과 목 명 : 빅데이터마이닝

1. 머신러닝 중 하나인 의사결정나무 분석 결과 다음과 같이 나타났다. (15점)

	Reference	
Predicted	**0**	**1**
0	300	55
1	40	400

1.1 Accuracy ? ( ) %

1.2 Sensitivity ? ( ) %

1.3 Specificity ? ( ) %

2. 'dataTransactions.tab' 데이터를 활용하여 연관규칙탐사를 수행하고 표 안을 채워라 (15점)

조건 : ① 일반식품과 가구 코너 제외, ② 지지도 20% 이상, 신뢰도 80% 이상만 고려
③ 스포츠 코너 중심 분석 (향상도 1.5 초과만 고려)

	lhs	rhs	support	confidence
[2]		**스포츠**		

3. **(머신러닝_판별분석)** Turkey.csv 데이터에서 'HUM'과 'RAD'를 활용하여 칠면조의 'TYPE'를 예측하는 모델을 만들고 다음에 답하여라 (20점)

3.1 LDA(Linear Discrimination Analysis) 적용 할 경우 (HUM, RAD)가 각각 (150, 135), (150, 150)인 TURKEY가 'DOMESTIC'일 확률은?
(150, 135) TURKEY → (%)
(150, 150) TURKEY → (%)

3.2 QDA(Quadratic Discrimination Analysis) 적용 할 경우 (HUM, RAD)가 각각 (150, 135), (150, 150)인 TURKEY가 'DOMESTIC'일 확률은?
(150, 135) TURKEY → (%)
(150, 150) TURKEY → (%)

4. **(머신러닝_신경망분석)** Hshopping2.txt 데이터, 신경망 분석, Training data 60%, Test data 40%, 고객나이, 고객구매금액, 출연자 고려, 신경망 7개 적용, maxit=1000, 고객 반품예측 분석 (20점)

4.1 고객 반품 예측에 가장 중요하게 생각하는 변수는 ? ( )

4.2 고객 반품 예측 모델의 신뢰구간 95% 예측 정확도는 ? (~ %)

5. **(머신러닝_의사결정나무)** Hshopping2.txt 데이터, 의사결정 나무 분석, Training data 60%, Test data 40%, 고객나이, 고객구매금액, 출연자 고려, winnow=FALSE, noGlobalPruning=FALSE, 고객 반품예측 분석 (20점)

5.1 고객 반품 예측에 가장 중요하게 생각하는 변수는 ? ( )

5.2 고객 반품 예측 모델의 신뢰구간 95% 예측 정확도는 ? (%)