	87.7	84.20	87	84.45	87.6
통합	83.59	82.77	83.92	83	84.11	88.75	84.12	82.65	84.19	82.71

〈표 3-15〉는 전체 데이터를 학습한 경우(표에서 ‘전체’로 표현)와 다면평가를 실시한 데이터만을 학습한 경우(표에서 ‘일부’로 표현)의 정확도이다.

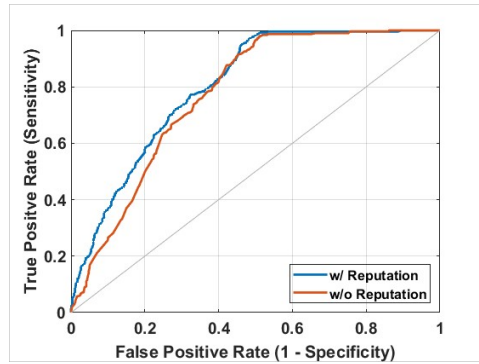
2019년에는 다면평가 실시자의 데이터만 가지고 정확도를 검증한 결과, 기존 83.7%에서 77.1%로 떨어졌다. 6, 9, 12, 15수준도 약 83%에서 77%대로 유사한 수준으로 떨어지는 결과를 보였다. 2020년은 모든 수준이 유사하게 약 84%에서 77%까지, 2021년도는 약 83%에서 약 81%까지 떨어졌다. 2022년에는 모든 수준의 정확도가 기존 약 84%에서 88%까지 상승하였다. 2023년도 또한 전 수준이 약 82%에서 87%까지 상승하였다. 2021년도부터 다면평가가 반영된 데이터로 학습한 경우 정확도가 80%를 넘게 되고 2022년도부터 기존 데이터와 비교 시 정확도가 더 높아지는 것을 확인할 수 있다.

따라서 2022년부터의 다면평가 데이터가 진급 대상자를 예측하는 모델에 큰 영향을 미치는 것으로 판단할 수 있다. 이는 향후 진급평가 예측에서 다면평가항목이 고려될 가능성을 시사한다.

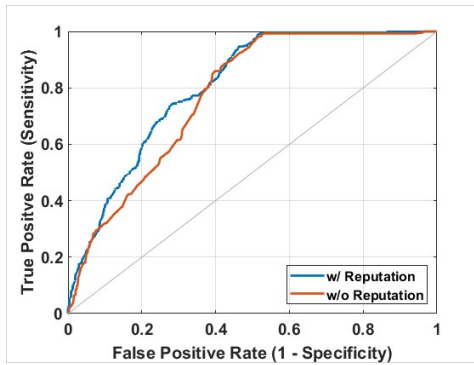
다면평가 요소의 포함 여부에 따른 ROC 곡선은 아래 그림에 나타나 있다. 이를 통해 다면평가를 포함했을 때와 포함하지 않았을 때의 모델 성능 차이를 비교할 수 있다.



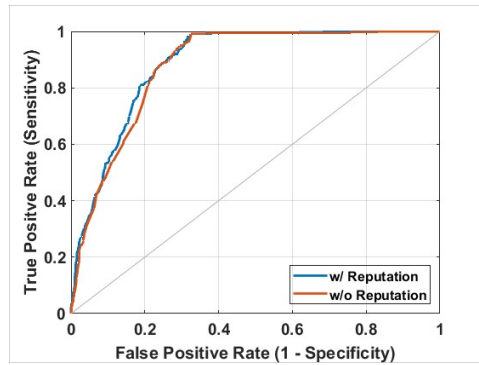
〈그림 3-19〉 2019년도 ROC



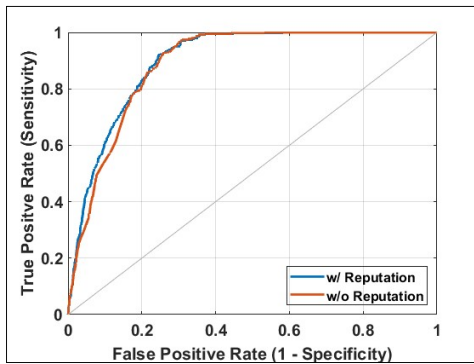
〈그림 3-20〉 2020년도 ROC



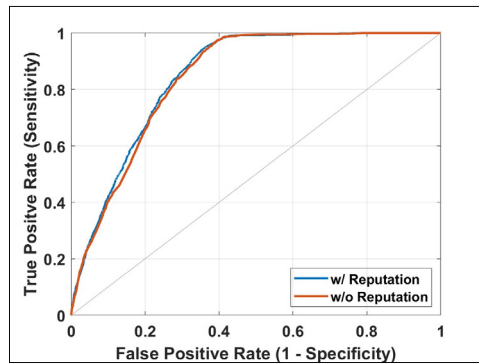
〈그림 3-21〉 2021년도 ROC



〈그림 3-22〉 2022년도 ROC



〈그림 3-23〉 2023년도 ROC

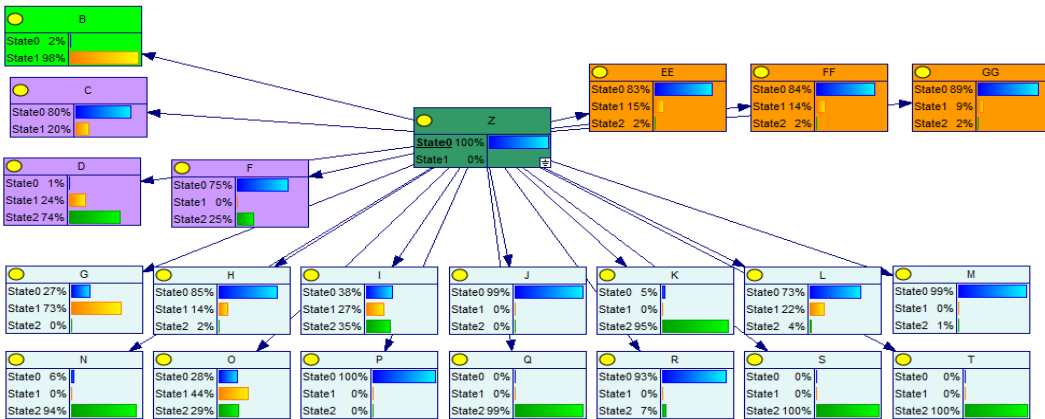


〈그림 3-24〉 통합 ROC

다면평가 실시 데이터를 분석한 결과, 다면평가 유무에 따른 ROC 곡선을 비교해 보면, 다면평가가 추가된 모델(파란선)이 더 좌상단으로 약간 이동한 것을 확인할 수 있다. 그러나 그 차이가 크지 않아, ROC 곡선이 좌상단으로 이동한 정도로 볼 때 다면평가 요소를 추가하여 학습한 모델이 진급 예측 성능에 미치는 영향은 미미한 것으로 판단된다.

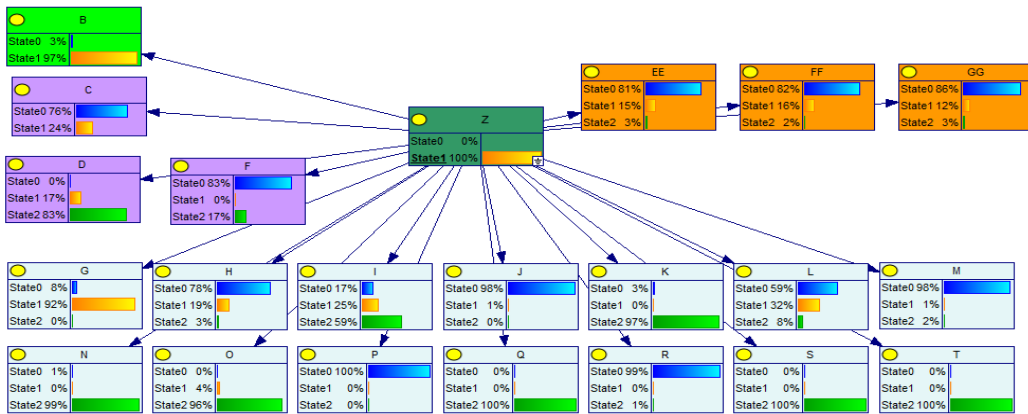
7.3 진급 결과에 따른 CPT 교정 결과

다면평가 실시 비율은 매년 증가하고 있다. 비록 다면평가는 참고 사항에 불과하지만, 이는 주변인들로부터의 평판을 반영하므로 향후 중요한 요소로 작용할 가능성이 있다. 따라서, 다면평가가 포함된 모델의 학습을 통해 조건부 확률(Conditional Probability Table, CPT)을 사용하여 다면평가의 영향도를 분석하였다. 먼저, 통합 데이터를 학습시킨 진급 모델의 Bar Chart는 다음과 같다.



〈그림 3-25〉 미진급(Z=0)인 경우의 Bar chart

〈그림 3-24〉에서는 미진급 시의 Bar chart를 표현하였다. 진급이 되지 않았을 경우(Z State = 0), 상급자의 다면평가(EE)의 확률분포는 State 0 = 83%, State 1 = 15%, State 2 = 2%이다. 이는 미진급 시 상급자로부터의 다면평가 점수 구간이 이산화 구간의 가장 낮은 State 0에 있을 확률이 83%, State 1에 있을 확률이 15%, State 2에 있을 확률이 2%임을 의미한다. 동료들의 다면평가(FF)의 확률 분포는 State 0~2까지 각각 84%, 14%, 1%이며, 하급자의 다면평가(GG)도 89%, 12%, 1%로 동일하다.

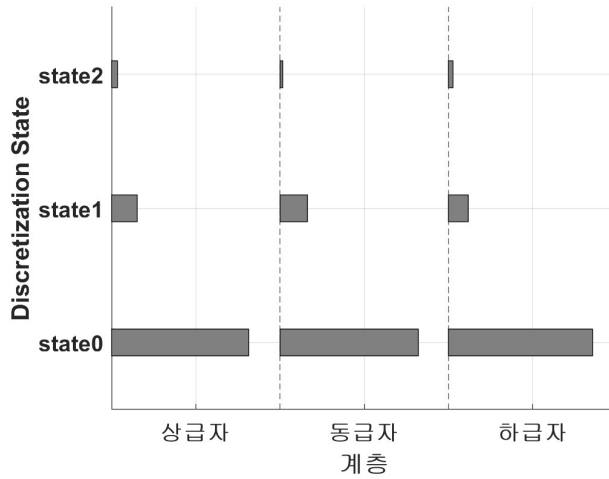


〈그림 3-26〉 진급(Z=1)인 경우의 Bar chart

〈그림 3-25〉에서는 진급이 되었을 때의 Bar chart를 표현하였다. 진급한 경우 (Z State = 1) 상급자의 다면평가(EE)의 확률분포는 State 0 = 81%, State 1 = 15%, State 2 = 3%이다. 이는 진급 시 상급자들이 평가한 다면평가 점수가 State 0에 있을 확률이 81%, State 1에 있을 확률이 15%, State 2에 있을 확률이 3%임을 의미한다. 동료들의 다면평가(FF)의 확률분포는 State 0부터 2까지 각각 82%, 16%, 2%이며, 하급자의 다면평가(GG)도 86%, 12%, 3%로 비슷하다. 이 산화 수준이 3인 경우의 CPT 값은 다음 〈표 3-16〉 및 그림 〈3-26〉과 같다.

〈표 3-16〉 통합년도 진급 유무에 따른 CPT (전체 데이터)

구분	CPT(%)					
	상급자		동료		하급자	
	미진급	진급	미진급	진급	미진급	진급
State 0	82.75	81.34	84.23	82.11	88.55	85.52
State 1	14.85	15.17	14.2	16.29	9.29	11.76
State 2	2.39	3.47	1.56	1.58	2.14	2.7

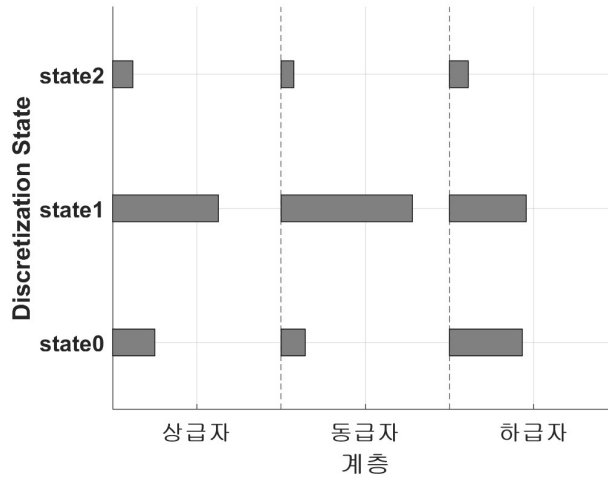


〈그림 3-27〉 통합연도 진급자 bar 그래프 (전체 데이터)

통합연도의 다면평가 CPT를 보면 상급자, 동급자, 하급자 모두 State 0에서 가장 높은 확률을 보인다. 하지만 이는 다면평가 미실시자의 데이터도 포함되어 있으므로 정확성이 다소 떨어진다. 따라서 다면평가를 실시한 대상의 데이터만을 추출하여 분석한 결과는 아래 〈표 3-17〉과 같다.

〈표 3-17〉 통합연도 진급 유무에 따른 CPT (다면평가 실시 데이터)

구분	CPT(%)					
	상급자		동료		하급자	
	미진급	진급	미진급	진급	미진급	진급
State 0	34.62	25.07	26.09	14.35	58.99	43.26
State 1	55.85	62.85	66.58	78.04	31.77	45.56
State 2	9.51	12.06	7.31	7.59	9.23	11.16

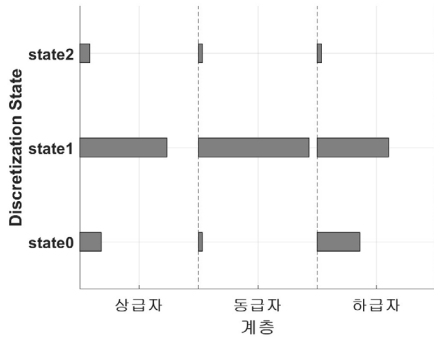


〈그림 3-28〉 통합연도 진급자 bar 그래프 (다면 실시 데이터)

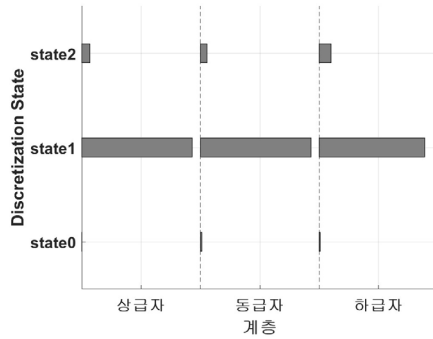
다면평가를 받은 인원들의 데이터를 분석한 결과, 상급자의 평가에서 중간 점수를 받은 인원이 상대적으로 많았다. 진급자의 경우, 상급자 평가에서 State 1에 속할 확률이 62.85%였으며, 동급자 평가에서는 78.04%, 하급자 평가에서는 45.56%로 각각 1수준에서 가장 높은 비율을 보였다. 이를 통해 진급자들의 대부분이 1수준에서 높은 평가를 받았음을 알 수 있다. 5개년간 진급에 따른 확률분포는 〈표 3-18〉에 제시되어 있다.

