Journal of the Ergonomics Society of Korea
Vol.21, No.1, pp.1-14, April 2002

단계별 비행훈련 성패 예측 모형의 성능 비교 연구

Comparison of Classification Models for Sequential Flight Test Results

손소영*, 조용관*, 최성욱**, 김영준**

ABSTRACT

The main purpose of this paper is to present selection criteria for ROK Airforce pilot training candidates in order to save costs involved in sequential pilot training. We use classification models such as Decision Tree, Logistic Regression and Neural Network based on aptitude test results of 288 ROK Air Force applicants in 1994-1996. Different models are compared in terms of classification accuracy, ROC and Lift-value. Neural network is evaluated as the best model for each sequential flight test result while Logistic regression model outperforms the rest of them for discriminating the last flight test result. Therefore we suggest a pilot selection criterion based on this logistic regression. Overall, we find that the factors such as Attention Sharing, Speed Tracking, Machine Comprehension and Instrument Reading Ability having significant effects on the flight results. We expect that the use of our criteria can increase the effectiveness of flight resources.

Keyword: Flight aptitude, Forecasting model, Decision Tree, Logistic Regression, Neural Network

* 연세대학교 컴퓨터과학, 산업시스템공학과
** 공군사관학교 산업공학과

Copyright (C) 2003 NuriMedia Co., Ltd.

1. 서 론

인간의 적성과 능력은 후천적인 노력에 의해 개발될 수 있으나 선천적인 자질에 크게 의존한다. 따라서 주어진 직무 수행에 요구되어지는 주요 요인과 개인의 적성을 논리적으로 분석하여 적절한 자원을 선별한다면 주어진 목표를 보다 효과적으로 달성할 수 있을 것이다. 비행이라는 특수한 상황, 특히 전투기 조종사에게는 일상에서 요구되지 않는 고도의 신체적 적응 능력과 순간적인 판단능력이 요구되기에 복잡한 선발 및 혹독한 훈련 과정이 필요하다. 군 관련 기관 및 일부 학회에서 조종 업무와 관련된 인간기능 요소를 도출하고 이러한 기능에 대한 수행도를 측정할 수 있는 검사항목을 개발하여 조종사 선발에 적용하려는 연구가 진행되어 왔다. 그러나 현재의 조종사 선발과정으로 미루어볼 때 비행훈련의 성패를 정확히 예측할 수 있는 모형의 개발이 미흡한 상태에 있으며 이를 적용하고자 하는 정책결정자의 의지도 부족하다고 판단된다. 과거와 현재의 조종사 양성체계는 학업성적이 우수하고 신체검사에 합격한 예비 조종사들을 대상으로 선 입과, 후 양성의 체제를 고수하고 있다. 그러나 비행이라는 특수한 상황이 이러한 몇 가지 조건만을 요구하는 것은 아니다. 조종사 양성을 위해서는 엄청난 예산이 소요되며, 특히 부적합한 조종사가 비행훈련을 수료하여 전투기 조종사로서 임무 중 인적 요인에 의한 사고를 유발 시는 단순히 비용의 문제를 떠나 그 파급 효과가 상상을 초월한다. 따라서 비행훈련을 무사히 수료하고 비행임무를 훌륭히 수행할 수 있는 대상자를 선발할 수 있는 모형을 제시한다는 것은 군의 발전과 국가예산절약 등 다양한 측면에서 중요한 의미를 갖는다고 볼 수 있을 것이다.

공군에서는 1963년도에 미 공군장교 적성검사(Air Force Offer Qualifying Test : AFOQT)를 번역하여 사관생도 및 간부 후보생들을 선발하는 자료로 사용하였으나, 보다 심층적인 연구의 발전은 이루어 지지 않았으며 적성검사 결과 자료를 조종사 선발에 적용한 기록은 없다. 그 이후 적성에 관한 연구는 더 이상 추진되지 못하다 90년대 초에 서서히 다시 태동하기 시작했다. 이러한 적성검사에 대한 연구 노력의 일환으로, 지필검사와 더불어 심리기능 검사를 실시하기 위한 여러 가지 관련 장비에 대한 개발이 시도되었다. 대표적인 연구로 이상원 (1992)은 컴퓨터를 이용한 조종사 선발 적성검사에서 심리기능 검사를 위한 도구를 개발하였으나 실행의 단계까지 이르지는 못하였다. 비슷한 시기에 이달호 (1992)는 조종사 선발용 검사기구의 개발에 관한 연구에서 기구를 이용한 심리기능검사 방법을 개발하였다. 이 장비는 1994년에 연구를 위하여 활용이 되었으나 실제로 조종적성을 판단하는 장비로는 사용되지 않았다. 이렇게 개발된 심리기능 검사용 기구들은 공군 내부에서 조종사가 되기 위한 지원자들의 조종적성을 연구하기 위한 획기적인 연구 방법이었으나 이 검사법들은 개발되고 한 차례 활용된 후 지속적 검사결과의 추적 및 후행연구가 이루어지지 않아 많은 아쉬움을 남겼다. 한편 Shub, et al. (1994)은 조종사 평가시스템 (PES : Pilot Evaluation System)을 이용하여 조종사

Copyright (C) 2003 NuriMedia Co., Ltd.

훈련 입과 후보생의 적성을 사전에 평가하고 기종 (전투기, 헬리콥터, 수송기, 경비행기) 분류를 위한 기준으로 적용하고자 하였다. 작전 임무 수행을 위해 요구되어지는 목표물 접촉, 고도/속도유지, 교신내용청취능력 등을 평가하기 위해 1차 과제와 2차 과제 그리고 이들 과제를 조합한 복합적인 임무를 부여하였다. 이러한 임무의 처리 능력을 통하여 전반적인 임무 수행가능성, 주의 집중력 및 분배력, 주의력 전환능력 그리고 주어진 상황 하에서 우선 순위에 따른 임무 배분능력을 측정함으로써 전투기 조종사에게 요구되어지는 적성을 판단하고자 하였다. 그러나 실제 비행상황 하에서는 비행자질뿐 아니라 정신적, 육체적 요인 등이 종합적으로 요구되기 때문에 비행훈련 결과를 모의비행 결과만으로 판단하기에는 무리가 따른다. 따라서 비행과 관련된 모든 변수를 고려하여 종합적인 비행적성을 판단해야 할 것이다.

