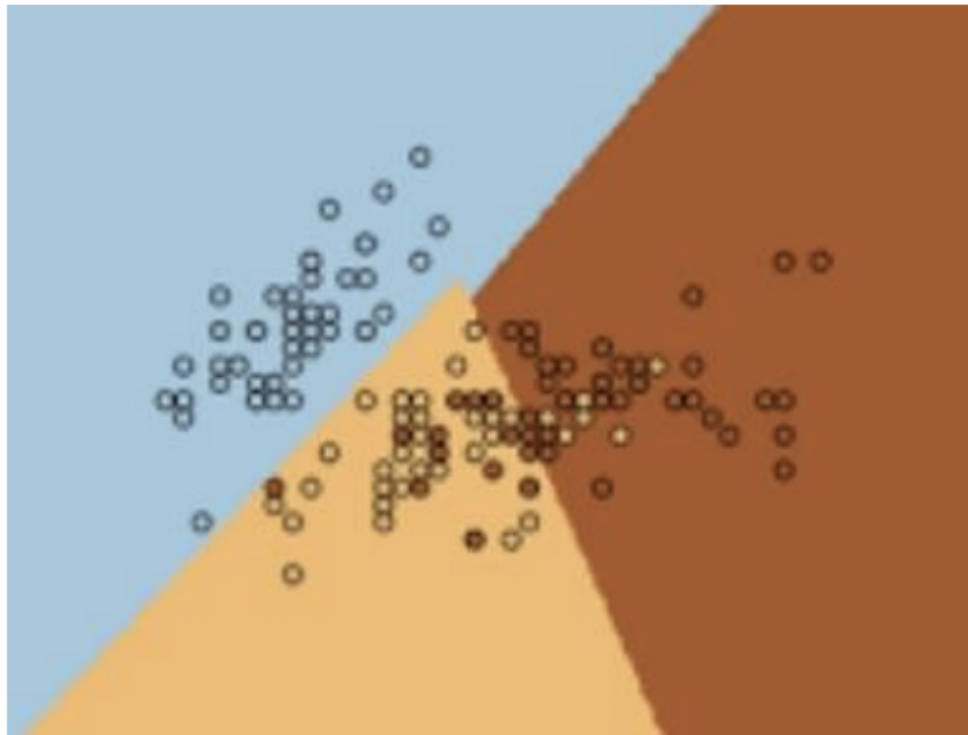




3 / 16

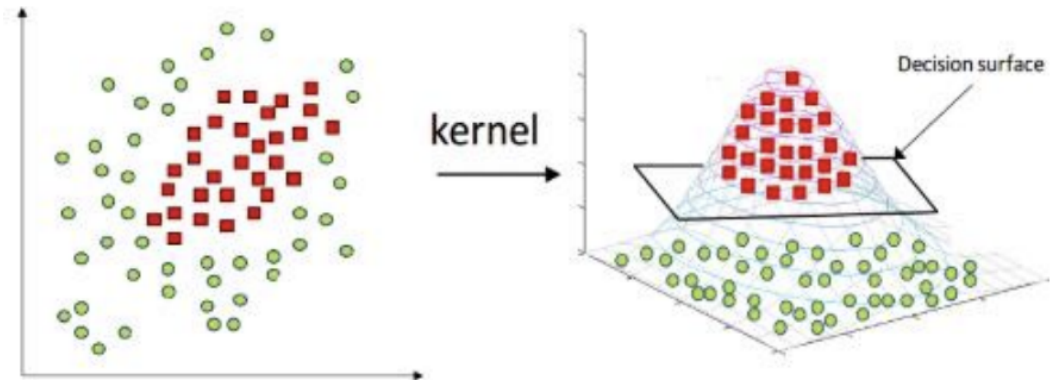
- 모델과 데이터 집합
 - 예측 모델을 만드는 일은 다른 속성을 이용해 타겟 변수에 대한 모델을 알아내는 과정
 - 전체 데이터 집합을 세분화해 더욱 정확한 정보를 전달하는 속성을 찾아가면서 감독 세분화 모델을 생성하는 과정
 - 기하학적 관점에서 보면 객체 공간을 더 정확한 정보 영역으로 분할하는 과정
 - 감독 학습, 비지도 학습으로 나눈다. 어떤게 좋은가? 비교사가 더 좋다. 보완
 - chatgpt도 교사학습, 정보를 학습한다.
 - 인공지능은 데이터에 접근할 때, 데이터 마이닝으로 접근한다.
 - 데이터 마이닝은 실시간 처리가 안됨. 새로운 정보를 추출하는게 어렵다
 - 20년 정도 역사를 가짐
 - 실제로 코딩을 해봐야 한다.
 - 유전 알고리즘
 - 모호하고, 부정확하며 불확실한 지식과 데이터를 다루는 중요한 기술 중 하나는 퍼지 논리임.
- 모델 구조
 - 일부 수치형 파라미터를 지정하지 않고도 모델구조만 만들어 놓은 상태에서 데이터 세트에 대한 예측 모델을 학습
 - 데이터 마이닝 기법을 적용해 주어진 훈련 데이터 세트에 최적화된 파라미터를 계산
 - 일반적으로 모델 구조는 여러 수치형 속성에 대해 파라미터화된 수학적함수나 방정식으로 정의

