영수증 데이터와 SVM을 이용한 고객의 구매의도 예측

빅데이터 청년인재 단국대학교 8조

INDEX

01

"CRM 기업승패를 좌우한다."

·CRM의 도입배경

고객 니즈의 다양성

인터넷과 소셜미디어의 발달

동종업계 경쟁의 가속화

▶ 사회적, 경제적, 기술적 변화에 따라 고객과 소통의 중요성은 커지고 있다.

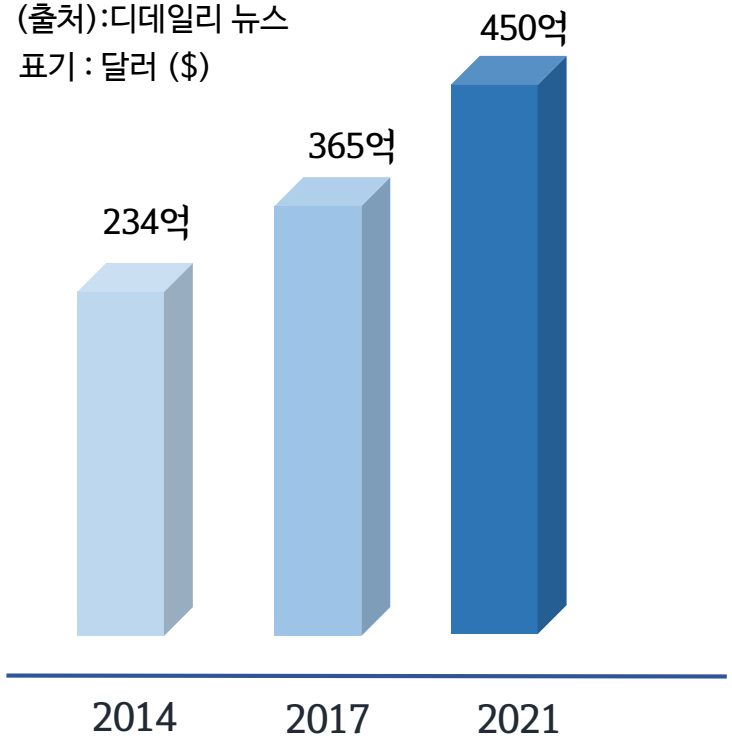
01

"CRM 기업승패를 좌우한다."

· CRM사업 현황

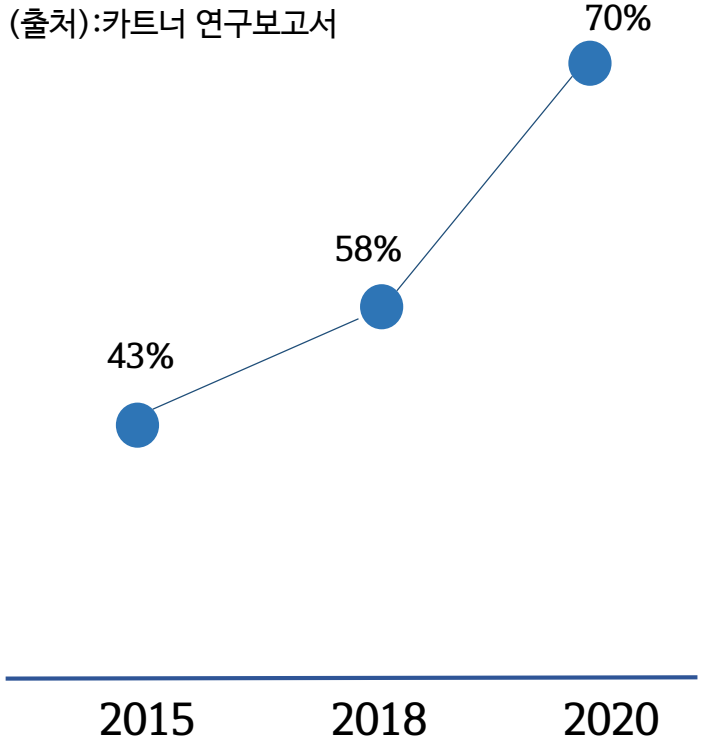
CRM시장규모

(출처):디데일리 뉴스
표기 : 달러 ($)