본 연구에서는 1994-1996년에 입과한 조종학생 288명의 31개 적성검사항목과 수료결과 자료를 바탕으로 비행 단계별 및 전 비행과정에 대한 비행훈련 성공 및 실패 예측 모형을 구축하여 각 단계별 순차적인 비행 특성과 전 비행과정에 따른 결과 예측 모형을 비교 분석하였다. 또한 의사결정나무, 로지스틱회귀모형, 신경망 등의 다양한 기법을 적용, 결과를 비교 분석함으로써 최적의 모형을 제시하였다. 이러한 분석 결과를 바탕으로 비행훈련 입과자 선발을 위한 기준을 제시하여 비행훈련에 입과 하는 자원의 양질화 및 비행훈련의 효율화 증대를 도모하고자 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 연구방법 및 자료의 성격을 소개하였으며 3장에서는 의사결정나무, 로지스틱회귀모형, 신경망을 적용하여 각 단계별 및 비행 전 과정에 걸쳐 성패 여부를 결정하는 모형을 추정하였다. 더불어 각 모형의 성능을 분류정확도 및 ROC, Lift-Value 관점에서 비교 평가하였다. 또한 이러한 분석 결과를 바탕으로 최적의 비행훈련 성공 및 실패 예측 모형을 제시하였으며 4장에서는 결론 및 본 연구의 발전방향을 제시하였다.

2. 연구방법

2.1 자료

공군 전투기 조종사가 되기 위해서는 비행훈련 입과 후 초등-중등-고등의 단계적인 비행 절차를 모두 수료해야 한다. 즉, 초등비행훈련 입과 후 전 단계의 비행을 무사히 수료해야 다음 단계의 훈련에 입과 할 수 있는 자격이 주어지며 3단계의 비행을 모두 수료해야 조종사의 자격이 주어진다. 조종사가 되기 위해 매년 3-4개 차수의 훈련 후보생이 입과하며 각 차수는 60여명으로 구성되어 있다. 본 연구의 대상은 1994년부터 1996년도에 조종사 양성 비행 훈련에 입과한 9개 차수 (94-2,3차, 95-1,2,3차, 96-1,2,3,4차) 539명 중 본 연구를 위한 5개 분야의 적성검사를 모두 실시한 288명의 훈련생이다. 이들은 세부적으로 최초 초등비행훈련 입과자 288명, 초등비행 수료후 중등비행훈련 입과자250명 그리고 중등비행 수료후 고등비행 입과자 181명으로 구분된다.

Copyright (C) 2003 NuriMedia Co., Ltd.

비행자질검사 및 심리기능검사는 조종사 적성 연구실에서 실시 되었으며 나머지 항목에 대한 자료는 공군사관학교 생도대 및 212비행대대에서 실시한 자료를 인계 받아서 활용하였다. 본 연구를 위하여 사용된 5개 적성검사 분야의 변수별 31개 항목은 표1과 같다. 채력검사 5개 항목 및 지능검사 6개 항목은 일반적으로 실시하고 있는 검사와 동일하므로 변수성격에 대한 설명은 생략하였다.

표 1. 각 항목별 변수

검사 종류	변수	항목	변수성격
심리 기능 검사	X1	수지 정밀도	계기판상의 원하는 조작기기를 신속정확히 제어할 수 있는 능력을 판단
	X2	추적 능력	공중 기동시 HUD나 MFD에 제시되는 목표물과 비행정보에 대한 정밀한 추적 및 비행조작의 능력을 판단
	X3	속도 추적능력	비행시 접근율, 침하율 등의 변동속도를 정확히 추정하여 공간거리를 예측하는 능력을 판단
	X4	단기 기억능력	비행 중 교신내용을 기억하고 지시된 대로 임무를 수행할 수 있는가를 판단
	X5	주의 분배력 1	비행 중 입력되는 많은 정보에 대해 주의력을 적절히 배분하여 이를 효과적으로 처리할 수 있는 능력을 판단
	X6	주의 분배력 2	
비행 자질	X7	수표해독	조종석의 많은 계기판을 정확히 판독하고 적절한 상황판단을 할 수 있는 능력을 판단
	X8	전기미로	비행 상황 하에서 명확한 판단을 할 수 있는 능력을 판단
	X9	토막수 세기	3 차원의 비행 상황 하에서 조종사에게 필수적인 공간지각능력을 판단
검사	X10	척도판독	복잡한 계기를 판독할 수 있는 능력을 판단
	X11	공구용도	많은 계기 중 필요한 계기를 선택하여 판독할 수 있는 능력을 판단
	X12	기계원리	항공기 조작에 필요한 각종 기계의 작동원리를 이해할 수 있는 능력을 판단
	X13	계기판독	항공기 자세를 계기로 판독할 능력이 있는가를 판단
체력 검사	X14	100m 달리기	
	X15	1500m 달리기	
	X16	멀리뛰기	
	X17	윗몸 일으키기	
	X18	턱걸이	
인성 검사	X19	감수성	자신의 욕구와 감정에 대한 민감도를 측정
	X20	자발성	자신의 감정을 행동으로 표현할 수 있는 능력
	X21	자기 긍정성	자신의 가치/장점 및 자존심에 대한 지각 정도 측정
	X22	자기 수용성	자신의 약점/결점을 수용할 수 있는 능력
	X23	인간관	인간의 본성에 대한 지각도
	X24	포용성	타인과의 친근한 인관관계
	X25	자아 성취도	현재 소속한 조직이나 집단내에서 자기가 많은 임무를 수행함에 있어 자기 능력을 발휘 하고 계발하여 자기의 이상을 성취하려는 정도
지능 검사	X26	어휘적용	
	X27	언어추리	
	X28	산수추리	
	X29	수열추리	
	X30	도형추리	
	X31	종 합	