〈표 3-18〉 5개년도 진급 유무에 따른 CPT (다면평가 실시 데이터)

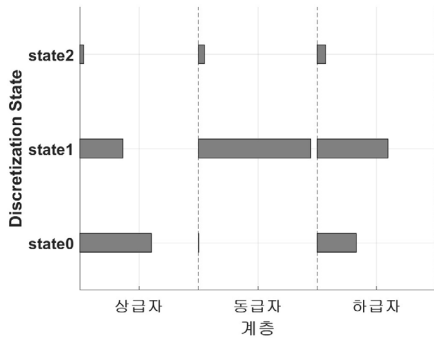
구분	2019년(%)					
	상급자		동료		하급자	
	미진급	진급	미진급	진급	미진급	진급
State 0	32.7	18.1	2.7	3.3	48.8	36
State 1	63.9	73.4	92.2	93.2	49.6	60.2
State 2	3.3	8.4	4.9	3.3	1.4	3.7
	2020년(%)					
State 0	1.1	0.1	3.4	1.2	2.6	0.9
State 1	96.1	93	92.4	93.2	93	88.9
State 2	2.7	6.8	4	5.5	4.3	10
	2021년(%)					
State 0	76.9	60.4	3.1	0.3	51.7	33
State 1	22.7	36.3	92.4	94.6	46.8	59.7
State 2	0.2	3.2	4.3	5	1.4	7.1
	2022년(%)					
State 0	22.8	19.2	39.8	26	80.1	75.4
State 1	62.2	59.7	51.8	63.9	6.3	9.7
State 2	14.9	21	8.2	10	13.4	14.7
	2023년(%)					
State 0	42.6	30.7	43.1	35.3	70.3	64
State 1	43.2	51	46.4	51.8	15.5	17.6
State 2	14	18.2	10.4	12.7	14	18.2
	통합(%)					
State 0	34.62	25.07	26.09	14.35	58.99	43.26
State 1	55.85	62.85	66.58	78.04	31.77	45.56
State 2	9.51	12.06	7.31	7.59	9.23	11.16



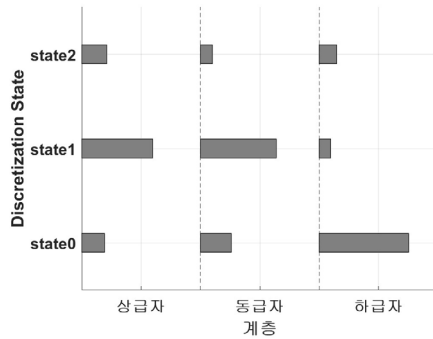
〈그림 3-29〉 2019년 진급자 bar 그래프
(다면 실시 데이터)



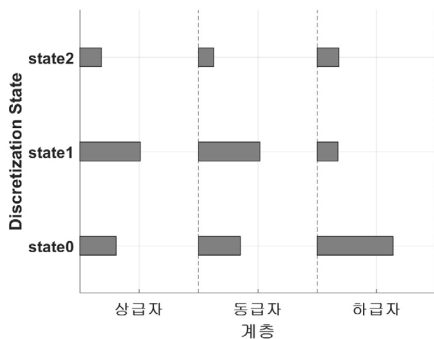
〈그림 3-30〉 2020년 진급자 bar 그래프
(다면 실시 데이터)



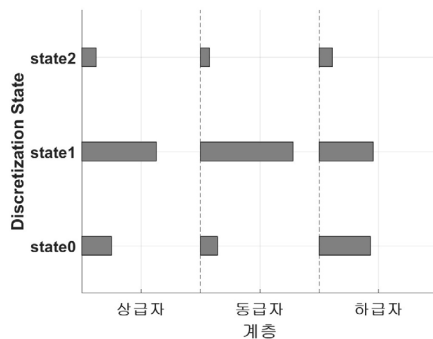
〈그림 3-31〉 2021년 진급자 bar 그래프
(다면 실시 데이터)



〈그림 3-32〉 2022년 진급자 bar 그래프
(다면 실시 데이터)



〈그림 3-33〉 2023년 진급자 bar 그래프
(다면 실시 데이터)



〈그림 3-34〉 통합연도 진급자 bar 그래프
(다면 실시 데이터)

2020년에는 진급 유무와 관계없이 State 1이 대다수였다. 특히 2020년에는 상급자의 평가에서 각각 96.1%, 93%였다. 동급자도 각각 92.4%, 93.2%를 기록하였고, 하급자의 평가에서도 93%, 88.9%로 매우 높은 경향을 보였다.

2021년에는 상급자의 평가에서 State 0인 경우가 미진급자는 76.9%, 진급자는 60.4%이었다. 하지만 동급자 간의 평가에서는 State 1의 확률이 미진급자는 92.4%, 진급자는 94.6%로 매우 높았다. 하급자의 평가에서는 미진급자의 경우 State 0과 1이 각각 51.7%, 46.8%로 유사했으며, 진급자의 경우 하급자로부터 State 1을 받을 확률이 59.7%로 가장 높았다.

2022년에는 상급자의 평가에서 State 1의 확률이 미진급자는 62.2%, 진급자는 59.7%로 높았다. 동급자의 평가도 각각 51.8%, 63.9%로 가장 높게 나타났다. 하지만 하급자의 평가에서는 State 0이 미진급자의 경우 80.1%, 진급자의 경우 75.4%로 매우 높았다.

2023년에는 상급자의 평가에서 State 1이 미진급자는 43.2%, 진급자는 51%로 다소 높았다. 동급자 간 평가에서도 State 1이 미진급자는 46.4%, 진급자는 51.8%로 높게 나타났다. 하지만 하급자의 평가에서는 State 0이 미진급자는 70.3%, 진급자는 64%로 매우 높았다.

〈표 3-18〉이 시사하는 바는 진급자의 경우 전체적으로 높은 상태(State 1 또는 2)의 확률이 높다는 것이고, 이는 평판이 미진급자에 비해 상대적으로 우수하다는 것이다. 이는 우리의 상식과 일치하는 결과로 정성적 평가인 다면평가가 정량적인 인자들을 보완하는 데 사용될 수 있음을 나타낸다. 특히 다면평가 인원의 비중이 확대된 2022년과 2023년의 정확도(〈표 3-14〉)가 크게 개선된 것으로 판단할 때 다면평가의 적용은 기존의 정성평가인 위원회 역할을 대체할 가능성을 시사한다.

8. 소결론

부사관 진급 절차를 보면 임무 수행에 대한 보상 개념으로 주어진 임무를 수행하는 데 동기를 유발하기 위한 개념이 강하다. 격오지 근무 점수나 근속 진급이 평가항목에 포함되어 있다는 점이 이를 뒷받침한다. 근무를 꺼리기 쉬운 격오지에서 근무 경험이나 오랜 기간 성실히 근무한 근속기간이 긴 인원에게 가산점을 부여하여 동기를 부여하는 측면을 확인할 수 있다. 아울러 전반적인 관리 능력보다는 전문성을 더 높이 평가하기 때문에 학위 점수나 자격증 점수가 평가항목으로 포함되어 있다는 점도 확인할 수 있었다.

다양한 머신러닝의 적합성을 비교하였다. 그중 베이지안 네트워크 모델의 AUC가 다른 모델보다 높은 수준을 유지해 성능 검증에는 문제가 없다고 판단되었다. 또한, 연도별로 다면평가 실시 여부에 따른 정확도를 비교한 결과, 2020년과 2021년에는 다면평가가 포함된 데이터의 정확도가 낮았으나, 2022년과 2023년에는 소폭 증가하는 경향을 보였다.

평가항목을 학습한 결과를 기반으로 진급을 예측한 결과, 예측률이 높은 수준을 유지하였는데, 평가항목의 다각화와 기준이 세분되어 대상자들에 대한 변별력이 높았다고 볼 수 있다. 다른 의미로는 선발 위원회 등 기초 평가항목의 영향력이 다소 적다고 할 수 있겠다.

민감도 분석을 통해 평가항목별 영향력을 확인한 결과는 절차상 배점과 상이한 결과를 보여주었는데, 근무평점의 배점이 가장 높지만 진급추천, 자격증, 상훈, 어학 순으로 영향력이 강한 것으로 나타나 실질적인 영향력과 배점이 차이가 있는 것으로 나타났다. 이는 배점이 높은 항목이 영향력이 높게 나타나기 위해서는 평가방법의 개선이나 배점의 재조정이 필요하다고 할 수 있다.

다면평가의 유효성을 검증한 결과를 살펴보면, ROC의 차이가 크지 않아 다면평가를 바로 적용하기에는 아직 시기상조라고 할 수 있다. 다만, 다면평가 실시 비율이 매년 증가하고 있으며, 진급자의 경우 대부분 높은 상태(state 1 또는 2)의 평가를 받았다. 다면평가의 중요성이 커질수록 이는 다면평가가 포함된 모델의 활용 가능성이 더욱 커질 것으로 예상된다.

다면평가가 진급과의 연관성을 더 구체적으로 보기 위해 CPT 분석을 추가로 실시하였다. 진급자의 다면평가 점수가 중간 점수대의 비율이 크게 나타났다. 연도별로 살펴보면 2019년, 2020년에는 진급자 중에서 다면평가에서 중간 정도의 점수 평가 비율이 모두 80% 이상이었다. 그러나 연도에 따라 차이가 있지만 다면평가가 낮아 나오는 비율이 증가하는 경우도 발견되었다. 이는 다면평가 점수가 좋은 경우가 드물고, 연도에 따라 일관되지 않은 경우도 발생한 것이 확인된 만큼 부사관의 다면평가 점수를 진급 항목에 반영하는 것은 신중할 필요가 있다고 볼 수 있다.

IV. 장교 진급 모델 연구

1. 공군 장교 진급제도

공군 장교 진급심사는 장병 진급관리 규정을 따르고 있다.³⁶⁾ 첫째, 기본 평가항목으로 자력점수를 산정한다. 자력점수는 인사참모가 주관하여 평가점수를 산정하고 본인 확인 절차를 거친다. 둘째, 추천 심사위원회를 소집하며, 대상 계급에 따라 위원회의 구성이 달라진다. 중령 진급심사 대상자의 평가는 대령이 위원장을, 중령이 심사위원으로 구성된다. 추천 심사위원회의 만장일치에 의해 진급 공식의 2~3배수의 대상자를 추천한다. 이때, 갑, 을, 병 3개의 독립된 심사위원회가 각각 점수를 부여하고 이는 자력 점수와 합산된다. 셋째, 2차로 선발 심사위원회가 소집되어 평가점수를 부여한다. 이때 대상자는 추천 심사위원회에서 추천된 대상자에 한한다. 선발 심사위원회의 평가점수는 만장일치 혹은 다수결에 의하여 개인별 평가에 따라 부여된다. 자력점수를 포함한 최종 점수가 높은 순으로 대상자의 순위가 부여된다. 진급 순위가 결정되면, 인사권자의 승인을 거쳐 진급 선발이 종료되고 선발 결과가 공시된다.

진급심사의 기본이 되는 자력평가는 근무평정, 교육점수, 부대 지휘추천 점수, 체력검정 등 다양한 항목으로 구성되어 있다. 자력평가 점수가 기본이 되며, 이에 추천 심사위원회의 점수와 선발 심사위원회의 점수가 가산된다. 기본 점수의 비중이 가장 크며, 심사위원이 평가하는 데 기초 자료로 활용되므로 중요한 의미가 있다.

2. 공군장교 데이터 기술통계

본 연구는 2019년부터 2023년까지 5년간 공군 중령 진급 대상자 2,013명의 자료가 대상으로 연구하였다. <표 4-1>은 기술통계 분석한 결과이다. 근무평정부 터 체력검정까지 항목은 자력 점수에 해당한다. 추천 심사위원회는 갑, 을, 병으로 독립적으로 구성되며, 핵심역량 갑, 을, 병 점수로 반영되었다. 선발위원회의 점수는 선발점수로 반영되어 있다. 진급 결과는 진급 여부를 말하는데, 더미 변수로 코

36) 공군규정 2-23, (2023) 장병 진급관리

〈표 4-1〉 장교 진급 항목 기술통계

변수	평균	표준편차	최소값	최대값
근무평정	27.6	0.91	17.4	30
보직경력	9.86	0.55	0	10
포 상	4.76	0.83	0	5
교 육	8.33	1.16	0	10
지휘추천	19.33	0.89	0	20
감 점	-0.07	0.50	-7	0
보안 별점	-0.003	0.04	-1	0
체력검정	4.63	0.22	3.8	5
핵심역량 갑	3.05	1.10	1	5
핵심역량 을	3.02	1.11	1	5
핵심역량 병	3.02	1.11	1	5
선발점수	3.80	2.14	0	5
진급결과	0.43	0.49	0(Fail)	1(Pass)
상급자 다면평가*	16.47	2.89	6.19	25
동급자 다면평가*	14.78	2.24	10.35	24.36
하급자 다면평가*	16.51	2.58	1.75	25

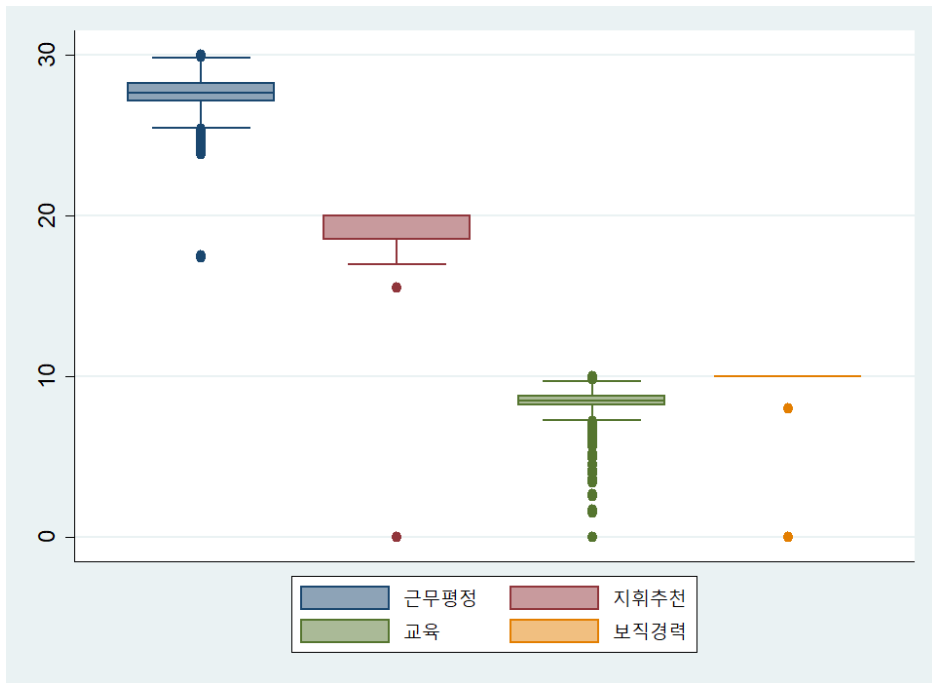
* 다면평가 점수는 참고사항으로 평가점수에 포함되지 않음

당되었다. 1은 진급을, 0은 선발되지 않은 경우를 의미한다. 평균이 0.43이므로 해당 기간 진급율은 43%라고 할 수 있다. 다면평가는 대상자에 대한 상급자, 동기, 하급자들이 평가한 점수로 진급심사에 평가항목으로 반영되지는 않으며, 심사 위원들의 참고 자료로만 활용되는 점수이다.

〈그림 4-1〉은 평가항목 중 배점이 높은 항목을 상자그림(Boxplot)으로 표현한 도표이다. 제시된 표에서는, 근무평정과 지휘추천의 배점이 가장 비중이 크다는 것을 알 수 있다. 근무평정의 최대값과 최소값에 대한 차이는 12.6점, 지휘추천은 20점으로 지휘추천의 범위가 넓은 것처럼 보이지만, 표준편차를 보면 포상이나, 교육과 큰 차이가 없는 것으로 나타나 평균 점수를 기준으로 밀집한 분포가 예상된다. 〈그림 4-1〉에서 상자 형태로 나타나는 IQR(inter Quatile Range)의 길이는 중간 50%의 범위를 의미한다. 근무평정과 지휘추천을 보면 둘 다 IQR이 매우 작게 나

타나, 많은 대상자들의 점수가 가운데 밀집한 형태라는 것을 확인할 수 있다.

평균 점수는 근무평정, 지휘추천, 보직경력 순으로 높은 점수가 나왔다. 이 항목이 배점이 크고 중요하게 평가된다고 할 수 있다. 하지만, 표준편차는 교육, 근무평정, 지휘추천, 포상 순으로 높게 나타나 교육 점수가 더 차이가 크게 날 가능성이 크다. 표준편차가 높은 항목 간의 수치 상위 4개 항목의 차이가 크지 않은 것으로 나타났다. 표준편차는 평균에 비례하여 커지는 것이 일반적이는데, 평균 점수는 근무평정과 지휘추천이 교육에 2~3배 차이 나는 데 반해 표준편차는 더 낮다는 것은 개인별 점수 차가 매우 낮고, 해당 항목의 변별력이 떨어진다는 것을 의미한다.