- 모델에 사용될 속성은 어느 속성이 타겟 변수를 예측하는 데 정보를 주는지 해당 영역의 지식을 통해서 선택
- 속성 선택을 데이터 마이닝 기법으로 찾아낼 수 있음
- 데이터 마이닝 작업을 수행하는 데이터 과학자는 모델의 형태와 속성을 지정
 - 데이터 마이닝의 목표는 모델이 가능한 한 데이터에 잘 맞도록 파라미터를 튜닝
- 단순화
 - 분류와 계층 확률을 추정하는 데 단순 계층만 고려
 - 여러 값을 갖는 다양한 계층에도 쓸 수 있도록 일반화
 - 복잡도
 - 구분자
 - 명확성



- 객체 공간 분류
 - 수학적 함수를 활용하여 객체 공간을 구분
 - 각 영역에 있는 객체들은 유사한 타겟 변수값 소유

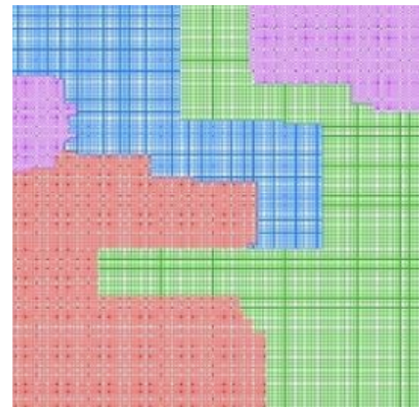
- 동질성을 가진 영역으로 분할
 - 동질 속성 추출과 새로운 객체의 타겟값 예측



○ 분할

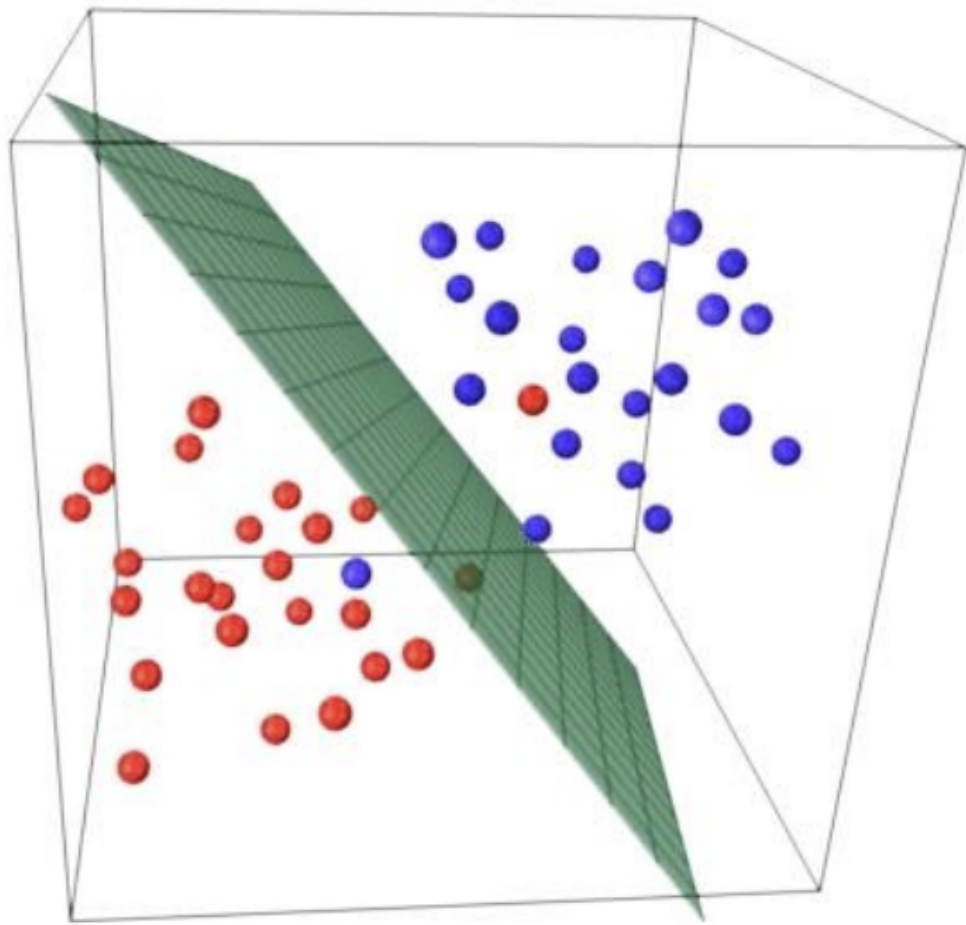
- 공집합이 아닌 집합 A 의 분할은 A 의 부분 집합의 모임 $\{A_1, A_2, A_3, \dots, A_N\}$ 을 말하며, 다음 조건을 만족 시킨다.