모형구축의 효과를 높이기 위하여 모형화 단계 이전에 데이터에 대한 기초통계, 이상치 및 결측치를 확인하여 데이터를 수정하였다. 또한31개 변수의 상관관계를 이용하여 이해하기 쉬운 형태로 변수를 축소하고 각 변수들간의 측정개념의 유사성을 파악하기 위해 요인분석을 실시하였다. 요인적재량 50(상관관계 0.5)과 최소고유값(MINEIGEN) 1.5로 하여 요인분석을 실시한 결과 6개의 요인이 추출되었다. 요인1은 X31, X29, X10, X28로 적성검사 종류 중 지능검사 항목이 다수 포함되었다. 요인2는 X24, X19, X22, X20, 요인3은 X16, X18, X14, 요인4는 X25, X23, 요인5는 X15, X17, X2, 요인6은 X6, X11의 변수 항목으로 분류되었다. 이 결과는 요인6을 제외하고 표1의 검사종류별 항목의 구분과 유사하다고 볼 수 있다. 요인에 의해 새로 형성된 변수의 명칭은 각 요인별 구성 항목의 특성을 고려하여 정하였으며, 요인1은 지능, 요인2는 감정표현력, 요인3은 신체 반응속도, 요인4는 자기인식, 요인5는 지구력, 그리고 요인6은 비행자질로 명명하였다. 이렇게 축약된 6개의 요인은

Copyright (C) 2003 NuriMedia Co., Ltd.

비행 훈련 결과 예측을 위한 설명변수로 이용하였으며 31개 변수를 설명변수로 하였을 경우와 모형의 성능을 비교하여 예측력이 우수한 모형을 구축하고자 하였다.

표 2. 요인분석결과

Rotated Factor Pattern

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5	FACTOR6
X31	84 *	16	-15	-36	3	2
X29	64 *	17	-18	5	1	-4
X10	60 *	-7	-2	-8	-5	-13
X28	54 *	-8	2	-34	6	9
X30	48	4	10	7	11	19
X9	48	-5	13	25	8	-18
X8	46	-9	23	42	-4	5
X12	41	-26	18	0	-5	39
X24	2	85 *	-2	11	-2	12
X19	7	78 *	7	11	-1	4
X22	1	69 *	1	-11	-3	-5
X20	-7	67 *	7	18	-12	0
X4	-7	-32	-20	31	-14	16
X16	1	4	78 *	13	-20	12
X18	-4	13	66 *	-1	23	-2
X14	2	-3	-75 *	8	23	-10
X25	-1	3	4	66 *	3	-6
X23	9	21	-27	53 *	7	6
X21	-4	38	11	48	6	-10
X27	32	15	-12	-40	9	-12
X26	38	10	-20	-48	-14	2
X15	20	8	-14	18	65 *	12
X17	6	-2	38	-18	63 *	-20
X1	18	11	-7	0	-28	2
X7	33	-7	8	-3	-45	-6
X5	-18	6	0	20	-45	34
X2	-11	6	13	-7	63 *	-20
X6	-7	-2	17	-8	2	66 *
X11	10	-4	-7	4	-2	63 *
X13	8	-6	-18	-3	-16	-23
X3	26	-14	16	14	-16	-36

2.2 분석 과정

본 연구는 과거에 실시하였던 적성검사 결과 자료를 학습용과 검증용 자료로 구분하여 의사결정나무, 로지스틱회귀모형, 신경망 등의 분류기법을 적용하였다.

의사결정나무를 위해서는 카이제곱 통계량의 P-value에 의해 분리규칙이 형성되는 CHAID알고리즘을 사용하였다. 로지스틱회귀모형은 목표변수가 이항형일 때 선형회귀모형의 단점을 극복하기 위해 확률에 대한 로짓변환 (logit transformation)을 고려하여 분석하는 것으로 목표변수의 범주가 2개일 때는 다중 로지스틱회귀모형을 사용하였으며 목표변수의 범주가 3개 이상일 때는 폴리토모스(Polytomous) 로지스틱회귀모형을 사용하였다.신경망의 구조는 그림1과 같이 2개의 은닉층(Hidden Layer) 과 3개 및 2개의 뉴런(Neuron)을 갖는 MLP구조를 사용하였다.

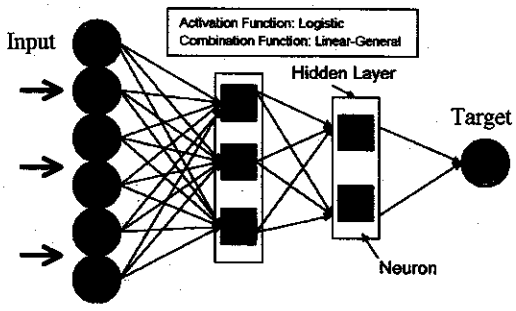


그림 1. 신경망 구조

이러한 분류기법을 평가하고 비교하기 위한 한 가지 전략은 데이터를 학습용 (Training)과 검증용 (Test)으로 분할하여, 분석용 데이터를 이용하여 모형을 구축하고 검증용 데이터를 이용하여 모형의 비교나 최종적인 평가를 수행하는 것이다. 학습용과 검증용 데이터의 분할은 분류모형에서 가장 일반적으로 적용하는 학습용 60%, 검증용 40%로 하였다 (강현철 등, 1999). 학습용 자료를 이용, 비행훈련을 성공적으로 이수할 수 있는가를 판단하는 모형을 추정함에 있어 비행훈련단계별 및 전 단계로 나누어 분석하였다. 각 단계별 순차적인 분석은 비행 자료를 초등, 중등, 고등비행 과정으로 구분하여 각 과정별 수료여부를 결정하는 주요변수 관점에서 실시하였다. 비행 전 과정에 대하여는 두 가지 모형을 고려하였는데 초등비행 입과후 고등비행 수료까지의 결과를 수료와 도태 두개의 범주로

Copyright (C) 2003 NuriMedia Co., Ltd.