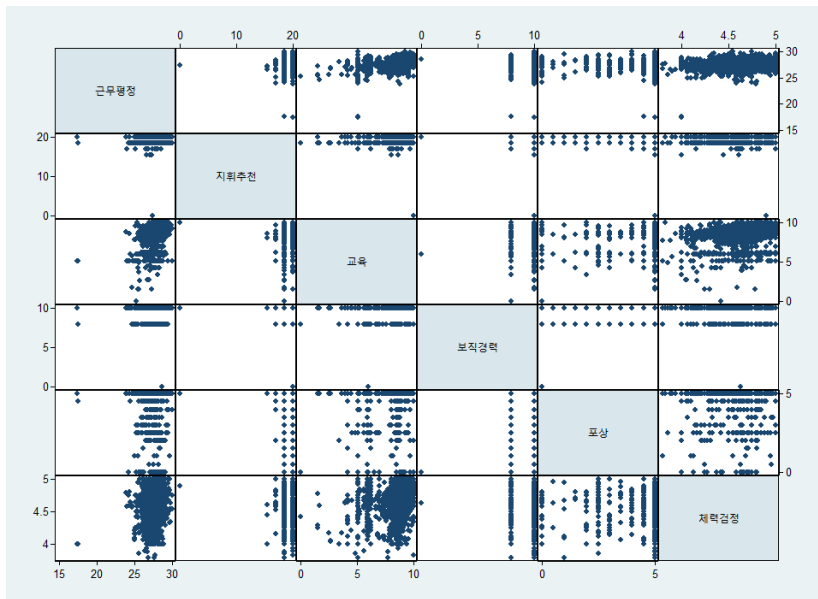


〈그림 4-1〉 주요 평가항목 상자도표

3. 공군장교 데이터 상관관계 분석

평가항목 간의 연관성을 분석하기 위하여 상관관계를 분석하였다. 분석 도구는 STATA 14이다. 상관관계 분석결과는 <표 4-2>와 같다. 먼저 상관관계가 높은 관계는 추천위원회 갑, 을, 병 점수로 상관관계가 0.7 정도 수준이며 유의수준 0.05에서 통계적으로 유의한 것으로 나타났다. 추천 위원회가 독립적으로 구성되고 평가됨에도 불구하고 위원회 간 점수가 유사한 경향이 있다고 할 수 있다.

다음으로 진급 여부에 영향을 미치는 항목을 살펴보면, 가장 높은 상관관계는 선발점수로 상관관계수가 1로 나타났다. 즉 선발점수를 부여받은 사람은 모두 진급이 되었다고 할 수 있다. 인사권자의 승인 절차가 있으나, 특별한 사유가 없는 이상 선발 위원회가 실질적인 선발을 결정한다고 말할 수 있다. 그밖에 선발 여부 결정에 영향을 주는 항목은 추천위원회 점수로 0.5 정도의 상관관계를 갖고 있어, 다른 항목에 비해 수치가 높지만, 상관관계가 강하다고 보기는 어렵다.



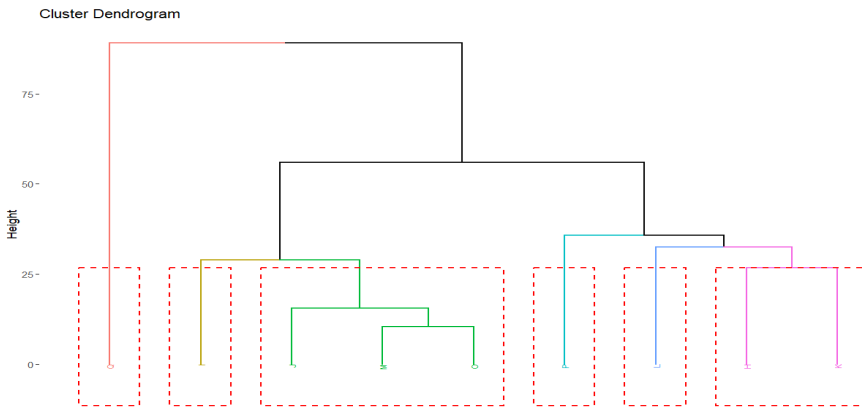
<그림 4-2> 주요 평가항목 산점도

〈표 4-2〉 상관관계 분석결과

변수	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1) 근무평정	1												
(2) 보좌경력	0.035	1											
(3) 포상	0.080*	0.181*	1										
(4) 교육	0.275*	0.204*	0.273*	1									
(5) 지휘추천	0.180*	-0.02	0.112*	0.079*	1								
(6) 감점	0.080*	0.004	0.102*	0.025	0.120*	1							
(7) 보안별점	0.060*	-0.009	-0.018	0.003	0.022	0.064*	1						
(8) 체력검정	0.186*	0.039	0.085*	0.184*	0.036	0.014	0.025	1					
(9) 감점수	0.372*	0.088*	0.184*	0.283*	0.383*	0.197*	0	0.156*	1				
(10) 음점수	0.382*	0.089*	0.207*	0.297*	0.399*	0.205*	0.049*	0.185*	0.708*	1			
(11) 병점수	0.417*	0.115*	0.213*	0.323*	0.408*	0.203*	0.054*	0.195*	0.719*	0.745*	1		
(12) 선발점수	0.262*	-0.047	0.158*	0.118*	0.220*	0.136*	0.006	0.126*	0.426*	0.418*	0.439*	1	
(13) 선발여부	0.366*	-0.028	0.143*	0.166*	0.290*	0.117*	0.019	0.168*	0.534*	0.541*	0.560*	1.000*	1

* p<0.05

〈그림 4-2〉는 자력점수의 분포에 대하여 산점도(scatter plot)를 표시한 것이다. 산점도에서 선형 관계가 나타나는 경우 두 항목의 관련성이 높다고 볼 수 있다. 하지만, 제시된 그림에서는 강한 상관관계를 갖는 그래프는 확인되지 않아 평가항목 간의 상관관계는 크지 않은 것으로 볼 수 있다. 주요 항목인 근무평정과 지휘추천은 높은 점수대에 다수의 데이터가 표시되어 상향 점수가 매우 평균화되었다는 것을 확인할 수 있다.



〈그림 4-3〉 클러스터 덴드로그램

전체적인 데이터 간에 유사성을 판단하기 위하여 클러스터링 분석을 시행하였다. R을 활용하여 계층적 클러스터링을 한 결과는 〈그림 4-3〉과 같다. 초록색으로 표시된 부분은 상점, 감점, 보안 별점이고, 분홍색으로 묶인 부분은 근무평정과 교육이다. 보안 별점, 감점, 상점이 유사한 패턴이 있고, 근무평정과 교육성적이 유사한 패턴을 보인다. 하지만, 클러스터링은 패턴을 기준으로 하는 것이므로 변수 간 직접적인 상관관계를 보기 위해서는 〈표 4-2〉의 직접 상관관계를 재확인해야 한다. 일반적으로 0.6 이상의 상관관계가 있을 때 상관관계가 보통이라고 판단하는데, 해당 변수들의 관계는 상당히 낮은 수준을 유지하므로, 관계가 있다고 보기는 어려우며, 유사할 가능성이 있다고 판단할 수 있다.

여러 상관분석을 하였을 때, 추천 심사위원회의 점수는 각각 독립적인 평가단이 판단하는 결과임에도 상관관계가 높은 것으로 나타나 평가단이 대상자에 대하여 유사한 관점을 갖고 있다고 할 수 있다. 자력점수의 경우 상관관계 분석결과 관련성 찾기는 어려웠으며, 클러스터링을 통한 덴드로그램 분석 시 근무평정이 교육점수와 유사한 패턴을 보이고, 보안과 감점 간 유사한 패턴을 보인다고 할 수 있다.

4. 장교 머신러닝 모델 비교

부사관과 마찬가지로 다면평가를 추가한 모델에 대해 다양한 머신러닝 알고리즘의 성능을 AUC로 비교하여, 베이지안 네트워크 사용의 적절성을 평가하였다.

4.1.1. 2019년 데이터 머신러닝 AUC 분석결과

〈표 4-3〉 2019년 모델별 AUC

Model	AUC	Model	AUC
Matern 5/2 GPR	0.94	Linear Regression Robust Linear	0.88
Rational Quadratic GPR	0.94	Coarse Tree	0.83
Bayesian Network	0.93	Bagged Trees	0.8
Squared Exponential GPR	0.93	Boosted Trees	0.79
Exponential GPR	0.93	Fine Tree	0.78
Tri-layered Neural Network	0.92	Medium Tree	0.78
Medium Neural Network	0.91	SVM Linear	0.77
Narrow Neural Network	0.91	SVM Quadratic	0.75
Stepwise Linear Regression	0.9	Fine Gaussian SVM	0.74
Wide Neural Network	0.9	Kernel SVM	0.74
Linear Regression	0.89	SVM Coarse Gaussian	0.74
Interactions Linear Regression	0.89	SVM Medium Gaussian	0.74
Bi-layered Neural Network	0.89	SVM Cubic	0.72

Bayesian Network(0.93)는 Matern 5/2 GPR, Rational Quadratic GPR에 이어 우수한 예측 성능을 보였으며, 머신러닝 알고리즘 전체에서 매우 높은 수준의 AUC를 기록하였다. 신경망 계열 모델도 AUC 값이 0.9대를 기록하여 우수한 성능을 보이고 있으며, 비선형 패턴을 잘 학습하는 경향을 보여주고 있다.

4.1.2. 2020년 데이터 머신러닝 AUC 분석결과

가우시안 과정 회귀 계열 모델이 가장 우수한 성능을 보이고 있으며, 신경망 모델 계열이 다음으로 높은 AUC를 나타내고 있다. Bayesian Network(0.86)는 다른 머신러닝 알고리즘과 비교하여 중간 수준의 AUC를 나타내고 있지만, 여전히 높은 수준의 예측 성능을 보인다.

〈표 4-4〉 2020년 모델별 AUC

Model	AUC	Model	AUC
Rational Quadratic GPR	0.93	Bayesian Network	0.86
Squared Exponential GPR	0.93	Coarse Tree	0.8
Matern 5/2 GPR	0.93	SVM Linear	0.79
Exponential GPR	0.93	Kernel SVM	0.79
Wide Neural Network	0.92	Fine Gaussian SVM	0.79
Narrow Neural Network	0.91	SVM Medium Gaussian	0.79
Medium Neural Network	0.91	SVM Coarse Gaussian	0.79
Tri-layered Neural Network	0.91	Medium Tree	0.79
Stepwise Linear Regression	0.89	Boosted Trees	0.79
Interactions Linear Regression	0.89	Bagged Trees	0.79
Linear Regression	0.89	Fine Tree	0.79
Linear Regression Robust Linear	0.88	SVM Cubic	0.72
Bi-layered Neural Network	0.87	SVM Quadratic	0.71

4.1.3. 2021년 데이터 머신러닝 AUC 분석결과

신경망, 가우시안 과정 회귀 계열 모델은 여전히 높은 예측 성능을 발휘하고 있으며, 이번 연도에는 특히 선형 회귀 기반 모델이 우수한 성능을 보인다. Bayesian Network(0.88)는 중상위 수준의 AUC를 나타내었으며, 결정 트리, SVM 계열 모델보다 더욱 나은 예측 능력을 보였다.

〈표 4-5〉 2021년 모델별 AUC

Model	AUC	Model	AUC
Bi-layered Neural Network	0.95	Bayesian Network	0.88
Linear Regression Robust Linear	0.95	Tri-layered Neural Network	0.84
Narrow Neural Network	0.94	Fine Tree	0.84
Exponential GPR	0.94	Medium Tree	0.84
Matern 5/2 GPR	0.94	Boosted Trees	0.83
Squared Exponential GPR	0.94	Coarse Tree	0.82
Rational Quadratic GPR	0.94	Bagged Trees	0.79
Linear Regression	0.93	SVM Cubic	0.74
Interactions Linear Regression	0.93	Kernel SVM	0.7
SVM Linear	0.92	Fine Gaussian SVM	0.7
Wide Neural Network	0.92	SVM Coarse Gaussian	0.7
Medium Neural Network	0.91	SVM Medium Gaussian	0.7
Stepwise Linear Regression	0.91	SVM Quadratic	0.65

4.1.4. 2022년 데이터 머신러닝 AUC 분석결과

2022년은 다른 연도에 비해서 AUC 값이 전체적으로 낮아진 경향을 보인다. 가우시안 과정 회귀, 신경망 계열 모델이 가장 높은 예측 성능을 보여주며, Bayesian Network(0.79)는 GPR, 신경망 모델보다는 낮지만, 안정적인 예측 성능을 발휘하고 있다.

〈표 4-6〉 2022년 모델별 AUC

Model	AUC	Model	AUC
Matern 5/2 GPR	0.84	Bagged Trees	0.78
Squared Exponential GPR	0.84	SVM Linear	0.77
Rational Quadratic GPR	0.84	Boosted Trees	0.76
Narrow Neural Network	0.84	SVM Medium Gaussian	0.75
Exponential GPR	0.83	SVM Quadratic	0.75
Stepwise Linear Regression	0.82	SVM Coarse Gaussian	0.75
Wide Neural Network	0.82	Kernel SVM	0.74
Linear Regression	0.81	Fine Gaussian SVM	0.74
Interactions Linear Regression	0.81	Medium Tree	0.73
Linear Regression Robust Linear	0.81	Fine Tree	0.73
Medium Neural Network	0.8	Bi-layered Neural Network	0.71
Bayesian Network	0.79	SVM Cubic	0.67
Coarse Tree	0.78	Tri-layered Neural Network	0.63

4.1.5. 2023년 데이터 머신러닝 AUC 분석결과

2023년에는 선형 회귀 계열 알고리즘이 0.9대의 AUC를 나타내며 가장 높은 예측 성능을 보였다. 이어서 가우시안 과정 회귀, 신경망 계열 모델이 높은 AUC 값을 나타낸다. Bayesian Network(0.82)는 신경망 모델과 유사한 수준으로 준수한 성능을 보인다.

〈표 4-7〉 2023년 모델별 AUC

Model	AUC	Model	AUC
Linear Regression Robust Linear	0.9	Medium Neural Network	0.83
Linear Regression	0.89	Bayesian Network	0.82
Interactions Linear Regression	0.89	SVM Linear	0.82
Squared Exponential GPR	0.88	SVM Cubic	0.76
Rational Quadratic GPR	0.88	Fine Tree	0.75
Exponential GPR	0.87	Medium Tree	0.74
Matern 5/2 GPR	0.87	Kernel SVM	0.72
Bi-layered Neural Network	0.86	SVM Coarse Gaussian	0.72
Tri-layered Neural Network	0.86	SVM Medium Gaussian	0.72
Boosted Trees	0.85	Fine Gaussian SVM	0.71
Stepwise Linear Regression	0.85	Coarse Tree	0.71
Narrow Neural Network	0.85	Wide Neural Network	0.7
Bagged Trees	0.84	SVM Quadratic	0.67

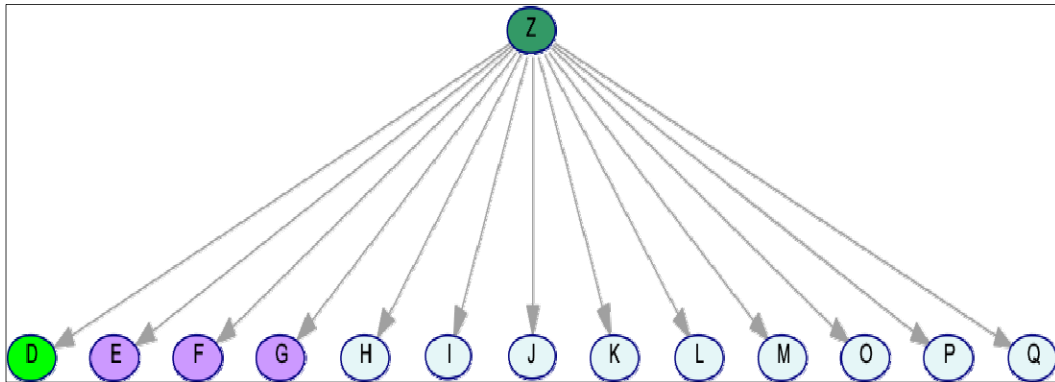
4.1.5. 머신러닝 AUC 분석결과 고찰

전년도 머신러닝 AUC 분석결과를 종합해 보면, 다양한 머신러닝 알고리즘들이 각기 다른 성능을 보여주었으며, 특히 가우시안 과정 회귀(GPR) 계열과 신경망 계열 모델들이 대부분의 연도에서 우수한 예측 성능을 기록했다. GPR 계열 모델은 복잡한 비선형 패턴을 잘 학습하며, 전반적으로 높은 AUC 값을 기록하여 데이터에 대한 강력한 예측력을 보여준다. 신경망 계열 모델 역시 비선형 패턴을 효과적으로 학습하며, 다수의 연도에서 AUC 0.9대 이상의 성능을 보인 점에서 우수성을 입증하였다.

Bayesian Network는 매년 안정적인 예측 성능을 보여주며, 이러한 결과는 Bayesian Network가 데이터 세트의 변동성에 영향을 받지 않고 신뢰할 수 있는 모델임을 나타낸다. 특히 진급 예측 모델은 여러 변수를 동시에 고려하여 복잡한 의사결정을 내려야 하는데, Bayesian Network는 변수 간의 관계를 명확하게 모형화하고 예측할 수 있다는 점에서 진급 예측에 적합한 모델로 판단된다.