- (1) $A_i \neq \phi, 1 \leq i \leq n$
- (2) $A = \bigcup_{i=1}^n A_i$
- (3) $A_i \cap A_j = \phi, i \neq j$



○ 선형 분할

- 선형 함수를 활용하여 분할에 적용
- 간단한 처리로 작업 완성
- 해석의 편리성
- 자동화



◦ 프레임

- 객체 또는 개념에 관한 지식
- 데이터 구조
- 이름과 연관된 속성(또는 슬롯)의 집합을 포함
- 슬롯은 특정 값
 - 속성 값을 결정할 수 있는 프로시저

A 항공사 탑승권		B 항공사 탑승권	
항공사	A	항공사	B
이름	Mr ABC	이름	Mrs ABCDE
항공기	AP123	항공기	BP1234
날짜	20Dec 2020	날짜	21Nov 2020
좌석	A24	좌석	B23

○ 프레임의 필요성

- 지식을 구조화하고, 정확하게 표현하는 합리적인 방법을 제공
 - 단일 개체(entity)에서 프레임은 특정 객체 또는 개념에 대한 모든 필수 지식을 결합
 - 프레임은 지식을 표현하는 방법으로 슬롯을 사용하며, 슬롯을 통해 객체의 다양한 특성과 특징을 기술
- 조건절과 귀결절을 이용하여 규칙을 합리적으로 표현
- 프레임은 객체 지향성 유지

○ 지식 표현 기법으로의 프레임

- 일반적인 슬롯 정보
 1. 프레임 이름
 2. 프레임과 다른 프레임간의 관계
 - IBM Aptiva S35 프레임은 컴퓨터 클래스의 소속이 될 수 있고, 컴퓨터는 하드웨어 클래스에 속할 수 있다.
 3. 슬롯 값
 - 슬롯 값은 심벌, 숫자, 불린이 될 수 있다.
 - 예를 들어, 이름 슬롯에는 심벌 값이, 탑승구 슬롯에는 숫자 값이 들어간다.



- 패킷

- 프레임의 속성에 대한 확장된 지식을 제공하는 방법
- 패킷은 속성 값을 확정하고, 이용자의 질문을 제어하며 추론 엔진에 속성을 처리하는 방법을 알려준다.