구분한 모형과 초등도태, 초등수료, 중등수료, 고등수료의 4개 범주로 구분 분석한 모형의 형태를 취함으로써 각 과정별 중간 과정의 수료 여부를 분석하였다. 이러한 모형을 수립할 때 설명변수로는 이미 소개된 31개 변수와 더불어 요인 분석을 통하여 얻은 축소된 차원의 요인을 이용하여 각 모형의 성능을 비교 평가하였다.

모형평가란 예측을 위해 만든 모형이 임의의 모형보다 과연 우수한지, 고려된 서로 다른 모형들 중 어느 것이 가장 우수한 예측력을 보유하고 있는지를 비교, 분석하는 과정이라 말할 수 있다. 분류모형의 성능을 평가하기 위한 여러 방법들이 있지만, 본 연구에서는 목표변수가 범주형일 때 가장 널리 이용되는 분류정확도, ROC 및 Lift-Value를 사용하였다 (강현철 등,1999). 특히 목표변수의 범주가 한 곳에 편중되어 있을 경우 분류정확도만으로 모형의 성능을 평가하는 데는 무리가 있으므로 ROC 와 Lift-Value를 함께 고려하여 구축된 모형을 평가함으로써 최적의 모형을 제시하고자 하였다. 이 값들을 구하는 과정은 다음과 같다.

분류정확도와 민감도, 특이도는 오분류표를 이용하여 구할 수 있다. 오분류표는 실제 범주와 모형에 의해 예측된 분류범주 사이의 관계를 나타내는 값으로 이는 목표변수의 범주가 c개인 경우 c c개의 셀로 이루어진 표 형식을 취한다. 오분류표의 빈도수는 모형에 의해 계산되는 사후확률의 경계가 되는 분류기준값 (Cut off-value)에 의해 결정된다. 즉, 분류 기준값을 변경할 때 마다 오분류표의 빈도수는 변하게 된다. 따라서 분류기준값을 정할 때는 분류정확도가 가장 높은 값을 선택한다. 목표변수의 범주수가 2개인 경우의 오분류표를 예를 들어 설명하면 다음과 같다.

		예측범위	
		0	1
실제 범주	0	a	b
	1	c	d

표 3. 목표변수가 2개인 오분류표

표3 에서 대각에 존재하는 셀은 목표변수와 원래 범주가 제대로 분류된 부분이고, 비대각에 존재하는 셀은 분류가 잘못된 부분으로 구분되어 각각의 빈도가 제시된다. 따라서 분류정확도는 (a + d)/(a + b + c + d)의 값이된다.

ROC (Receiver Operating Characteristic Curve)곡선은 신호감지이론 (Signal Detection Theory)으로부터 출발한 개념으로 구축한 모형의 성능을 민감도와 특이도에 의해 판단하고자 하는 곡선이다. 표3에서 목표변수를 1로 하였을 때 민감도는 (실제1, 예측1)인 관찰치의 빈도/실제 1인 관찰치의 빈도 즉, d/(c + d)가 되며 특이도는 (실제0, 예측0)인 관찰치의 빈도/실제 0 인 관찰치의 빈도 즉, a/(a + b)가 된다.

ROC곡선을 위해서는 각 분류기준값에서 오분류표가 필요하다. 각 분류기준값에서의 오분류표에 의해 민감도와 특이도의 값을 구하고 이 값을 이용하여 X축에는 1-특이도를 Y축에는 민감도를 각 분류기준값에 대응시킴으로 구할 수 있다. Lift-Value는 모형 평가에 있어서 사후확률을 사용하기는 하나, ROC와

Copyright (C) 2003 NuriMedia Co., Ltd.

는 다른 개념으로써 다음과 같은 절차에 의해 구해진다. 먼저 모형설정을 통해 사후확률을 구하고 사후확률값의 순서에 따라 전체 데이터를 정렬한다. 다음 전체 데이터를 균일하게 N등분하고 N등분의 각 집단에서 목표변수의 특정 범주에 대한 빈도를 구한다. N등분의 각 집단에서 Lift-Value통계량은 해당 집단에서 목표변수의 빈도/전체에서 목표변수의 특정범주 빈도*0.1로 구할 수 있다. 이러한 결과를 종합 분석하여 분류정확도, ROC, Lift- value, 입력변수의 수, 일반화의 가능성 등을 고려한 최적의 분류모형을 제시하였다.

3. 모형비교분석

본 절에서는 최적의 분류 모형을 얻기 위해 각 단계별(초등, 중등, 고등) 순차적인 결과 및 전 과정에 걸친 비행결과 자료에 대하여 의사결정나무, 로지스틱회귀모형, 신경망을 적용하여 그 결과를 분류정확도 및 ROC 와 Lift-Value관점에서 비교 평가하였다. 또한 요인 분석에 의해 축소된 6개의 factors를 이용한 분석과 31개의 모든 변수를 설명 변수로 하여 분석한 결과를 비교 평가하였다. 그 결과 고등비행 결과를 제외한 모든 분석에서 31개의 변수를 이용 시 우수한 결과를 보였다. 따라서 고등비행결과를 제외한 모든 자료에 대하여 31개 변수를 설명변수로 하여 분석을 실시하였다. 이러한 결과를 종합 분석하여 결과 예측을 위한 최적의 분류모형을 제시하였다.