5. 기본 모델 모델링

본 연구에서는 공군 중령 진급 대상자 2,013명의 데이터를 사용하여 베이지안 네트워크를 구성하였다. 평가항목 진급에 영향을 줄 수 있는 진급연도, 심사그룹, 임관년도 모델링에 추가하였다.



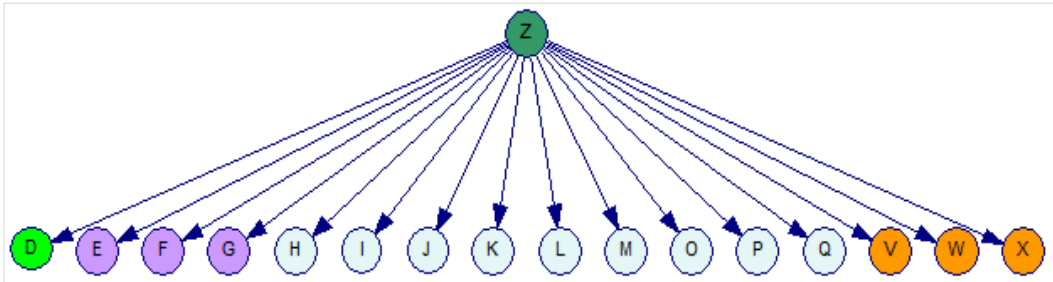
〈그림 4-4〉 장교 기본 모델³⁷⁾

위의 그림은 진급 모델을 나이브 베이지 모델로 구성한 것이다. 노드 Z는 진급 결과를 의미하고 노드 아들 노드인 D부터 Q까지는 진급심사에 평가항목을 나타낸다. 〈그림 4-4〉에서 각각의 엣지가 평가결과인 Z에서부터 이어져 나온다. 진급결과가 평가항목에 영향을 주는 모델이므로, 직관적인 방향과 반대라고 생각할 수 있겠지만 각 항목은 평가 결과에 센서 역할을 한다고 할 수 있으므로, 그림과 같이 구성하였다. 각 평가항목에 대한 정보가 주어질 때 진급 추천 대상이 될지 그렇지 않을지 예측할 수 있도록 고안하였다. 진급 결과인 부모 노드가 없으므로 간단히 연속 혹은 이산적인 확률분포로 표현된다. 일반적으로 이산 확률분포가 선호되므로 연속변수에 대하여 이산화하였다. 각각의 항목을 구성하는 노드는 조건부 확률분포(Conditional Probability Distribution)를 갖고, 분포의 크기는 부모 노드의 상태의 수(m)와 자식 노드 상태의 수(n)의 곱($m \times n$)이 된다.

37) 이두열, 백세일, 김민생, 김신곤(2021) “베이지안 네트워크 기반 전시 파손항공기 임무수행능력 예측 모델” Transactions of the Korean Society of Mechanical Engineers A · February 2021

6. 다면평가 추가 모델 모델링

본 연구에서 기본 모델에 다면평가를 추가로 포함한 모델을 <그림 4-5>와 같이 제시하였다. 현재 다면평가 점수는 평가항목이 아니므로 점수에 반영되지 않는다. 평가항목에 반영된 근무평정과 다면평가를 비교하면, 근무평정은 직속상관의 평가만을 나타내지만, 다면평가는 다양한 참여자가 평가한 점수이므로 객관성 확보에 장점이 있다고 할 수 있다. 따라서, 이러한 다면평가가 추가하는 것에 대한 필요성을 검토하기 위해 다면평가가 추가된 모델을 구성하였다. 다면평가는 상급자, 동급자, 하급자 평가로 구성되어 있다. <그림 4-5>의 가장 우측에 주황색으로 표시된 부분이 다면평가를 측정하는 3개의 변수이다.



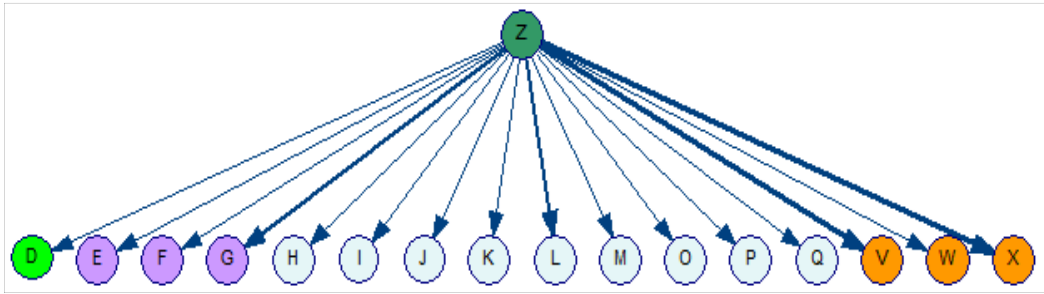
<그림 4-5> 장교 다면평가 추가 모델

7. 민감도 분석결과

민감도는 지니(GeNIe)의 영향력 강도(Strength of Influence) 기능을 통하여 분석하였다. 영향력 강도는 연결된 선의 굵기로 표현된다. 영향력 강도는 각 노드의 조건부 확률분포가 부모 노드로부터 받는 영향력 나타내는 것이다. 영향력의 강도는 각 평가항목의 조건부 확률분포 간의 거리를 기준으로 계산된다.³⁸⁾

모델의 민감도 분석결과 <그림 4-6>와 같이 나왔다. 그림에서 엣지의 두께가 굵을수록 영향력 강도가 강하다는 것을 의미한다. 해당 결과는 지니의 여러 분석 방법 중 EM(Expectation Maximization) 알고리즘을 사용하였다.

38) Bayesfusion, LLC. GiNIe Modeler Programmer's Manual, 2023, Available From <https://support.bayesfusion.com/docs/GeNIe>(accessed Feb. 14, 2024)



〈그림 4-6〉 민감도 분석결과

민감도 분석결과, 영향력이 큰 항목은 다면평가(상, 하), 임관년도, 부대 추천, 다면평가(중), 진급년도 순으로 나타났다. 여기서 다면평가(상, 중, 하)는 각각 상급자, 동료, 하급자로부터의 평가를 의미한다.

첫 번째와 두 번째로 중요한 항목은 다면평가(상)와 다면평가(하)로, 다면평가가 매우 높은 영향력을 보였다. 그다음으로 임관년도가 큰 영향을 미쳤으며, 그 뒤로 부대 추천, 다면평가(중), 진급년도가 순차적으로 높은 영향력을 가진 것으로 나타났다. 민감도 분석결과는 해마다 차이가 있지만, 다면평가가 추가되었을 때 영향력이 크게 증가하는 경향을 확인할 수 있었다. 5개년도의 민감도 분석결과는 〈표 4-8〉에 제시되어 있다.

〈표 4-8〉 연도별 민감도 분석결과

민감도 순위	연도					
	19	20	21	22	23	통합
1	다면상	다면상	다면상	다면상	다면상	다면상
2	추천	다면하	다면하	다면하	근평	다면하
3	다면하	다면중	임관년도	추천	다면중	임관년도
4	다면중	진급년도	근평	진급년도	다면하	추천
5	진급년도	임관년도	추천	임관년도	임관년도	다면중
6	임관년도	체검	진급년도	다면중	진급년도	진급년도
7	성과	성과	다면중	성과	성과	성과
8	체검	보직	성과	근평	체검	체검
9	보직	상훈	체검	체검	상훈	근평
10	교육	교육	교육	상훈	보직	상훈

〈표 4-8〉은 민감도 순위를 1위부터 10위까지 나타내고 있으며, 다면평가 상 및 하가 상위권에 위치한 것을 확인할 수 있다. 상급자의 평가는 5개년 동안 일관되게 1순위로, 가장 높은 영향력을 보였다. 이는 진급자의 경우 대부분 상급자로부터 긍정적인 다면평가를 받았음을 의미한다. 또한, 하급자의 평가도 높은 영향력을 발휘했다. 동료 간 평가(중)는 일부 연도(21, 22)에서는 영향력이 다소 낮았지만, 여전히 강력한 영향력을 가지고 있다.

8. 교정 및 타당성 분석

8.1. 다면평가 추가모델 타당성 분석

민감도 분석 시 사용한 EM 알고리즘은 최초 각 평가항목 노드의 사전 분포가 상태별로 동일한 확률을 갖는다고 가정한다. 이후 실제 데이터를 활용하여 각각의 확률을 업데이트하는 과정을 반복적으로 거친다.

타당성 분석을 통하여 K-fold Cross Validation을 이용하여 모델의 타당성을 확인하였다. 〈표 4-9〉는 진급확률이 0.5 이상이면 진급했다고 판단할 때, 다면평가를 포함 모델과 기본 모델의 정확도를 비교한 결과이다.

〈표 4-9〉 진급 예측 정확도 비교

기본 모델		예 측	
		비선발	선발
실제	비선발	850	305
	선발	351	507
다면평가 추가모델		예 측	
		비선발	선발
실제	비선발	918	237
	선발	232	626

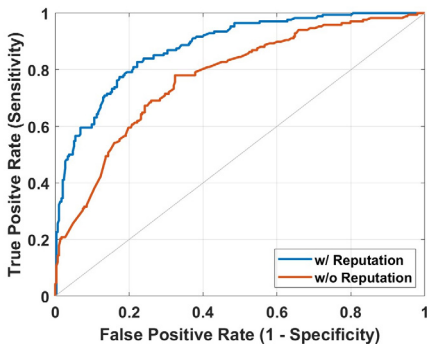
5개년도 진급대상자 데이터를 분석한 결과, 기본 모델의 예측 정확도는 67.7%(1357/2013)였으나, 다면평가를 추가한 모델의 경우 76.7%(1544/2013)로 상승한 것을 확인할 수 있었다.

〈표 4-10〉 연도별 정확도 비교

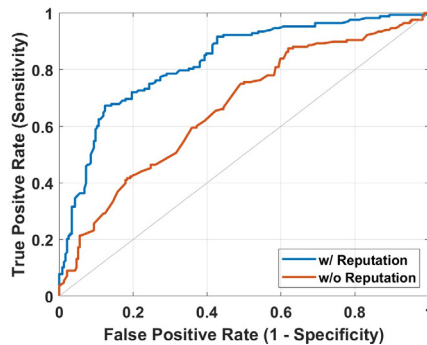
진급연도(년)	기본모델	다면평가 추가
19	72.2%	80.7%
20	62.9%	74.4%
21	70.2%	82.6%
22	67.4%	71.8%
23	67.3%	70%
통합	67.7%	76.7%

〈표 4-10〉은 연도별로 다면평가 유무에 따른 예측 정확도를 비교한 표이다. 기본 모델의 정확도는 연도별로 최소 67.3%(23년)에서 최대 72.2%(19년)이었지만, 다면평가를 추가한 모델은 최소 70%(23년)에서 80.7%(19년)로 정확도가 상승하였다. 따라서 진급 예측 모델에서는 다면평가가 충분히 고려될 요소임을 확인할 수 있다.

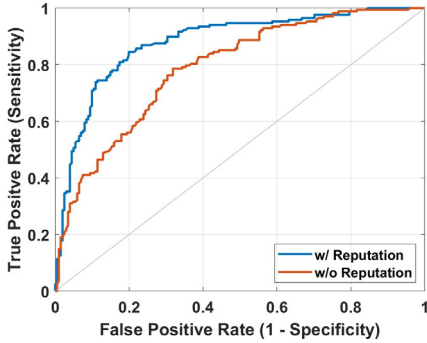
아래 그림은 다면평가 모델과 기본 모델의 1종 오류에 대한 민감도 분석을 실행한 결과이다. 파란 선은 다면평가를 포함한 모델이며, 붉은 선은 기본 모델이다. 모든 문턱값에서 파란 선의 민감도가 높은 것을 확인할 수 있다. 이는 다면평가 포함 모델의 정확도가 더 높다는 것을 나타내므로, 다면평가를 포함한 모델을 사용하는 것이 더 적합하다고 말할 수 있다.



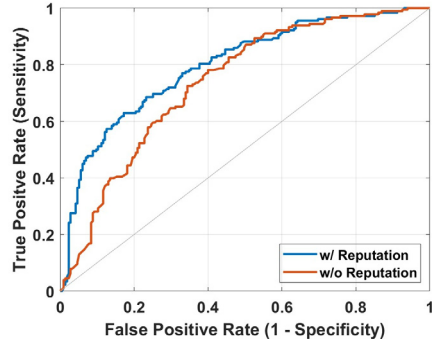
〈그림 4-7〉 19년도 ROC



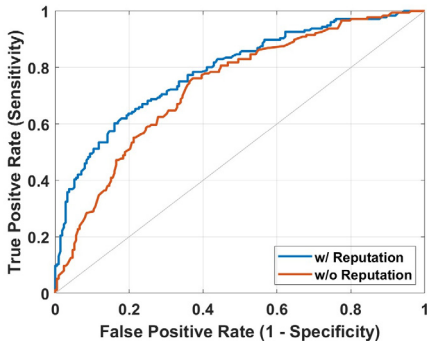
〈그림 4-8〉 20년도 ROC



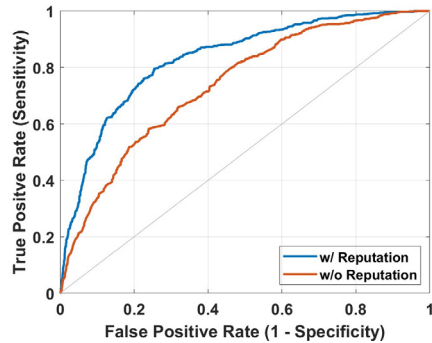
〈그림 4-9〉 21년도 ROC



〈그림 4-10〉 22년도 ROC



〈그림 4-11〉 23년도 ROC



〈그림 4-12〉 통합 ROC

위 그림은 2019년부터 통합년도까지 3수준일 때 ROC 곡선이다. 매년 및 통합 ROC 곡선을 보면, 파란선(다면평가가 포함된 모델)이 좌측 상단에 더 가까이 위치한 것을 확인할 수 있다. 이는 다면평가가 추가된 경우 진급 예측 성능이 더 우수하다는 것을 보여주는 결과다. ROC 곡선의 좌측 상단으로의 이동은 모델이 더 높은 민감도와 특이도를 보인다는 의미로, 다면평가 요소가 예측력을 향상시킨다는 점을 시사한다.

장교 진급 데이터를 기반으로 평가항목의 중요도와 다면평가의 효용성을 살펴보았다. 연구 결과에서 확인할 수 있듯이, 첫 번째, 현재 진급심사의 배점과 영향력에 차이가 있다는 점이다. 두 번째는, 현재 참고사항으로만 사용되는 다면평가가 평가항목으로 포함될 경우 진급예측 정확도가 상승한다는 점이다. 다양한 인원이 평가한 다면평가를 평가항목으로 활용하는 것이 진급심사에 긍정적인 영향을 줄 수 있다. 단, 주의해야 하는 점은 동급자 평가는 포함시키는 것이 좋지 않다는 것이다. 동급자는 진급 경쟁자가 포함될 수 있어 오히려 신뢰도를 떨어뜨릴 가능성

이 있다. 따라서, 상급자와 하급자의 다면평가만 활용하는 것이 적절하다고 할 수 있다.

8.2 이산화 계획의 효과

장교 모델도 부사관 모델과 같이 동일하게 이산화 과정을 통해 수준에 따른 모델 정확도를 비교하였다.

〈표 4-11〉 연도별 다면평가 이산화 수준에 따른 모델 정확도

진급연도(년)	이산화 수준(States) 및 확률(%)				
	3	6	9	12	15
19	80.69	80.26	79.17	80.26	82.21
20	74.37	74.12	73.38	75.12	72.88
21	81.57	83.19	82.11	83.19	80.75
22	71.75	72.01	70.99	71.75	71.5
23	70.1	70.36	71.64	71.9	71.64
통합	76.6	76.5	76.35	76.7	77.04

이산화 후 연도별 정확도는 19년도 80.69% ~ 82.21%, 20년도 72.88% ~ 75.12%, 21년도 80.75% ~ 83.19%, 22년도 70.99% ~ 72.01%, 23년도 70.1% ~ 71.9%, 통합 데이터는 76.35 ~ 77.04%로 이산화 효과가 크지 않았다. 이는 다면평가를 미실시한 대상자들의 데이터가 다수 포함되었기 때문이다. 이에 따라, 다면평가를 실시한 데이터만을 사용하여 재학습을 진행한 후 추가적인 분석을 수행하였다.