- 일반적으로 프레임기반 전문가 시스템은 속성에 값 패킷, 프롬프트 패킷, 추론 패킷을 할당할 수 있게 한다.
- 값 패킷은 속성에 기본 값과 초기 값을 규정한다.
- 프롬프트 패킷은 시스템이 동작할 때, 사용자가 온라인으로 속성 값을 입력할 수 있다.
- 추론 패킷은 고정된 속성 값이 바뀔 때 추론 절차를 멈추게 한다.
- 문제 분할
 - 패킷은 컴퓨터 클래스의 플로피, 전원 공급기, 보증기간과 재고 슬롯에 기본 값과 초기 값을 할당한다.
 - 클래스의 이름, 종류, 기본 값과 초기 값 속성은 인스턴스에 의해 상속되는 특성이다.
- 클래스와 인스턴스
 - 프레임은 특정 객체를 지칭하는 인스턴스 프레임과 유사한 객체 그룹을 의미하는 클래스 프레임으로 나뉜다.
 - 클래스 프레임은 공통 속성을 포함한 객체의 그룹을 의미한다.
- 객체와 클래스
 - 객체를 클래스로 묶으면, 추상적인 형태로 문제를 표현할 수 있다.
 - 프레임 기반 시스템은 클래스 상속을 지원
 - 상속이라는 절차를 통해 클래스 프레임의 모든 특징들이 인스턴스 프레임으로 전달
 - 기본적으로 클래스 프레임의 속성은 클래스 내의 모든 객체에 보편적으로 적용
 - 인스턴스의 슬롯에는 각 인스턴스에 해당되는 실제 데이터 값이 저장



클래스: 승용차
[C] 엔진 종류 4기통: 6기통:
[N] 마력:
[C] 동력전달 종류 후륜 구동: 전륜 구동: 4륜 구동:
[C] 변속기 종류 5단 수동: 4단 자동:
[N] 연료 소모(마일/갤런):
[N] 좌석 수:

클래스: 모델
상위 클래스: 승용차
[C] 엔진 종류 4기통: 6기통:
[N] 마력:
[C] 동력전달 종류 후륜 구동: 전륜 구동: 4륜 구동:
[C] 변속기 종류 5단 수동: 4단 자동:
[N] 연료 소모(마일/갤런):
[N] 좌석 수:
[Str] 생산 국가: 일본

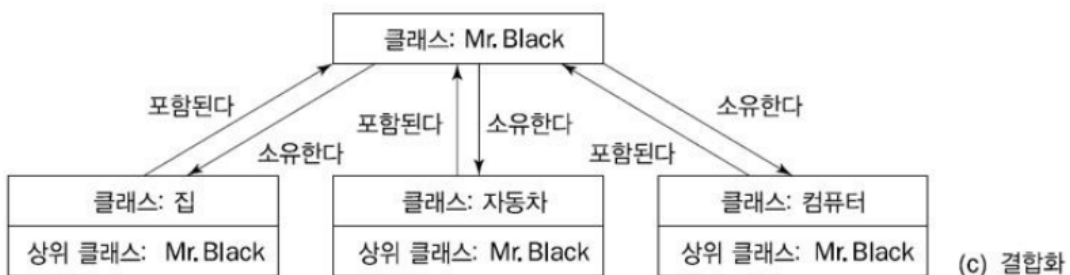
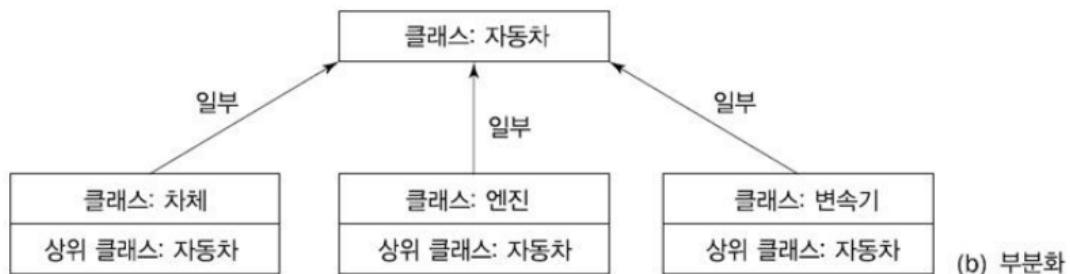
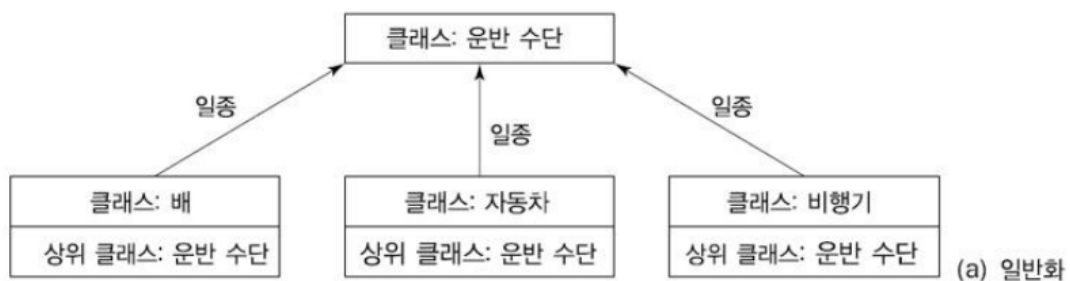
클래스: Mazda 626
상위 클래스: Mazda
[C] 엔진 종류 4기통: 6기통:
[N] 마력: 125
[C] 동력전달 종류 후륜 구동: 전륜 구동: 4륜 구동:
[C] 변속기 종류 5단 수동: 4단 자동:
[N] 연료 소모(마일/갤런): 22
[N] 좌석 수: 5
[Str] 생산 국가: 일본
[Str] 모델명:
[C] 차량 색상 흰색(빙하): 메탈릭 회색: 메탈릭 회청색: 오닉스 검정색:
[Str] 소유주:

클래스: Mazda DR-1216
상위 클래스: Mazda
[C] 엔진 종류 4기통: TRUE 6기통: FALSE
[N] 마력: 125
[C] 동력전달 종류 후륜 구동: FALSE 전륜 구동: TRUE 4륜 구동: FALSE
[C] 변속기 종류 5단 수동: FALSE 4단 자동: TRUE
[N] 연료 소모(마일/갤런): 28
[N] 좌석 수: 5
[Str] 생산 국가: 일본
[Str] 모델명: DX
[C] 차량 색상 흰색(빙하): FALSE 메탈릭 회색: TRUE 메탈릭 회청색: FALSE 오닉스 검정색: FALSE
[Str] 소유주: Mr. Black

(b) 자동차 프레임과 각 프레임의 슬롯

- 프레임기반 시스템에서 객체의 관계
 - 일반적으로 객체 간의 관계는 일반화, 부분화, 결합화로 구분

- 일반화는 상위 클래스와 하위 클래스 간의 일종(is-a 또는 a-kind-of) 관계를 맺는다.
 - 차량은 일종의 운반 수단이다.
 - 차량은 운반 수단이라는 좀 더 일반적인 상위 클래스의 하위 클래스이다.
 - 각 하위 클래스는 상위 클래스의 모든 특성을 상속받는다.
- 부분화는 전체를 나타내는 상위 클래스와 컴포넌트를 의미하는 하위 클래스 간에는 일부(a-part-of 또는 part-whole) 관계를 맺는다.
- 결합화는 관련이 없는 서로 다른 클래스 사이의 일종의 연결 고리 기능을 제공한다.



• 상속

- 상속은 인스턴스 프레임에 클래스 프레임의 모든 특징을 부여하는 과정
 - 일반적인 상속이란 모든 인스턴스 프레임에 기본 값을 대응시키는 과정

- 프레임은 그와 관련된 상위 레벨 프레임의 특성을 상속
 - 객체 또는 개념의 고유한 특성을 담고 있는 클래스 프레임을 한 번 생성하면, 클래스 레벨의 특성을 일일이 지정하고 앎고도 인스턴스 프레임을 다수 확보
- 추상화의 가장 상위 레벨은 트리의 루트로 가장 높은 곳에 표시
 - 루트 아래 가지는 좀 더 낮은 추상화 레벨을 의미
 - 맨 마지막에는 인스턴스 프레임을 표시



- 메소드와 데몬

- 메소드란 속성 값이 변할 때 속성 값을 결정하거나, 필요한 연속 동작을 실행시켜주는 특별한 속성
- 데몬은 IF-THEN구조로 IF 문의 속성 값이 변할 때마다 실행
- 데몬과 메소드는 매우 유사
 - 메소드는 복잡한 프로시저를 구현할 때, 더 적합
 - 데몬은 보통 IF-THEN 문에 한정

Credit Evaluation Advisor

대출 신청 금액	50000
현금 자산	50000
주식	9000
저당 가치	12000
순 자산 가치	40
작년 판매 성장률	20
이익 증가율	45
단기 판매 부채	9
예상 수익률	Excellent

신청자 목록

No	신청자 이름
CN001-98	Mrs White, J.
CN002-98	Mr Black, N.
CN003-98	Mr Green, P.
CN004-98	Mr Clark, E.