3.1 비행 단계별 순차적인 결과 분석

각 비행 단계별 (초등, 중등, 고등) 성공 여부를 결정짓는 주요 설명 변수를 분석하고 적용 기법간의 성능을 비교 평가하기 위하여 순차적인 분석을 실시하였다. 초등비행훈련자료 분석 결과 분류정확도에서는 의사결정나무가 94%로 가장 높았으며 신경망이 91.3%, 로지스틱회귀모형이 87%로 나타났다. 민감도, 특이도를 고려한 ROC도표를 이용하여 각 모형의 성능을 비교 평가한 결과 신경망의 성능이 가장 우수한 것으로 나타났다. 결과 예측을 위한 주요 설명 변수로는 의사결정나무의 경우 수표해독 (x7), 공구용도 (x11)가 신경망의 경우 속도추적능력 (x3)외 4개 변수가 선별되었다.

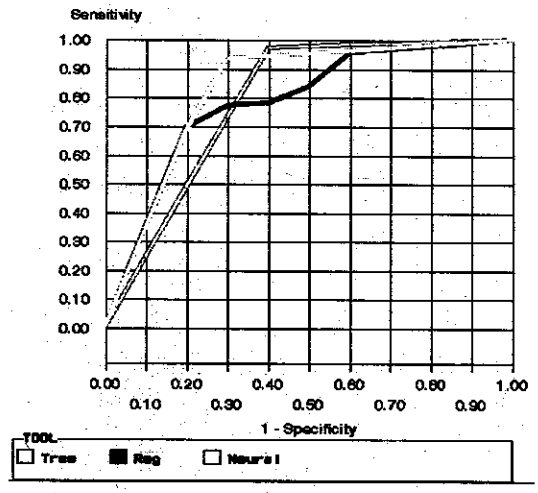


그림 2. 초등 ROC 도표

중등비행훈련자료 분석 결과 분류정확도 및 ROC관점에서 신경망이 가장 우수한 모형으

Copyright (C) 2003 NuriMedia Co., Ltd.

로 평가되었다. 분류정확도는 신경망이 82%, 로지스틱회귀모형이 79%, 그리고 의사결정나무는 78%로 나타났다. 의사결정나무의 경우 모든 관측치를 수료로 분류하였으며 ROC도표 분석결과 모형 구축의 의미가 없는 것으로 분석되었다. 결과 예측을 위한 주요 설명 변수로는 로지스틱회귀모형의 경우 추적능력(x2)외 4개 변수가 선별되었다.

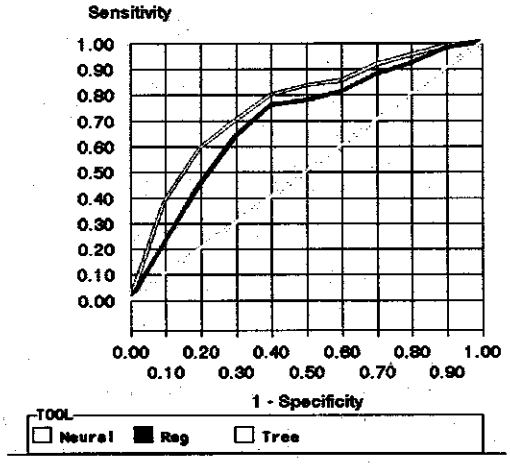


그림 3. 중등 ROC 도표

고등비행훈련자료 분석 결과 ROC관점에서 신경망이 다소 우수한 모형으로 나타났으나 전반적으로 모형 구축의 의미가 없는 것으로 분석되었다. 따라서 요인분석에 의해 축소된 6개 변수를 이용하여 분석을 실시하였다. 그 결과 신경망과 로지스틱회귀모형이ROC관점에서 의미 있는 모형으로 분석되었다. 분류정확도는 신경망 92.8%, 로지스틱회귀모형 91%로 나타났으며 의사결정나무는 모든 관측값을 수료로 분류하였고 분류정확도는 91.7%로 나타났다. 결과 예측을 위한 주요 설명 변수로는 로지스틱회기모형의 경우 인간관 (x23)과 자아성취도(x25)에 의해 축약된 자기인식과 주의 분배력2 (x6)와 공구용도 (x11)에 의해 축약된 비행자질이 선별되었다.

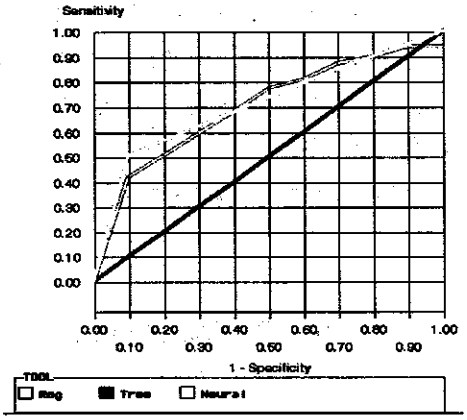


그림 4. 요인분석 결과를 이용한 고등 ROC 도표

이상의 결과를 요약하면 표4와 같다.

표4의 수료율 항목은 각 비행 훈련 단계에서의 실제 입과 인원 대비 수료한 인원의 비율로써 각 분류기법의 분류정확도와 비교하기 위해 제시하였다.

각 비행단계의 목표변수 범주가 한 곳(수료)에 편중되어 있으므로 분류 정확도로 모형의 성능을 판단하기에는 무리가 있으며, 민감도와 특이도를 고려한ROC 관점에서 모형의 성능 평가 시 신경망이 가장 우수한 것으로 나타났다. 신경망모형이 우수한 모형으로 평가되었으나 해석상의 문제점으로 로지스틱회귀모형과 의사결정나무를 이용하여 주요 설명

Copyright (C) 2003 NuriMedia Co., Ltd.