〈표 4-12〉 다면평가 실시 데이터 학습 후 정확도

진급연도(년)	이산화 수준(States)에 따른 정확도(%)				
	3	6	9	12	15
19	82.91	87.26	84.47	86.64	86.64
20	74.9	77.45	78.54	78.54	80.36
21	83.76	82.47	85.04	81.62	82.47
22	71.33	70.09	71.65	70.71	69.78
23	73.95	75.88	74.27	72.66	74.91
통합	75.39	76.89	76.62	77.17	76.76

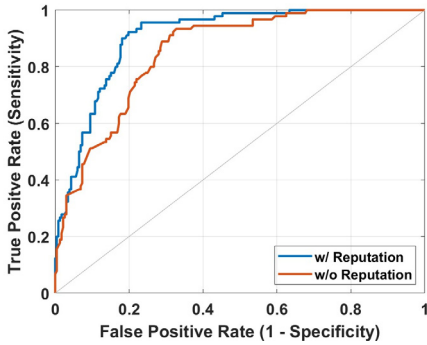
이산화 후 연도별 정확도를 살펴보면, 19년도와 20년도에서는 각각 15수준에서 정확도가 86.64%, 80.36%로 가장 높게 나타났다. 그러나 21년도에서는 12수준에서 81.61%, 22년도에서는 15수준에서 69.78%로 이산화 이후에도 정확도가 하락하는 경향을 보였다. 23년도는 15수준에서 74.91%로, 최초 3수준의 73.95%에 비해 약 1% 상승했으나, 매년 전체적으로 정확도가 떨어지고 있다. 이를 바탕으로, 장교 모델의 경우 이산화 계획이 모델의 정확도에 큰 영향을 미치지 않는 것으로 판단된다.

〈표 4-13〉 전체 및 일부(다면평가 실시 데이터) 반영 후 정확도 비교

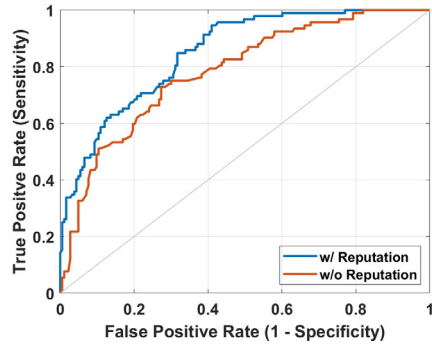
진급연도	이산화 수준(States)에 따른 정확도(%)									
	3		6		9		12		15	
	전체	일부	전체	일부	전체	일부	전체	일부	전체	일부
19	80.69	82.91	80.26	87.26	79.17	84.47	80.26	86.64	82.21	86.64
20	74.37	74.9	74.12	77.45	73.38	78.54	75.12	78.54	72.88	80.36
21	81.57	83.76	83.19	82.47	82.11	85.04	83.19	81.62	80.75	82.47
22	71.75	71.33	72.01	70.09	70.99	71.65	71.75	70.71	71.5	69.78
23	70.1	73.95	70.36	75.88	71.64	74.27	71.9	72.66	71.64	74.91
통합	76.6	75.39	76.5	76.89	76.35	76.62	76.7	77.17	77.04	76.76

전체 데이터를 학습한 경우(‘전체’)와 다면평가를 실시한 데이터만 학습한 경우(‘일부’)의 정확도를 비교한 결과, 22년도 및 통합 데이터를 제외하고는 다면평가를 실시한 데이터만 학습한 경우의 정확도가 더 높았다. 특히, 19년과 20년의 경우 15수준에서 각각 86.4%와 80.36%로 상승했다. 이는 다면평가 실시율이 높아질 경우, 해당 데이터를 충분히 활용할 수 있을 것으로 판단된다.

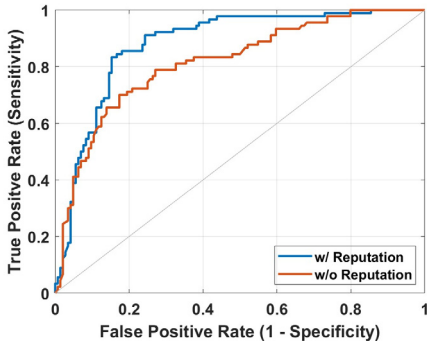
2019년부터 통합년도까지 다면평가 실시한 데이터만을 활용한 ROC 곡선은 아래 그림과 같다.



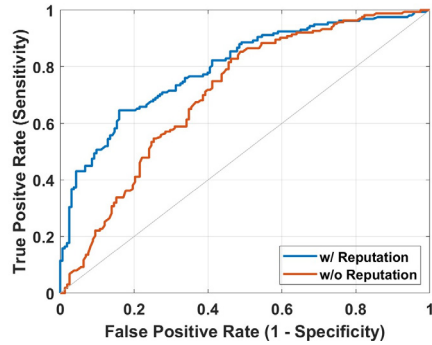
〈그림 4-13〉 19년도 ROC



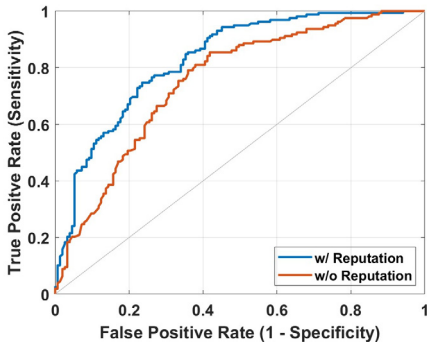
〈그림 4-14〉 20년도 ROC



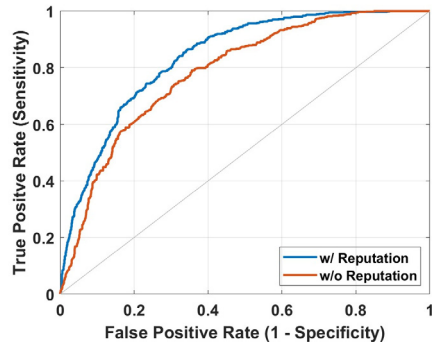
〈그림 4-15〉 21년도 ROC



〈그림 4-16〉 22년도 ROC



〈그림 4-17〉 23년도 ROC

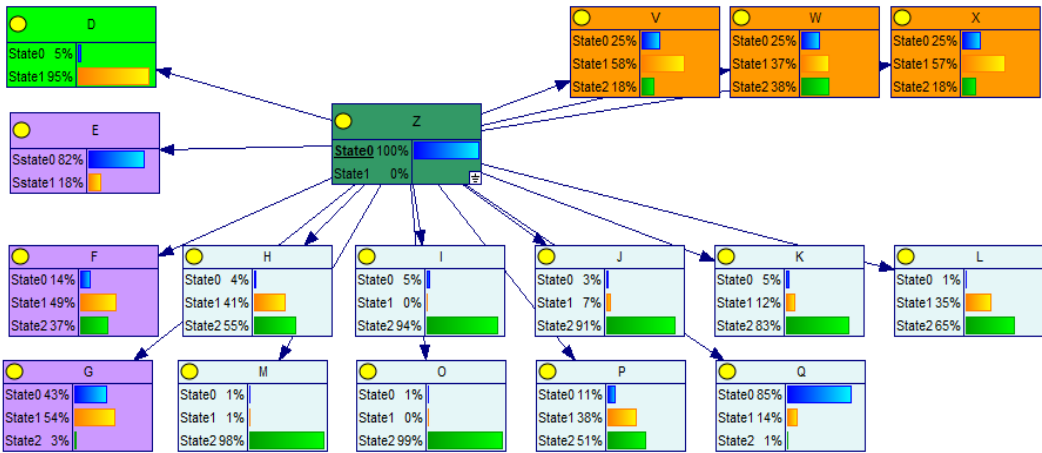


〈그림 4-18〉 통합 ROC

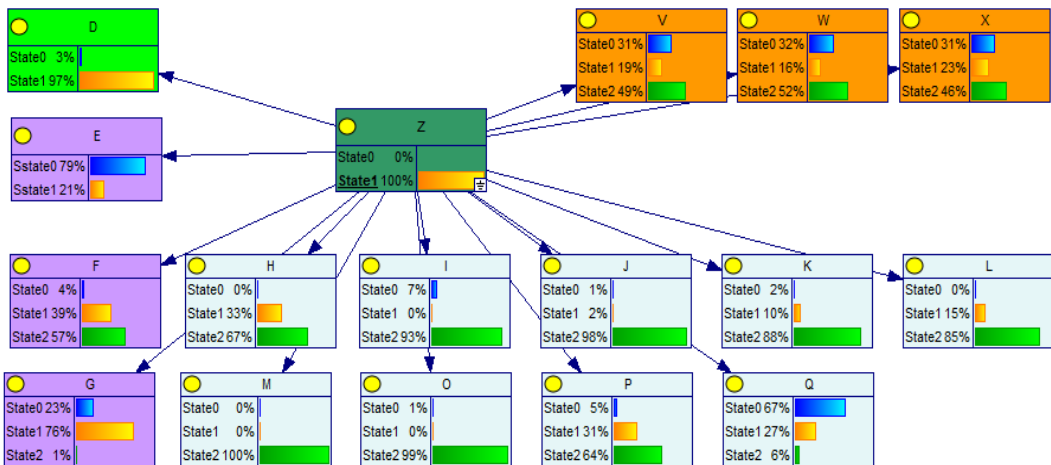
다면평가만 반영된 데이터로 분석한 결과, ROC 곡선에서 파란선이 좌상단으로 더 올라가는 것을 확인할 수 있다. 이는 다면평가를 포함한 모델이 더 높은 민감도와 특이도를 보이며, 예측 성능이 우수하다는 것을 의미한다.

8.3 진급 결과에 따른 CPT 교정 결과

다면평가가 포함된 모델을 학습시킨 후, 조건부 확률 테이블(CPT)을 사용하여 다면평가의 영향도를 분석하였다. 먼저, 통합 데이터의 진급 대상자 데이터를 기반으로 학습된 진급 모델의 Bar chart는 다음과 같다.



〈그림 4-19〉 미진급(Z=0)인 경우의 Bar chart

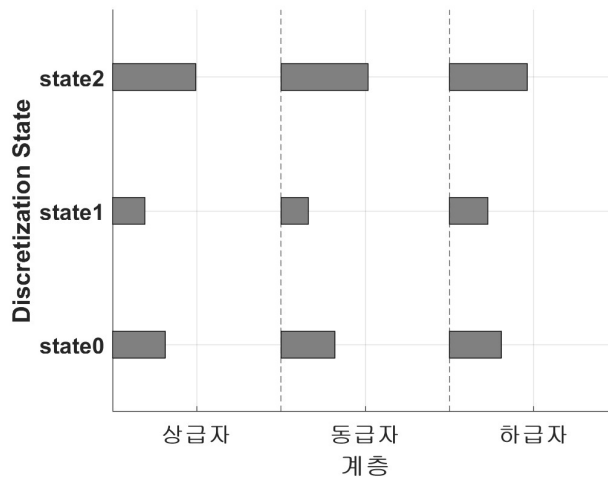


〈그림 4-20〉 진급(Z=1)인 경우의 Bar chart

〈표 4-14〉 통합연도 진급 유무에 따른 CPT (전체 데이터)

구분	CPT(%)					
	상급자		동료		하급자	
	미진급	진급	미진급	진급	미진급	진급
State 0	24.8	31.3	25.4	32	24.8	30.8
State 1	57.5	19.2	36.7	16.2	56.7	22.8
State 2	17.5	49.4	37.8	51.7	18.3	46.2

5개년도 전체 데이터로 학습한 경우 49.4%, 동료의 경우 51.7%, 하급자의 경우 46.2%의 확률로 2수준이 제일 높았다. 이는 대다수의 진급자들이 모든 계층으로부터 좋은 평가를 받은 것으로 알 수 있다.



〈그림 4-21〉 통합연도 진급자 bar 그래프 (전체 데이터)

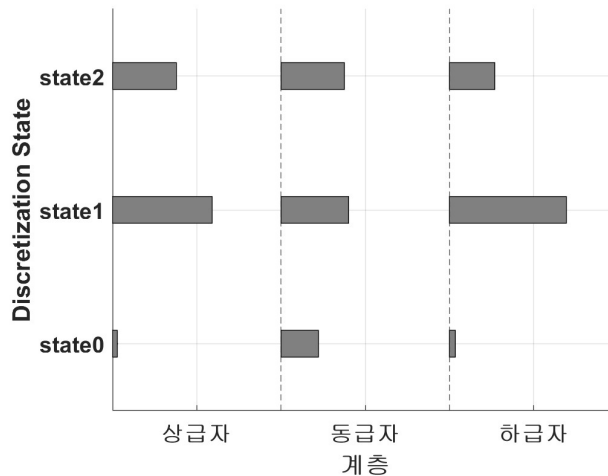
〈그림 4-21〉 진급자의 다면평가 분포를 나타낸 것이다. 상가한 바와 같이 점수가 가장 높은 State 2의 비중이 가장 높지만, 0의 비율이 생각보다 많이 분포하고 있는 것이 확인된다. 이는 다면평가 미실시자의 데이터가 '0'으로 반영되어 진급자 중 다면평가를 실시하지 않은 원인이 있기 때문이다. 따라서, 미실시자가 다면평가

결과에 대한 정확도를 떨어뜨리고 있다고 볼 수 있다. 다면평가 미실시자의 영향을 제거한 결과는 다음과 같다.

〈표 4-15〉 통합연도 진급 유무에 따른 CPT (다면평가 실시 데이터)

구분	CPT(%)					
	상급자		동료		하급자	
	미진급	진급	미진급	진급	미진급	진급
State 0	31.2	2.8	46.4	22.2	31.8	3.5
State 1	60.4	59.1	36.8	40.1	60.6	69.4
State 2	8.2	37.9	16.6	37.6	7.4	26.9

진급자의 경우 1수준에서 상급자 59.1%, 동료 40.1%, 하급자 69.4%의 확률로 가장 높았다. 이는 진급자 모두 중간 정도의 평가를 받은 것으로 확인할 수 있다.

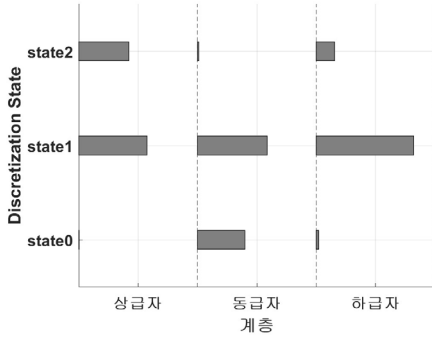


〈그림 4-22〉 통합연도 진급자 bar 그래프 (다면실시 데이터)

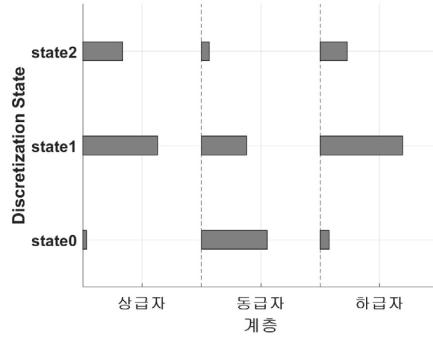
〈그림 4-22〉에서 보면 State 0에 대한 부분이 확연히 줄어든 것을 확인할 수 있었으며, 진급자들의 다면평가가 우수하다는 것을 재확인할 수 있었다.