이전 다음
처음 마지막

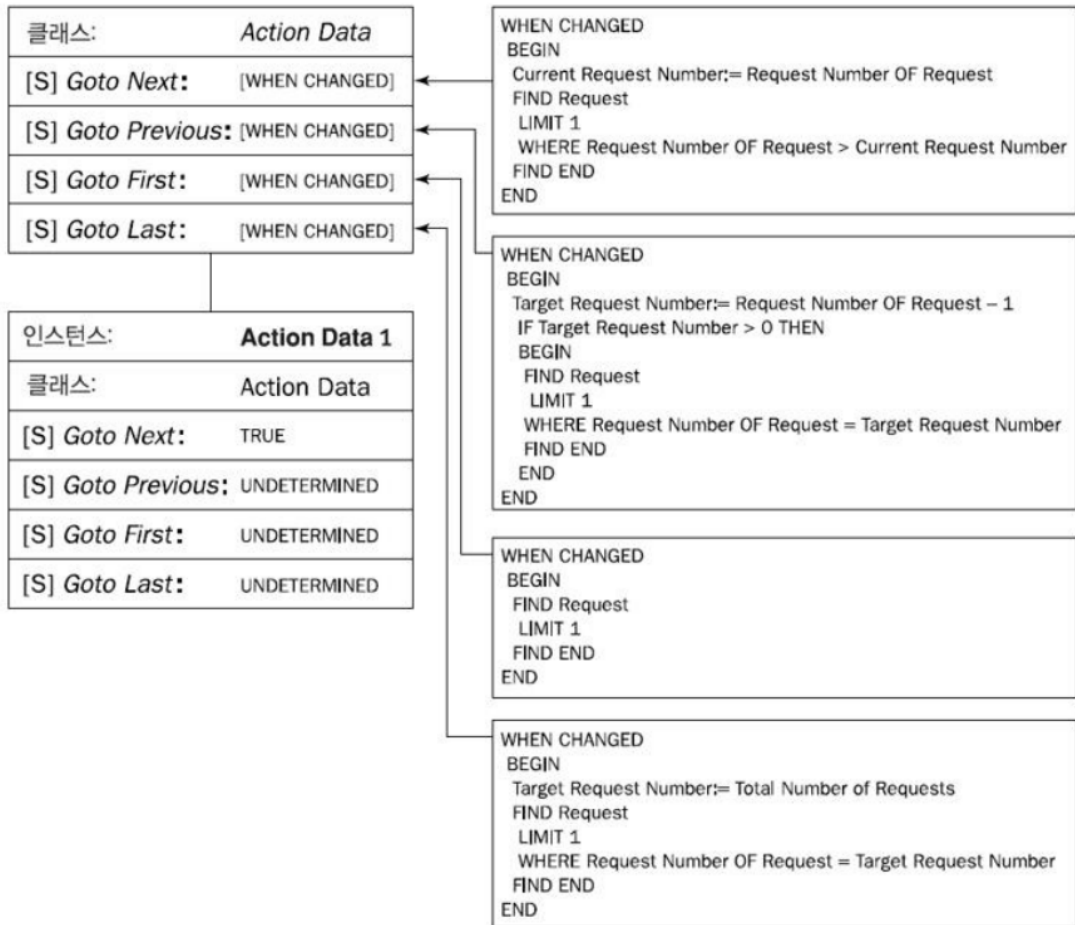
대출 조회 평가 도우미는 대출 요청에 대한 등급을 다음과 같이 결정하였습니다.

대출 가능

대출 평가

담보 등급	Excellent
재정 등급	Good

도움 재시작 종료



- 과제?

테스트

- 저장된 6개의 좌표 데이터들을 이용하여 2개의 그룹(a group, b group)으로 나눈다.

	다양한 방법으로 다음 데이터를 분류하시오.
저장	i1(1,1) i2(2,2) i3(3,3) i4(70,70) i5(80,80) i6(90,90)
입력	집단 수량 2개
출력1	a group, b group의 중심 좌표 변화를 연속 출력
출력2	a group, b group에 소속된 데이터 출력