변수를 분석하였다. 로지스틱회귀모형의 경우 MLE(Maximum Likelihood Estimates) 분석에 의해 유의수준 5%에서 유의한 변수를, 의사결정나무의 경우 분리규칙을 형성시키는 변수를 주요설명 변수로 선별하였다. 그 결과 초등 비행의 경우 수표해독(x7), 공구용도(X11), 속도추적능력(X3), 기계원리(X12), 계기판독(X13), 감수성(X19) 등이 수료 여부를 설명하는 중요 변수로 선별되었다. 중등 비행의 경우 추적능력(X2), 주의 분배력(X6), 기계원리(X12), 계기판독(X13), 100M달리기(X14) 등이 중요 변수로 분석되었다. 고등 비행의 경우 요인분석을 실시하여 6개로 축약된 변수를 이용하여 분석을 실시하였다. 그 결과 X23(인간관)과 X25(자아성취도)에 의해 축약된 자기인식과 X6(주의 분배력2)와 X11(공구용도)에 의해 축약된 비행자질이 중요한 변수로 분석 되었다.

표 4. 단계별 분석 결과

구분	수료율 (%)	분류정확도(%)			ROC 성능 비교	주요설명변수	
		의사결정나무	로지스틱	신경망		의사결정나무	로지스틱회귀모형
초등비행	86.8	94	87	91.3	신경망, 의사결정나무	X7,X11	X3,X11, X12,X13, X19
중등비행	72.4	78	79	82	신경망, 로지스틱		X2,X6, X12,X13, X14
고등비행 (축소된 변수 이용)	86.7	91.7	91	92.8	신경망, 로지스틱		자기인식 (X23,X25) 비행자질 (X6,X11)

3.2 전 비행과정 결과 분석

비행훈련 성공 및 실패 예측모형의 궁극적인 목적은 어떠한 자원이 비행을 고등훈련까지 무사히 마치고 조종사가 될 수 있는가 하는 문제일 것이다. 따라서 초등에서 고등훈련까지 비행 전 과정을 종합적으로 분석하는 것은 중요한 의미를 갖는다. 전 비행 과정 결과에 대한 분석은 먼저 288명의 자료를 수료와 도태로 구분하여 분석함으로써 입과 후 최종 고등 훈련과정까지 수료 여부를 판단할 수 있는 예측모형을 제시하고 결과에 중요하게 영향을 미치는 변수를 선별하였다. 그리고 각 과정별 중간 과정의 수료 여부를 분석하기 위해 결과를 초등 도태 (범주1), 초등 수료 (범주2), 중등 수료 (범주3), 고등 수료 (범주4)의 4개 범주로 구분하여 분석을 실시하였다. 각 비행 결과에 대하여 세가지 모형을 적용하여 그 결과를 분류정확도와 ROC 및 Lift- Value 관점에서 비교평가 하였으며, 각 과정별 중요 변수를 분석하였다.

전 비행 과정을 수료와 도태로 구분하여 분석한 결과 분류정확도는 로지스틱회귀모형이 74%, 신경망이 71.3%, 의사결정나무는 66.1%로 나타났다. 로지스틱회귀모형의 경우 최종 수료율 54.5%를 고려시 비교적 높은 적중율을 보였다. 로지스틱회귀모형이 우수한 모형으로 나타나 로지스틱회귀모형을 적용시 비교적 정확한 결과를 예측할 수 있을 것으로 판단되었다.

Copyright (C) 2003 NuriMedia Co., Ltd.

그림 5. 전과정 2개 범주 ROC도표

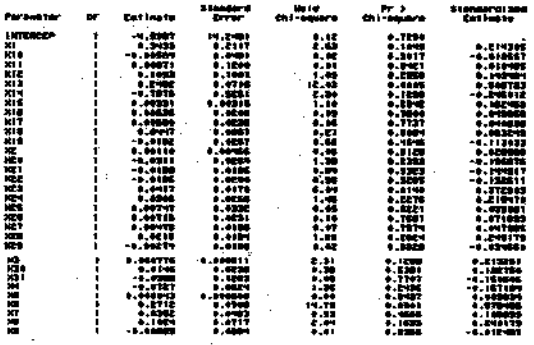
비행결과에 중요한 영향을 미치는 변수로는 의사결정나무의 경우 주의분배력(X6), 계기판독능력(X13), 인간관(X23)이 선별되었으며 로지스틱회귀모형에서는 주의분배력 (X6), 계기판독능력(X13), 기계원리(X12)가 선별되었다. 따라서 최종 고등비행 훈련까지 수료하는데 가장 중요한 변수는 비행 중 입력되는 많은 정보에 대해 주의력을 적절히 배분하여 이를 효과적으로 처리할 수 있는 주의분배능력과 항공기 자세를 계기로 판독하는 계기판독 능력으로 분석된다. 위의 분석결과 ROC 성능과 분류정확도가 가장 우수한 로지스틱 회귀모형을 이용하여 비행훈련 성공 및 실패 예측모형을 구축한다면 비행결과예측 및 조종훈련 후보생의 선발에 크게 기여할 수 있을 것으로 판단된다. 로지스틱회귀모형의 MLE (Maximum Likelihood Estimates) 분석결과는 표5와 같다.

표5 결과의 사후확률에 대한 추정식은 식 1이 되며

$$\hat{P}(y=1|X_1,..,X_n)=\frac{\exp(-4.936740.3433X_1+0.0011X_2,.....-0.0368X_n)}{1+\exp(-4.936740.3433X_1+0.0011X_2,.....-0.0368X_n)} \quad (1)$$

절단값(Cutoff-Value) 60%에서 분류정확도가 가장 높게 나타났다.

표 5. MLE 분석 결과



다음은 입과한 자원이 어느 과정까지 수료할 수 있는가를 판단하기 위해 비행 전 과정의 결과를 4개 범주로 구분하여 비행 결과를 분석하였다. Lift-Value 관점에서는 의사결정나무가 가장 우수한 것으로 평가되었으며, 분류정확도는 신경망이 66.1%, 의사결정나무는 64.7%, 로지스틱회귀모형이 63.3%로 나타났다. 어느 과정까지 수료할 수 있는가를 예측하기 위해 중요한 변수로는 로지스틱모형의 경우 속도추적능력(X3), 주의분배력(X6), 수표해독(X7), 기계원리(X12), 계기판독(X13)이, 의사결정나무의 경우 속도추적능력(X3), 주의분배력(X6), 기계원리(X12)가 선별되었다.