〈표 4-16〉 5개년도 진급 유무에 따른 CPT (다면평가 실시 데이터)

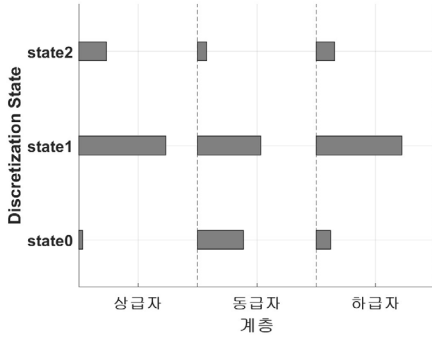
구분	2019년(%)					
	상급자		동료		하급자	
	미진급	진급	미진급	진급	미진급	진급
State 0	17.2	0.3	70.2	40	38.7	2.2
State 1	75.4	57.5	24.5	58.8	57.7	82.2
State 2	7.3	42.1	5.1	1.1	3.4	15.5
	2020년(%)					
State 0	44.8	3.2	79.2	55.4	49.7	7.6
State 1	50.2	63	18	38	44.2	69.5
State 2	4.9	33.6	2.7	6.5	6	22.8
	2021년(%)					
State 0	59	3.3	60.4	38.8	56.2	12.2
State 1	37.5	73.3	36.1	53.3	40.2	72.2
State 2	3.4	23.3	3.4	7.7	3.4	15.5
	2022년(%)					
State 0	0.6	2.4	0.2	0.6	8.2	0.6
State 1	78.4	33.7	64.9	44.7	82.9	74.8
State 2	20.8	63.8	34.8	54.6	8.8	24.5
	2023년(%)					
State 0	41.8	4.4	5.8	5.6	1.3	0.2
State 1	52.9	75.9	49.6	18.3	81	54.9
State 2	5.2	19.6	44.4	75.9	17.6	44.8
	통합년도(%)					
State 0	31.2	2.8	46.4	22.2	31.8	3.5
State 1	60.4	59.1	36.8	40.1	60.6	69.4
State 2	8.2	37.9	16.6	37.6	7.4	26.9



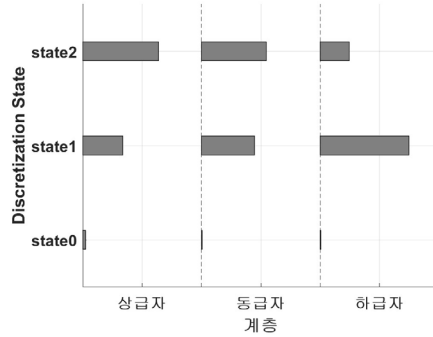
〈그림 4-23〉 2019년 진급자 bar 그래프 (다면 실시 데이터)



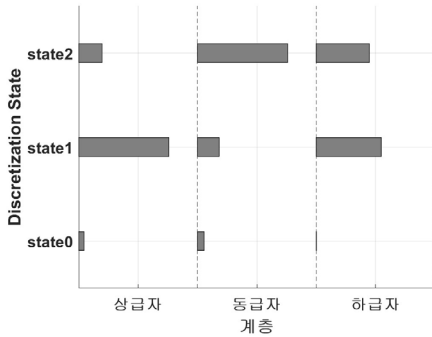
〈그림 4-24〉 2020년 진급자 bar 그래프 (다면 실시 데이터)



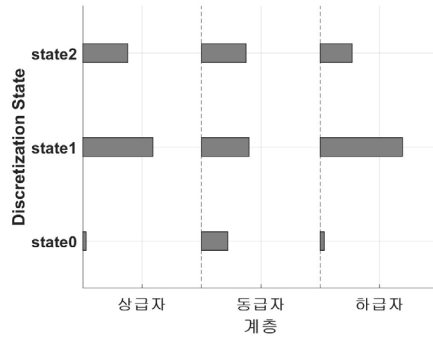
〈그림 4-25〉 2021년 진급자 bar 그래프 (다면 실시 데이터)



〈그림 4-26〉 2022년 진급자 bar 그래프 (다면 실시 데이터)



〈그림 4-27〉 2023년 진급자 bar 그래프 (다면 실시 데이터)



〈그림 4-28〉 통합연도 진급자 bar 그래프 (다면 실시 데이터)

19년부터 통합년도까지 진급자 중 다면평가가 0수준에 있는 경우는 20년도 동료들의 평가만 있다. 그 외 나머지 진급한 원인은 대부분 1 또는 2 수준에 있었다. 이 표가 시사하는 바는 부사관 모델과 같이 진급자는 전체적으로 높은 상태(state 1 또는 2) 확률을 가진다. 즉, 다면평가의 점수를 추가할 경우, 모델 성능이 조금 더 우수해지며 다른 정량적인 인자들을 보완하는데 사용될 수 있다.

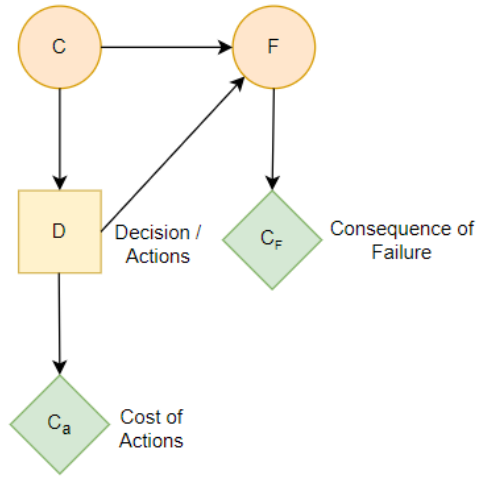
9. 향후 발전 방향

장교 진급 모델과 관련하여 향후 발전 방향을 제시하고자 한다. 조직에 있어서, 선택의 결과가 조직에 긍정적인 영향을 주었는지 그렇지 못한지를 확인하려는 노력은 지속될 필요가 있다. 본 모델에 따라 진급을 시킬 경우 이 결과가 조직에 유익한 결과를 가져올 수 있다는 점을 확인할 수 있다면 진급 모델 연구가 갖는 유용성을 확인할 수 있을 것이다. 이를 위하여 베이지안 의사결정 개념을 활용한 방안을 제시하고자 한다.

9.1. 베이지안 의사결정 개념

베이지안 의사결정은 불확실성을 고려하여 최적의 의사결정을 내리는데 활용할 수 있는 방법이다.³⁹⁾ 베이지안 의사결정 이론은 베이지 정리를 기반으로 비용의 최소화, 이익의 최대화와 같이 특정 목표를 달성하기 위해 발생 가능한 결과와 그에 따른 확률을 분석하고, 행동의 비용과 이익을 고려하여 최적의 의사결정을 지원한다. 베이지안 의사결정 이론에서는 의사결정자가 정보를 얻은 후 최고의 기대 효용을 실현할 수 있는 행동을 선택한다고 가정하는데, 효용은 결정에 따른 비용과 실패의 결과에 반비례한다. 즉, 비용이 많이 발생하면 실패할 확률은 줄어들고 비용이 적게 발생하면 실패할 확률이 늘어나게 되는 것이다. 따라서, 베이지안 의사결정은 성공과 실패에 따른 비용분석을 통해 최적의 행동 과정을 찾는 최적화 문제로도 정의할 수 있다.

39) Economou T, Stephenson DB, Rougier JC, Neal RA, Mylne KR, 2016, "On the Use of Bayesian Decision Theory for Issuing Natural Hazard Warnings," Proceedings of the Royal Society A, vol. 472



〈그림 4-29〉 베이지안 의사결정 모델 예시

〈그림 4-29〉는 베이지안 의사결정 분석을 위한 영향도로 구성요소의 실제 상태, 관찰, 발생하건 등 여러 요소간의 확률적 관계를 묘사한다. 〈그림 4-29〉에서 원형 노드는 임무의 성공 또는 실패 등의 사건을 나타내며, 사각형 노드는 어떠한 행동을 결정하는 결정 지점을 나타낸다. 또한, 마름모 노드는 의사결정 과정에서 발생하는 비용을 의미한다.

베이지안 의사결정의 최적화를 일반화하여 수식으로 나타내면 다음과 같이 나타낼 수 있다.⁴⁰⁾

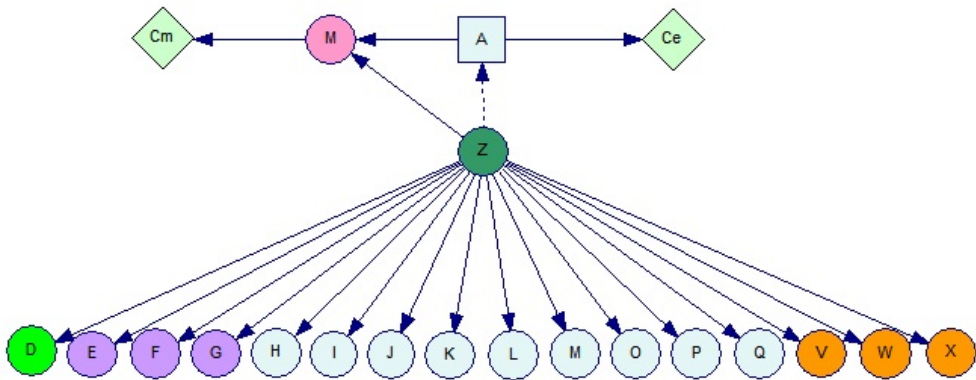
$$a_{opt}(z) = \arg_{a \in a_0, a_1, \dots} \min E_{c|z} [C_T(a, C)] \quad (5)$$

식 (5)에서 $a_{opt}(z)$ 는 관찰 결과 z 가 주어졌을 때 최적의 행동을 나타내며, 이는 의사결정에 따른 결과를 고려할 때 기대되는 총비용이 가장 낮은 행동을 나타낸다. $E_{c|z}$ 는 관찰 결과 z 가 주어졌을 때 시스템 상태 C 에 대한 조건부 확률 기댓값을 의미하고, $C_T(a, C)$ 는 행동 a 와 시스템의 실제 상태에 의존하는 총비용 함수를 의미한다. 베이지안 의사결정은 여러 발생 가능한 비용을 정량화하여 총비용 $E_{c|z} [C_T(a, C)]$ 을 최소화하는 의사결정을 의미하며, 의사결정을 함에 있어 불확실성이 존재하는 상황에서 유용하게 활용할 수 있다.

40) T. Economou, D. B. Stephenson, J. C. Rougier, R. A. Neal and K. R. Mylne, 2016, "On the use of Bayesian decision theory for issuing natural hazard warnings," Proc. R. Soc. A.472.

9.2. 수치 예시(Numerical Example)

장교 진급 모델 BN을 바탕으로 베이지안 의사결정 모델을 구현하면 <그림 4-30>와 같이 나타낼 수 있다. 베이지안 의사결정 모델에서 A는 의사결정 노드(Action Node), M은 임무 노드(Mission Node), C_m 은 진급 후 임무 성공과 실패에 따른 비용 노드, C_e 는 진급으로 인해 발생하는 비용 노드를 의미한다. 그리고, 점선 호(Dash Arc)는 의사결정 노드로 들어오는 정보의 흐름을 나타내는데, 이를 통해 의사결정 노드가 다른 노드들의 상태를 고려하여 정보에 기반한 의사결정을 내릴 수 있도록 하는 역할을 한다.

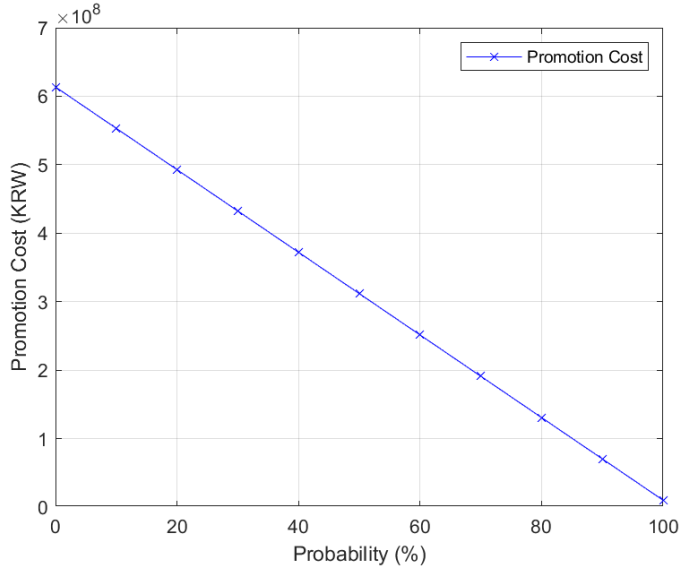


<그림 4-30> 장교 진급 베이지안 의사결정 모델

<그림 4-30>의 베이지안 의사결정 모델을 통한 의사결정 과정은 다음과 같다. 먼저 BN를 통해 진급확률이 산출되면, 이 정보는 임무 노드 M에 영향을 미치게 되고, 임무 노드의 정보를 바탕으로 C_m 노드에서는 진급 후 임무 성공 및 실패에 따른 비용을 산출한다. 이렇게 산출된 비용에 대한 정보들은 점선 호를 통해 의사결정 노드 A로 들어오고, 의사결정 노드에서는 C_m , C_e 값을 바탕으로 위험비용을 산출하고 총비용을 최적화하는 방향으로 의사결정이 진행된다.

베이지안 의사결정 모델을 기반으로 한 진급확률에 따른 기대 비용 그래프는 <그림 4-31>과 같이 나타낼 수 있다. 여기서는 진급확률에 따른 기대 비용을 계산하기 위해 베이지안 의사결정 모델의 비용 노드를 다음과 같이 설정하였다. 먼저, C_m 노드에서는 진급 후 주어진 임무의 성공과 실패에 따른 비용을 계산하기 위해 연봉 상승, 연금 증가, 정년 연장에 따른 추가 비용을 임무 성공과 실패에 따

른 비용으로 반영하였다. 또한, C_e 노드는 진급자가 새로운 계급과 직위에 맞는 필수 교육을 받아야 하므로, 이에 소요되는 교육비를 비용으로 포함시켰다.



〈그림 4-31〉 진급확률 vs. 기대 비용

위 내용을 바탕으로 분석한 결과, 베이지안 네트워크를 통해 산출한 진급확률이 높을수록 진급 후 발생하는 기대 비용은 감소하는 것으로 나타났다. 이는 진급확률이 높은 사람을 선택할수록 임무를 성공적으로 수행할 가능성이 높아지기 때문에, 임무 실패에 따른 추가 비용(연봉 상승, 연금 증가, 정년 연장, 교육 비용 등)이 줄어드는 결과로 해석할 수 있다. 즉, 진급에 영향을 미치는 요인들을 바탕으로 진급확률을 산출하고, 그 확률이 높은 사람을 진급시킴으로써 비용을 절감하고 효율성을 극대화할 수 있음을 보여준다.

하지만, 위의 결과는 예시로 제시한 것이며, 군의 성과를 측정하는데 한계가 있어 추후 과제로 제시하고자 한다. 정확한 분석을 위해서는 구체적인 비용 산출 모델이 필요하다. 해당 조직원이 진급 후 조직에서 지출하게 되는 비용은 산출할 수 있겠으나, 해당 조직원이 조직의 성과에 기여한 정도를 금액으로 환산하는 것에는 현실적인 한계가 존재한다. 다만, 여러 부서 중에서 업무 성격에 따라 업무 성과에 대한 금액을 유추할 수 있는 부서를 대상으로 한다면 일부 적용할 가능성이 있다고 할 수 있겠다.

10. 소결론

장교진급 모델 연구에서는 장교 진급 데이터를 바탕으로 베이지안 네트워크를 사용하여 평가항목에 대한 중요도를 분석하였다. 평가항목에 대한 중요도를 연구한 결과 근무평정, 부대 추천, 상점, 다면평가가 진급에 큰 영향을 준 요인으로 분석되었다. 상점은 점수 배점이 상대적으로 낮음에도 불구하고 진급에 중요한 요인으로 나타나는 등 현재 배점과 실제 영향성에 상이점이 있다는 것을 확인할 수 있다.

또한, 현행 방식에서 다면평가가 심사위원의 참고사항으로 사용되었으나, 다면평가 점수를 평가항목으로 사용하는 것이 진급 예측도를 높이는 것으로 나타나, 다면평가가 진급평가항목으로 가치가 있다는 점을 발견하였다.

추후 장교 진급 시 다면평가를 평가에 포함하는 방안을 제시하고자 한다. 다면평가는 다양한 평가자가 평가한 항목으로 직속상관의 평가로 이루어지는 근무평정보다 객관성이 높다. 다만, 잠재적 경쟁자가 평가하는 동급자의 다면평가는 객관성 유지가 어렵기 때문에 제외하고, 상급자와 하급자의 다면평가 점수를 평가에 포함하는 것이 바람직하다.

베이지안 의사결정 이론을 바탕으로 장교 진급 모델의 효용성 분석한 결과 장교진급 모델에서 예측한 진급확률이 높을수록 진급에 따른 기대 비용은 감소함을 확인할 수 있었다. 이는 장교 진급 베이지안 네트워크 모델이 진급 여부를 결정하는 과정에서 경제적 효율성을 고려한 객관적 기준을 제시하여, 진급 결정의 신뢰성을 높일 수 있음을 시사한다. 따라서, 베이지안 진급 분석 모델은 진급 여부를 결정하는 데 있어 추가적인 판단 자료로 활용될 수 있을 것으로 생각된다.

VII. 결론

그동안 진급 관련 연구들이 문헌을 통한 질적연구를 주로 하였다면, 본 연구는 정량적인 방법론을 사용하였다는 점에서 차별점을 가진다. 공군 간부 진급 데이터를 머신러닝 도구를 사용하여 분석함으로써 어떠한 평가항목이 진급에 실제로 큰 영향을 주었는지 확인 할 수 있었다. 연구 결과 조직의 의도한 바와 같이 배점을 높게 준 항목이 중요도가 높은 경우도 있었지만, 일부 항목은 예상과 달리 배점이 낮음에도 진급에는 영향력이 강한 항목이 있다는 점을 확인할 수 있었다. 아울러 논란이 많이 있었던 다면평가의 실효성을 실제 검증 과정을 거쳐 확인했으며 장교와 부사관에 있어서 다른 결과를 보여주었다는 부분도 앞으로 진급 정책 수립에 많은 시사점을 준다고 할 수 있다.

먼저, 부사관 진급 모델 분석결과에서는 진급에 영향력이 강한 항목의 배점이 모두 낮은 수준이라는 결과가 나왔다. 영향력이 강한 항목은 진급추천, 자격증, 포상으로 나타났는데, 이들 항목은 모두 개인 자력 점수에서 차지하는 비율이 낮다. 이는 배점이 높은 근무평정, 보직관리 등의 다양한 항목들이 평준화되어 실질적으로 작동하지 못하고 있을 수 있다는 의미이다. 부사관 진급 규정 중 보직관리는 평가방법이 매우 세분화 되어 있지만, 다양한 항목 중 유리한 사항을 선택하여 적용하다 보니, 많은 진급 대상자가 상향 평준화 되어 실질적인 영향을 주기보다는 기본적인 점수를 채우는 개념으로 활용되고 있을 가능성이 있다.

둘째, 장교 진급 연구 결과를 살펴보면, 배점이 높지만 영향력이 떨어지는 항목이 발견되었고, 해당 항목에 대한 조정 논의가 필요하다고 할 수 있다. 장교 진급에 영향력이 강한 변수는 근무평정, 지휘추천, 포상, 처벌, 다면평가(상급자, 하급자)인 것으로 나타났다. 근무평정과 지휘추천의 경우 배점이 높아 예상된 결과이나, 포상과 처벌은 배점이 가장 낮은 수준임에도 불구하고 진급에 영향력이 큰 것으로 나타났다. 반면, 교육과 보직경력은 배점이 높게 설정되었지만 진급에 대한 영향은 낮게 나타났다. 결과 중 흥미로운 부분은 일반적으로 점수 편차가 큰 항목일수록 더 점수 변별력이 커서 진급에 영향력이 강하게 나타날 가능성이 높은데, 가장 편차가 크게 나타난 교육이 모델 분석에서 영향력이 낮은 것으로 나타났다는 점이다. 이처럼 점수의 편차가 크다고 해서 진급의 직접적인 영향력이 강하다고 판단하면 잘못된 결과를 얻을 수 있다는 점을 보여준다. 연구결과에 배점보다 낮은 영향력이 있는 2개 항목에 대한 부분은 진급에 있어서 중요도를 재판단하고, 필요시 배점 상향 등을 고려할 필요가 있다고 할 수 있다.

장교와 부사관의 영향력을 검증한 결과를 보았을 때, 아쉽게도 평가 가중치가 설정된 것과 다르게 운영되는 항목들이 식별되었다. 즉, 배점이 높지만 영향력이 낮은 항목(장교의 경우 교육, 보직경력)은 배점을 상향 조정하여 자력 점수에서 영향력을 키우던지, 평가방법을 수정하여 변별력을 기르려는 노력이 필요하다.

셋째, 다면평가에 관한 결과는 장교와 부사관이 다른 결과를 보여주었다. 장교는 다면평가가 포함된 모델이 진급 예측력 수치가 더 우수한 것으로 나타났으며, 다면평가 항목에 대한 영향력을 분석하였을 때도, 상급자와 하급자의 평가가 영향력이 큰 것으로 나타났다. 즉, 상급자와 하급자의 다면평가 결과는 진급평가항목에 추가하는 것이 의미가 있다는 결과를 얻을 수 있었다. 하지만, 부사관의 경우 다면평가를 포함한 모델이 현재 방식보다 더 우수하다고 보기 어렵고, 영향력도 크지 않은 것으로 나타났다. 이는 다면평가 실시자의 비율이 낮아 진급 데이터가 정확히 반영되지 않은 것으로 확인하였다. 실제로 다면평가를 실시한 인원들의 데이터를 학습하였을 때, 21년도까지는 다면평가를 포함하지 않았을 때 더 정확도가 높았으나, 22년도부터는 다면평가를 포함한 데이터의 정확도가 더 높았다. 연도별로 다면평가 실시자의 비율이 증가하고 있어 추후에 다면평가를 포함한 자료 또한 부사관 진급 모델을 예측하는 데 높은 이바지를 할 것으로 보인다.

이번 연구의 한계점으로 수치화된 평가항목만을 연구 대상으로 선정하였기 때문에 반영되지 못한 외부요인에 대한 영향성은 확인하지 못하였다는 한계가 있다. 아울러, 본 연구에서는 베이지안 네트워크가 높은 정확도를 보여주었으며, 민감도 분석결과를 활용할 수 있다는 장점이 있어서 베이지안 네트워크를 사용하였으나, 다른 인공지능 방법론이 갖는 장점들이 있으므로, 추후 연구에서는 다른 모델의 장점을 활용한 후속 연구들이 진행될 필요가 있다고 하겠다.

이처럼 한계가 있지만, 본연구는 그동안 정성적인 연구가 주를 이루었던 진급 연구를 정량적으로 접근했다는 점에서 의의가 있다고 할 수 있다. 연구결과에서 중요도가 높은 요소와 배점이 높은 요소가 일치하지 않는다는 결과가 나왔으므로, 배점에 대한 비중 변화나, 평가방법 개선에 본 연구 결과를 활용할 수 있을 것이다. 또한, 장교의 경우 다면평가를 평가항목에 포함시켜 우수한 인재를 선발하는 진급심사의 신뢰성을 높이는 것이 필요하다.

기업에서 인사관리에 인공지능을 활용하려는 노력은 지속되고 있다. 면접자의 1차 선발 과정에서 인공지능을 활용한 면접을 활용하는 사례가 보도되었다. 그 밖에도 인사기록과 다양한 데이터를 종합적으로 평가하기 위해 빅데이터 분석 방법이 활용되고 있다. 진급자 선발에도 인공지능을 활용하면 평가 과정이 신속하고

정확하게 이루어질 수 있을 것으로 기대된다.

군조직의 인사관리도 인공지능을 활용하여, 진급은 물론 희망자가 많은 주요 보직에 적임자 가려내는 데 활용하는 등 활용할 수 있는 분야에서 적극적으로 활용할 필요가 있다. 인구절벽으로 인하여 군조직이 축소되는 것은 피할 수 없는 일이 되었다. 이러한 상황에서 각 조직의 중요도를 판단하여 조직의 개편하는 결정은 몇몇 전문가의 주관적 판단이 아니라 다양한 자료를 분석한 객관적인 분석결과에 기반할 필요가 있다. 지속적으로 변화하는 과정에서 성공적으로 인사관리 업무를 수행하기 위하여 군에서도 인사관리에 인공지능을 도입하려는 노력이 필요하다. 이에 대한 첫걸음은 현재에 대한 점검이어야 하고, 본 연구는 그러한 측면에서 의미가 있다고 생각된다. 머지않아 인공지능이 포함된 인사관리 시스템이 나올 것으로 기대되는데, 데이터를 활용한 인사관리에 관한 후속 연구들이 지속되어 점진적으로 올바른 방향으로 진행되길 기대한다.

참고문헌

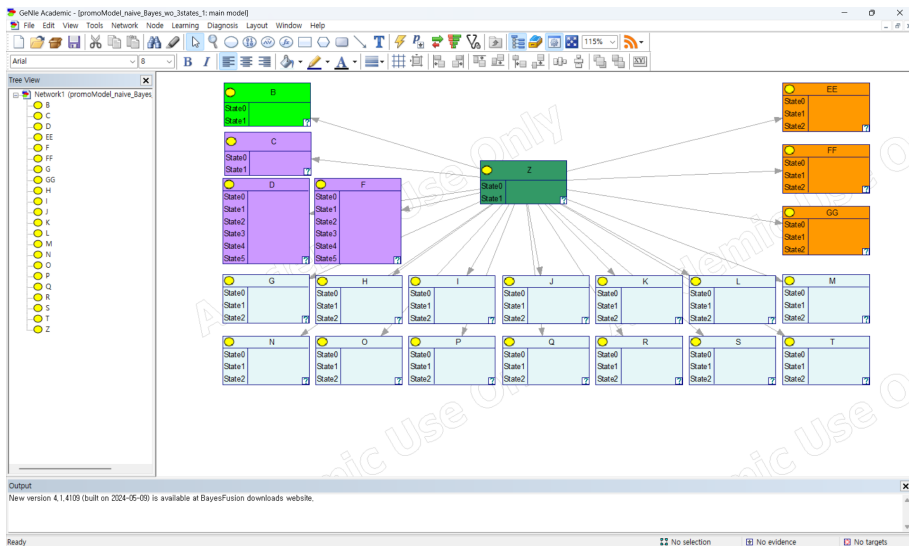
- 강성춘. (2020) 인사이드 아웃, 21세기 북스.
- 공군규정 2-23. (2023). 장병 진급관리.
- 구용희. (2014). 한국군 인사관리 제도의 변천 과정에 관한 연구: 진화론적 특성변화를 중심으로, 목원대학교 박사 학위논문.
- 김선형, 강성춘. (2017). 다면평가 연구 리뷰, 시사점 및 향후 연구 방향. 노사관계연구. 28. 123-153.
- 김애리, 이두열. (2022). 베이지안 네트워크 기반 상황인식 및 방책 결정 지원 모델. 한국 국방경영분석 학회지, 48(1), 131-141.
- 노명화(2011) B-플레이어를 위한 성공의 재발견, 북코리아
- 이두열, 백세일, 김민생, 김신곤. (2021). 베이지안 네트워크 기반 전시 파손항공기 임무수행 능력 예측 모델. 대한기계학회논문집 A, 45(2), 103-111.
- 최병순, 문영세. (2006). 직업군인제 발전을 위한 정책 대안 분석. 한국정책과학학회보, 10(1), 51-78.
- Atwater L. E., & Brett, J. F. (2006). 360-degree feedback to leaders: Does it relate to changes in employee attitudes?. Group & organization management, 31(5). 578-600.
- Borman, W. C. (1997). 360 ratings: An analysis of assumptions and a research agenda for evaluating their validity. Human Resource Management Review, 7(3), 299-315.
- Bayesfusion, LLC. GiNIE Modeler Programmer's Manual, 2023, Available From <https://support.bayesfusion.com/docs/GeNIE> (accessed Feb. 14, 2024)
- Margaritis, D. (2003). Learning Bayesian Network Model Structure from Data, Doctoral dissertation, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA,
- David P. Lepak, & Scott A. Snell, (2002). Examining the Human Resource Architecture: The Relationships Among Human Capital, Employment, and Human Resource Configurations, Journal of Management, 28(4), 517-543,
- Fletcher, (2001). Performance appraisal and management: The developing research agenda. Journal of Occupational and Organizational Psychology, 74(4): 473-487.

- Kiazad, K., Holtom, B. C., Hom, P. W., & Newman, A. (2015). Job embeddedness: a multifoci theoretical extension. *Journal of Applied Psychology*, 100(3), 641-699
- Lepak, D. P., & Snell, S. A. (1999). The human resource architecture: Toward a theory of human capital allocation and development. *Academy of management review*, 24(1), 31-48.
- Moore, D. W., & Trout, B. T. (1978). Military Advancement: The Visibility Theory of Promotion. *The American Political Science Review*, 72(2), 452-468.
- Porter, M. E. (1997). Competitive strategy. *Measuring business excellence*, 1(2), 12-17.
- Rasmussen, C. E. and Williams, C. K. I. (2007) *Gaussian Processes for Machine Learning*, The MIT Press, Cambridge, MA,
- Schuler, R. S. (1989). Strategic human resource management and industrial relations. *Human relations*, 42(2), 157-184.
- Sucar, L. E., Bielza, C., Morales, E. F., Hernandez-Leal, P., Zaragoza, J. H., & Larranaga, P. (2014). Multi-label classification with Bayesian network-based chain classifiers. *Pattern Recognition Letters*, 41, 14-22.
- Tichy, N. M. (1981). Strategic human resource management, *Sloan Management Review*, 23, 47-61.
- Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.
- Scholkopf, B., & Smola, A. J. (2001). *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth.
- Breiman, L. (1996). *Bagging Predictors*. Machine Learning.
- Freund, Y., & Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to Linear Regression Analysis*. Wiley.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

부 록

- 베이저안 네트워크 프로그램 GENIE 매뉴얼 : 기댓값 최대화(EM) 알고리즘 -

1. Step 1: Bayesian Network 모델 생성



〈그림 A-1〉 Bayesian Network 모델 생성

〈그림 A-1〉은 GeNIe 소프트웨어에서 생성한 Bayesian Network 모델로 각 노드는 확률 변수를, 노드 간의 방향성 화살표는 인과 관계를 나타낸다.

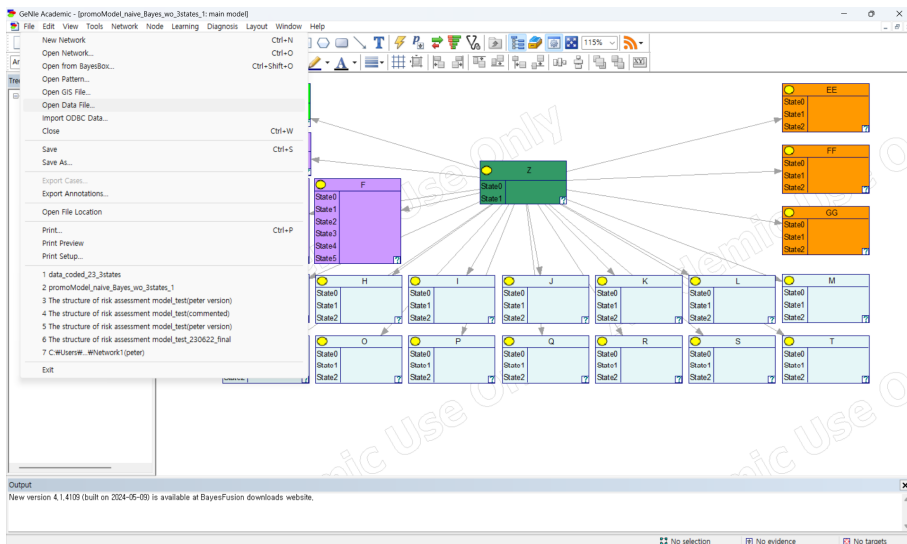
1.1. 노드 생성 및 연결

- GeNIe의 툴바에서 노드 추가 버튼을 사용해 변수(노드)를 생성한 후, 각 노드 간의 인과 관계를 설정한다.

1.2. 네트워크 구조 정의

- 노드들은 다양한 상태(State)를 가질 수 있으며, 네트워크 구조는 모델링하려는 문제에 따라 다르게 설정된다.

2. Step 2: 데이터 파일 불러오기



〈그림 A-2〉 데이터 파일 불러오기

이 단계에서는 Bayesian Network에 사용할 데이터를 GeNIe에 불러오는 과정이다.

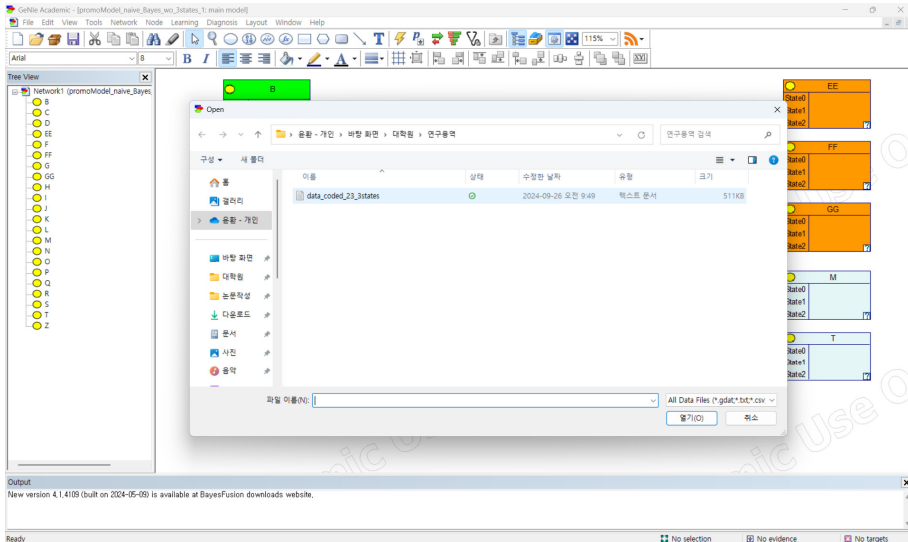
2.1. File 메뉴 열기

- 상단 메뉴에서 File을 클릭한다.
- 드롭다운 메뉴에서 'Open Data File...'을 선택한다. 이 옵션은 네트워크에 적용할 데이터 파일을 불러오기 위한 기능이다.