Copyright (C) 2003 NuriMedia Co., Ltd.

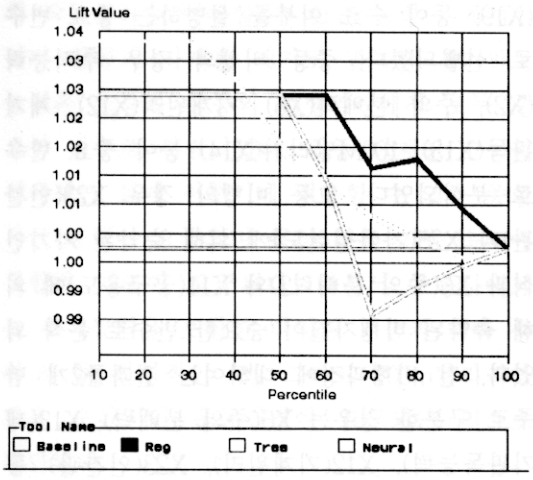


그림 6. 전과정 4개 범주 Lift 도표

그림 6의 Lift-Value 관점에서 가장 우수한 의사결정나무를 이용하여 입과한 인원이 어느 과정까지 수료할 수 있는가를 예측하여 보았다. 그림7에서 알 수 있듯이 전 과정에 걸친 수료 여부를 판단할 수 있는 주요 변수는 X12 (기계원리), X6 (주의분배력2) , X3 (속도추적능력)이며 X12가 8.5보다 같거나 클 경우 최종 고등비행 수료까지 가능한 것으로 예측하고 이 때 정확도는 86.6%이었다. 그리고 X12가 4.5보다 작고 X6가 6보다 작을 경우 초등에서 도태되는 것으로 예측하며 이 때 정확도는 100%이었다. 한편 X12가4.5보다 크고 8.5보다는 작으며 X3이 1048보다 크거나 같을 경우에는 고등비행까지 수료로 예측하며 정확도는 100%이었다. 그리고 X12가4.5보다 크고 8.5보다는 작으며 X3이 397과1048 사이에 있으며 X6가 10보다 작을 때는 초등 수료로, 10보다 크고 12.5보다 작을 경우는 중등수료로 분류하며 이 때 정확도는 100%이었다.

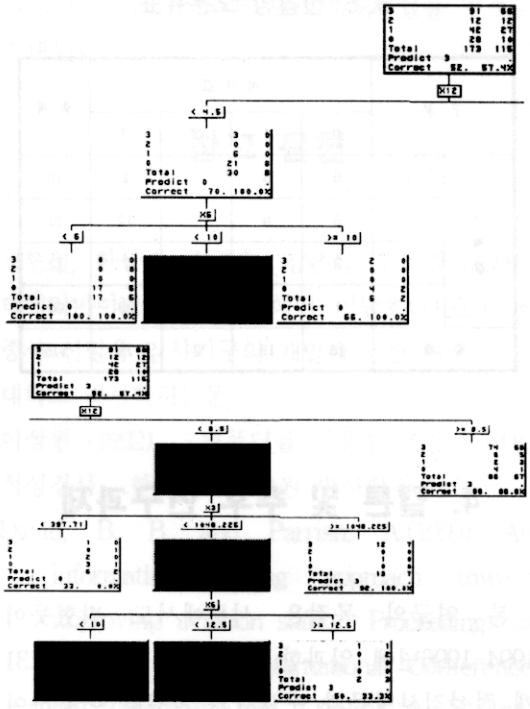


그림 7. 전 과정 의사결정나무

다음은 분류정확도가 가장 높은 신경망의 오분류표에 의해 분류결과를 알아보았다. 표 6에서 보는 바와 같이 분류 범주가 4개인 것을 고려시 범주3과 범주0에 대하여는 높은 정확도를 보이나 범주1과 범주2는 낮은 정확도를 보임을 알 수 있다. 특히 범주3의 분류정확도가 높은 것은 전체 데이터에서 범주3의 상대적인 비율(54.5%)이 높기 때문인 것으로 분석된다.

Copyright (C) 2003 NuriMedia Co., Ltd.

표 6. 신경망 오분류표

구 분		예측값				총 계
		0	1	2	3	
실제값	0	6	3		1	10
	1	6	9		12	27
	2	3		2	7	12
	3	1	6		59	66
총 계		16	18	2	79	115