2.2. 데이터 파일 선택

- 'Open Data File...'을 클릭하면 파일 탐색 창이 열리며, 여기서 데이터 파일(CSV, TXT 등)을 선택해 불러올 수 있다.
- 선택된 데이터는 Bayesian Network에서 학습에 사용되며, EM 알고리즘을 통해 파라미터 학습에 활용된다.

3. Step 3: 데이터 파일 선택



〈그림 A-3〉 데이터 파일 선택

이 단계에서는 학습에 사용할 데이터 파일을 선택하는 과정이다.

3.1. 파일 선택 창

- GeNIe에서 ‘Open Data File...’을 클릭하면 파일 탐색 창을 연다.

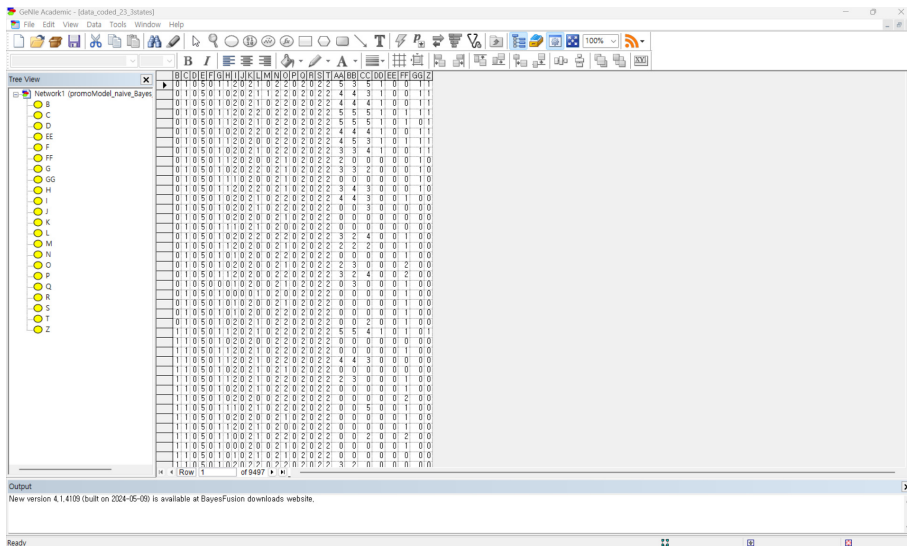
3.2. 데이터 파일 선택

- 네트워크에서 학습에 사용할 데이터 파일을 선택한다. 이 예시에서는 ‘data_coded_23_3states’라는 파일을 선택했다.

3.3. 파일 형식

- 하단의 파일 형식 옵션에서 CSV, TXT와 같은 형식을 지원하며, 올바른 데이터 파일 형식을 확인하고 선택할 수 있다.

4. Step 4: 데이터 불러오기 완료



〈그림 A-4〉 데이터 불러오기 완료

이 단계에서는 선택한 데이터 파일이 GeNIe에 성공적으로 불러와진 상태를 보여준다.

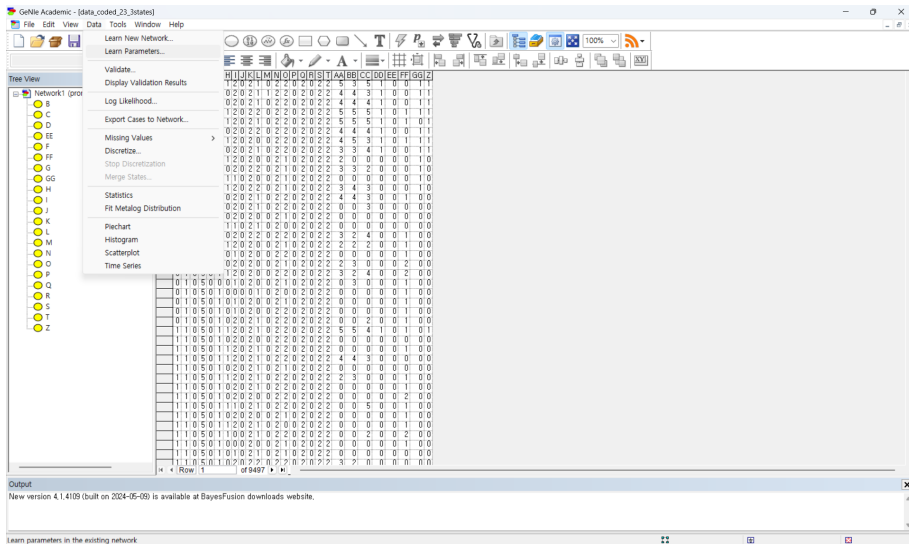
4.1. 데이터 확인

- 불러온 데이터는 각 열 (column)이 네트워크 변수에 해당하며, 각 행 (row)은 개별 데이터 레코드를 나타낸다.
- 이 데이터는 Bayesian Network의 학습에 사용될 관측치이며, 각 값은 해당 변수의 상태 (State)를 나타낸다.

4.2. 데이터 탐색

- 데이터가 제대로 불러와졌는지 확인한다. 모든 변수와 레코드가 정확하게 표시되어 있어야 하며, 데이터가 손상되거나 누락되지 않았는지 점검하는 것이 중요하다.

5. Step 5: 파라미터 학습 시작



〈그림 A-5〉 파라미터 학습 시작

이 단계에서는 EM 알고리즘을 통해 파라미터 학습을 수행하기 위한 메뉴를 선택하는 과정이다.

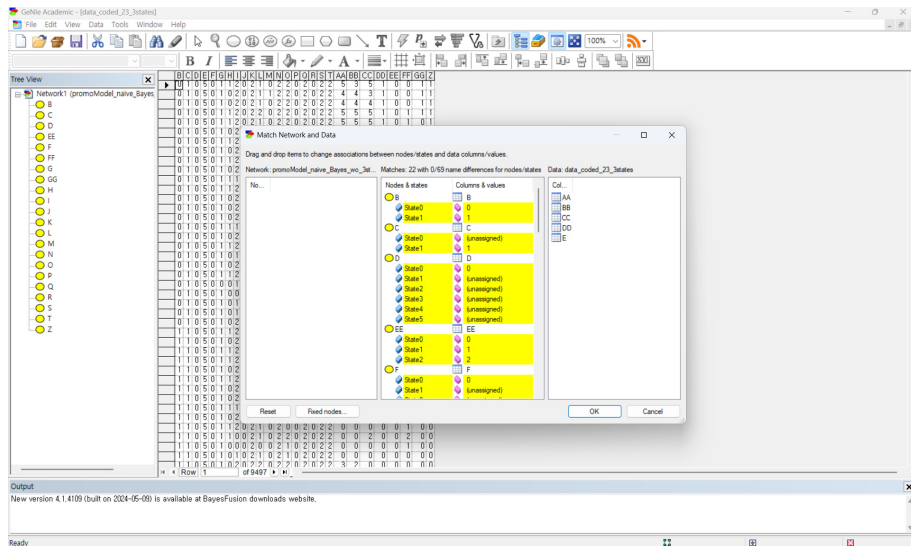
5.1. Learning 메뉴 열기

- 상단 메뉴에서 Learning을 클릭한다.

5.2. Learn Parameters 선택

- 드롭다운 메뉴에서 'Learn Parameters...'를 선택한다. 이 옵션은 현재 불러온 데이터를 바탕으로 네트워크의 파라미터 (확률 값)를 학습하는 기능이다.
- EM 알고리즘을 사용해 각 변수 간의 조건부 확률을 추정할 수 있다.

6. Step 6: 네트워크와 데이터 매칭



〈그림 A-6〉 네트워크와 데이터 매칭

이 단계에서는 네트워크의 변수와 데이터 파일의 열을 매칭하는 작업을 수행한다.

6.1. Match Network and Data 창

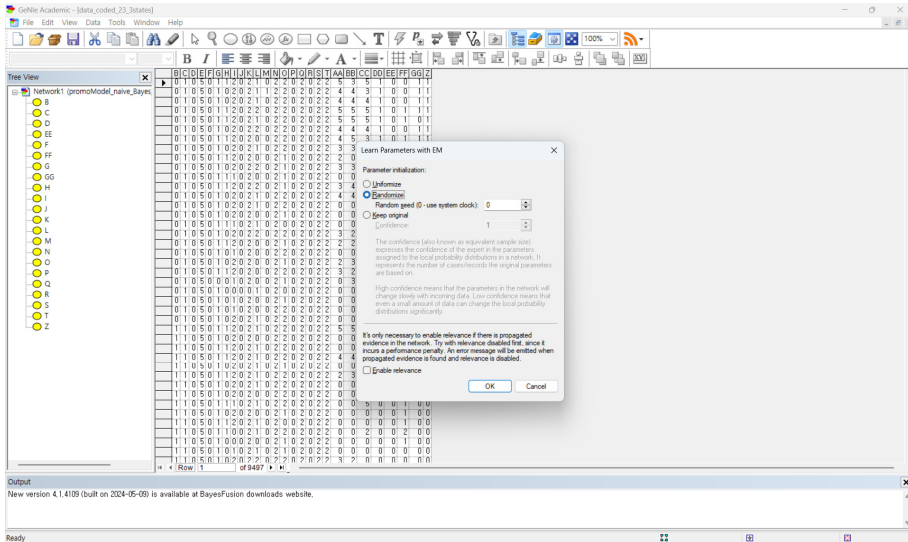
- 이 창은 '네트워크에서 사용되는 노드'와 데이터 파일에서의 열의 매칭을 설정한다.
- 노드의 각 상태 (State) 가 데이터 파일의 어떤 값과 대응되는지를 지정한다.

6.2. 매칭 설정

- 각 노드의 상태 (State) 가 데이터 파일의 어느 값에 대응되는지 드래그 앤 드롭 방식으로 매칭을 설정한다. 예를 들어, 노드 "B"의 상태 0과 1이 데이터 파일의 값 0과 1에 매칭되도록 설정해야 한다.

※ 일부 상태는 '(unassigned)'로 표시되며, 이는 아직 데이터와 연결되지 않았다는 의미이다.

7. Step 7: 파라미터 학습 설정



〈그림 A-7〉 파라미터 학습 설정

이 단계에서는 EM 알고리즘을 사용하여 파라미터 학습을 수행하기 위한 초기 설정을 지정하는 단계이다.

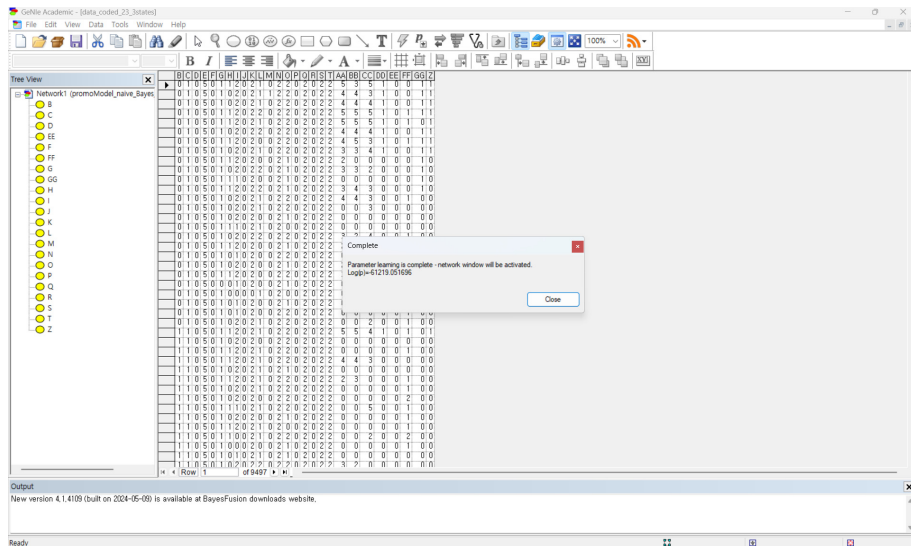
7.1. Parameter Initialization (파라미터 초기화)

- Uniformize: 모든 파라미터를 균등하게 초기화한다.
- Randomize: 파라미터를 무작위로 초기화한다. Random seed를 '0'으로 설정하면 시스템 클럭을 기준으로 랜덤 초기값이 생성되고 이는 재현 가능한 결과를 위해 특정 시드 값을 설정할 수 있는 옵션이다.
- Keep original: 기존의 파라미터 값을 유지한 상태에서 학습을 진행한다.

7.2. Confidence (신뢰도)

- 신뢰도는 데이터가 모델에 미치는 영향을 결정한다. 신뢰도가 높을수록 데이터가 기존의 파라미터 값을 천천히 변경하게 되며, 신뢰도가 낮을수록 데이터가 큰 영향을 미친다.

8. Step 8: 파라미터 학습 완료



〈그림 A-8〉 파라미터 학습 완료

이 단계에서는 EM 알고리즘을 통한 파라미터 학습이 완료된 상태를 나타낸다.

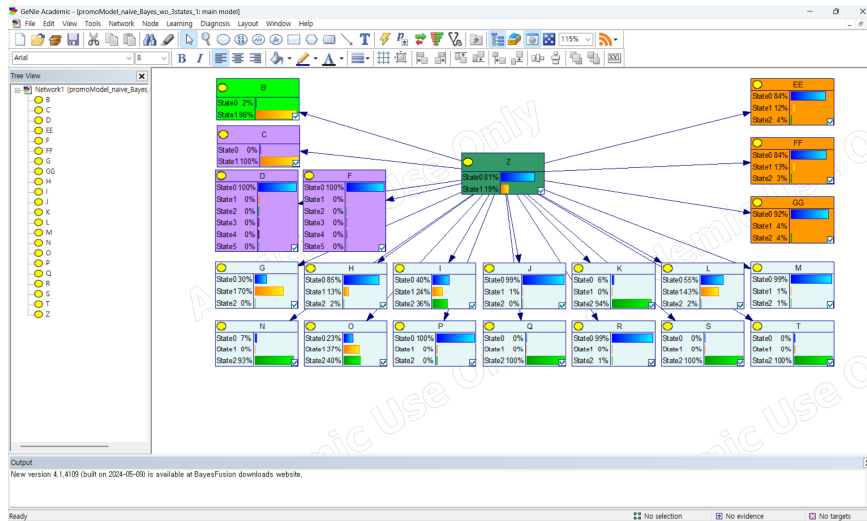
8.1. 완료 메시지

- "Parameter learning is complete": 파라미터 학습이 성공적으로 완료되었음을 알리는 메시지이다. 이제 학습된 Bayesian Network가 활성화 완료되었다.

8.2. 로그 우도 값 (Log-likelihood)

- $\text{Log}(p) = -61219.051696$: 이는 학습된 모델의 로그 우도(Log-likelihood) 값이다. 이 값은 학습된 모델이 주어진 데이터에 얼마나 잘 맞는지를 나타내며, 일반적으로 로그 우도가 클수록 모델이 데이터를 더 잘 설명한다는 의미이다.

9. Step 9: 학습 완료 후 네트워크 결과 확인



〈그림 A-9〉 학습 완료 후 네트워크 결과 확인

EM 알고리즘을 사용한 파라미터 학습 결과를 확인할 수 있다. Bayesian Network의 각 노드에 대한 조건부 확률 분포가 표시된 상태이다.

9.1. 확률 분포 확인

- 각 노드의 상태(State)에 대한 확률이 색상으로 시각적으로 표시된다.

9.2. 조건부 확률 테이블 (CPT)

- 각 노드는 학습된 데이터를 바탕으로 ‘조건부 확률 테이블(CPT)’을 계산한다. 이 CPT는 노드 간의 관계와 데이터 간의 상호작용을 바탕으로 추론된 결과다.

9.3. 네트워크 분석

- 학습된 모델을 기반으로 추가 분석을 수행할 수 있으며, 네트워크에서 특정 노드의 상태 변화에 따른 다른 노드의 영향을 시뮬레이션할 수 있다.
- 예측 및 추론 작업을 통해 모델의 성능을 확인하고, 새로운 데이터에 대해 예측할 수 있다.

유의사항

1. 본 연구보고서 내용은 연구진의 개인적인 견해로서 국방대학교 국가안전보장문제연구소의 공식입장과 다를 수 있습니다.
2. 본 연구보고서는 정책입안시 참고자료로만 활용하고 타 기관에 불필요한 자료유출을 삼가 주시기 바랍니다.

군사과학정책연구

제17권

2024년 12월 29일 인쇄

2024년 12월 31일 발행

저 자 : 장용진, 이재홍, 박경욱, 신진수,
양경민, 조윤환, 이현석

발 행 처 : 국방대학교 국가안전보장문제연구소
TEL. (041) 831-6414
FAX. (041) 831-0000

인 쇄 : 청맥기획 (042) 487-2589

ISSN 1976-5967