4. 결론 및 추후 연구과제

본 연구의 목적은 서론에서도 밝혔듯이 1994-1996년에 입과한 조종학생 288 명의 31개 적성검사항목과 수료결과 자료를 이용하여 비행 결과를 예측할 수 있는 최적의 모형을 제시하고 아울러 각 단계별로 수료 결과에 중요하게 영향을 미치는 변수가 무엇인가 분석하는 것이다. 그 결과 비행단계별 순차적인 결과에 대하여는 전반적으로 신경망이 분류정확도 및 ROC 판점에서 우수한 모형으로 평가되었다. 비행 전 과정의 결과에 대하여는 결과를 2개 범주로 구분하여 분석시는 로지스틱회귀모형이 분류정확도와 ROC관점에서 우수한 모형으로 평가되었다. 결과를 4개 범주로 구분하여 분석시는 분류정확도 관점에서는 신경망이 Lift-Value관점에서는 의사결정나무가 우수한 모형으로 평가되었다. 결과에 가장 큰 영향을 미치는 변수로는 초등 비행의 경우 수표해독(x7), 공구용도(X11), 속도추적능력(X3), 기계원리(X12), 계기판독(X13), 감수성(X19) 등이 수료 여부를 설명하는 중요 변수로 선별되었다. 중등 비행의 경우 추적능력(X2), 주의 분배력(X6), 기계원리(X12), 계기판독(X13), 100M달리기(X14) 등이 중요 변수로 분석되었다. 고등 비행의 경우 X23(인간관)과 X25(자아성취도)에 의해 축약된 자기인식과 X6(주의 분배력2)와 X11(공구용도)에 의해 축약된 비행자질이 중요한 변수로 분석 되었다. 전 비행과정에 대하여는 결과를2개 범주로 구분한 경우는 X6(주의 분배력), X13(계기판독능력), X12(기계원리), X23(인간관) 등이 중요 변수로 선별되었으며, 결과를4개 범주로 구분한 경우는 X3(속도추적능력), X6(주의 분배력), X13(계기판독능력), X12(기계원리), X7(수표해독) 등이 중요 변수로 선별되었다. 각 비행단계별 수료 결과 예측에 있어서 중요하게 작용하는 변수가 차이가 있는 것은 각 비행단계에서 요구되는 기능의 차이 때문인 것으로 분석 된다. 즉 초등 비행의 경우 단독 비행을 위한 안전한 이착륙 조작이 가장 중요한 요소이므로 비행시 접근율, 침하율 등의 변동 속도를 정확히 추정하여 공간 거리를 예측할 수 있는 능력의 판단 기준인 속도추적능력과 항공기 자세를 계기로 판독할 능력이 있는가를 판단하는 계기판독 능력 등이 중요 설명 변수로 선별되었다.

중등의 경우 전술구사를 위한 편대비행이나 생존능력을 위한 계기비행이 요구 된다. 따라서 공중 기동시 제시되는 목표물과 비행 정보에 대한 정밀한 추적 및 비행조작의 능력을 판단할 수 있는 추적능력과 순발력의 판단 기준인 100m 달기기가 중요한 변수로 선별되었다. 그리고 항공기 조작에 필요한 각종 기계

Copyright (C) 2003 NuriMedia Co., Ltd.

의 작동원리를 이해할 수 있는 능력인 기계원리와 항공기 자세를 계기로 판독할 계기판독 능력은 초등과 중등의 비행과정에서 모두 중요한 변수로 선별되었다. 최종 단계인 고등비행에서는 실제 전투기에 적용할 수 있는 종합적인 능력이 요구되며 특히 빠른 기동 속도에 적합한 기능이 요구된다. 따라서 조종사에게는 추적능력을 포함한 다양한 기능과 자아 성취도 등 인성적인 면, 그리고 100m달리기와 같은 순간적인 운동신경이 요구되는 것으로 판단된다.

분류정확도 관점에서는 단계별 순차적인 분석의 경우 목표변수의 범주가 한 곳(수료)에 편중되어 있는 점을 고려시 그리 높지 않은 것으로 나타났다. 전 비행 과정의 결과에 대해 2개 범주로 구분하여 분석한 결과 로지스틱회귀모형의 경우 분류정확도 74%로 나타났다. 이는 최종 수료율 54.4%를 고려시 비교적 높은 정확도로 판단된다. 따라서 이 모형을 조종후보생 선발 및 비행 결과 예측을 위한 분류모형으로 이용시 비교적 정확한 결과를 예측할 수 있을 것으로 판단된다.

그러나 좀 더 높은 분류정확도를 보장하기 위해서는 이 연구 결과 나타난 중요변수 및 추가적으로 실제 비행을 모사한 모의 비행장비를 이용한 비행적성판단 프로그램을 개발하는 등 복합적인 노력이 필요하다. 또한 실험단계에서 체계적인 실험계획과 완벽한 실험을 통하여 유용한 자료를 구하고 이를 효과적으로 분석한다면 좀 더 적중률이 높은 모형을 구축할 수 있을 것이다. 이렇게 구축된 모형을 실제 비행 입과 인원의 선발 및 부족한 적성을 판단하는 도구로 이용한다면 최초에 목적했던 소기의 성과를 달성할 수 있을 것으로 기대된다.

참고 문헌

강현철, 한상태, 최종후, 김은석, 김미경 (1999), 데이터마이닝 자유아카데미 이달호 (1992), 조종사 선발용 검사기구의 개발에 관한 연구, 서울대학교 박사학위논문

이상원 (1992), 컴퓨터를 이용한 조종사 선발 적성검사, 한국과학기술원 박사학위 논문

David, B. B. and Parrish, A.(1999) An information mining approach toward improving aviation safety Proceedings of the 5th ISSAT International Conference on Reliability and Quality in Design, 249-253

Gupta, A., Park, S., and Lam Sm. (1999) Generalized analytic rule extraction for feedforward neural networks IEEE Transactions on Knowledge & Data Engineering, V.11 N.6, 985-991

Jagielska, H.,(1998), Linguistic Rule Extraction from Neural Networks for Descriptive Datamining 1998 Second International Conference on Knowledge-Based Intelligent Electronic Systems, April 1998

Shub, Y., Kushnir, A., and Frenkel, J.(1994), Pilot evaluation system, Aerospace and Electronics Conference, 1994. Proceedings of the IEEE 1994 National, Page(s): 734

Copyright (C) 2003 NuriMedia Co., Ltd.

-741 vol.2

Vorko, A. and Jovic, F. (1999) Multiple attribute entropy classification of school-age injuries Accident Analysis and Prevention 32(2000) 445-454, May 1999

저자 소개

◈ **손소영**

연세대학교 컴퓨터과학·산업시스템공학 교수

◈ **조용관**

연세대학교 컴퓨터과학·산업시스템공학
석사 3학기

◈ **최성욱**

공군사관학교 조종적성연구실장

◈ **김영준**

공군사관학교 산업공학과 교수

논문접수일 (Date Received): 2001/5/3
논문게재승인일(Date Accepted): 2001/11/7

Copyright (C) 2003 NuriMedia Co., Ltd.