2017년 글로벌 소프트웨어 매출

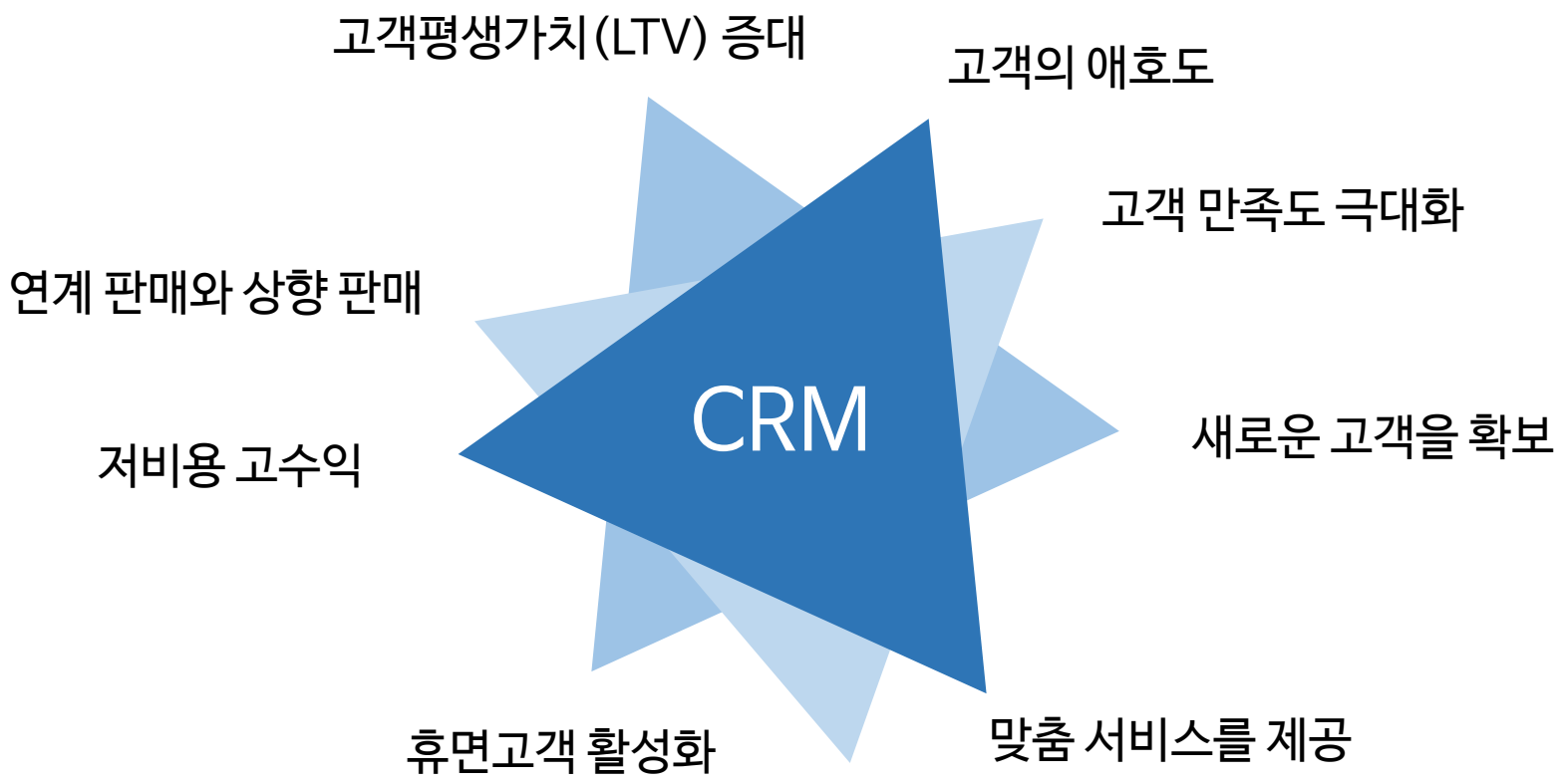
(출처):카트너 연구보고서



01

"CRM 기업승패를 좌우한다."

·CRM 효과



01

"CRM 기업승패를 좌우한다."

·CRM 성공사례

아마존의 성공 비결은

"휴머니즘이 결합한 마케팅 활동의 승리"

-아마존 Jeff Bezos 사장-

amazon.com

CRM의 기법 중 하나

YOUTUBE, 추천 검색 엔진

-Susan Wojcicki,아마존 사장-

YouTube

고객과 소통하는 길
CRM

고객과 소통하며, 통하다

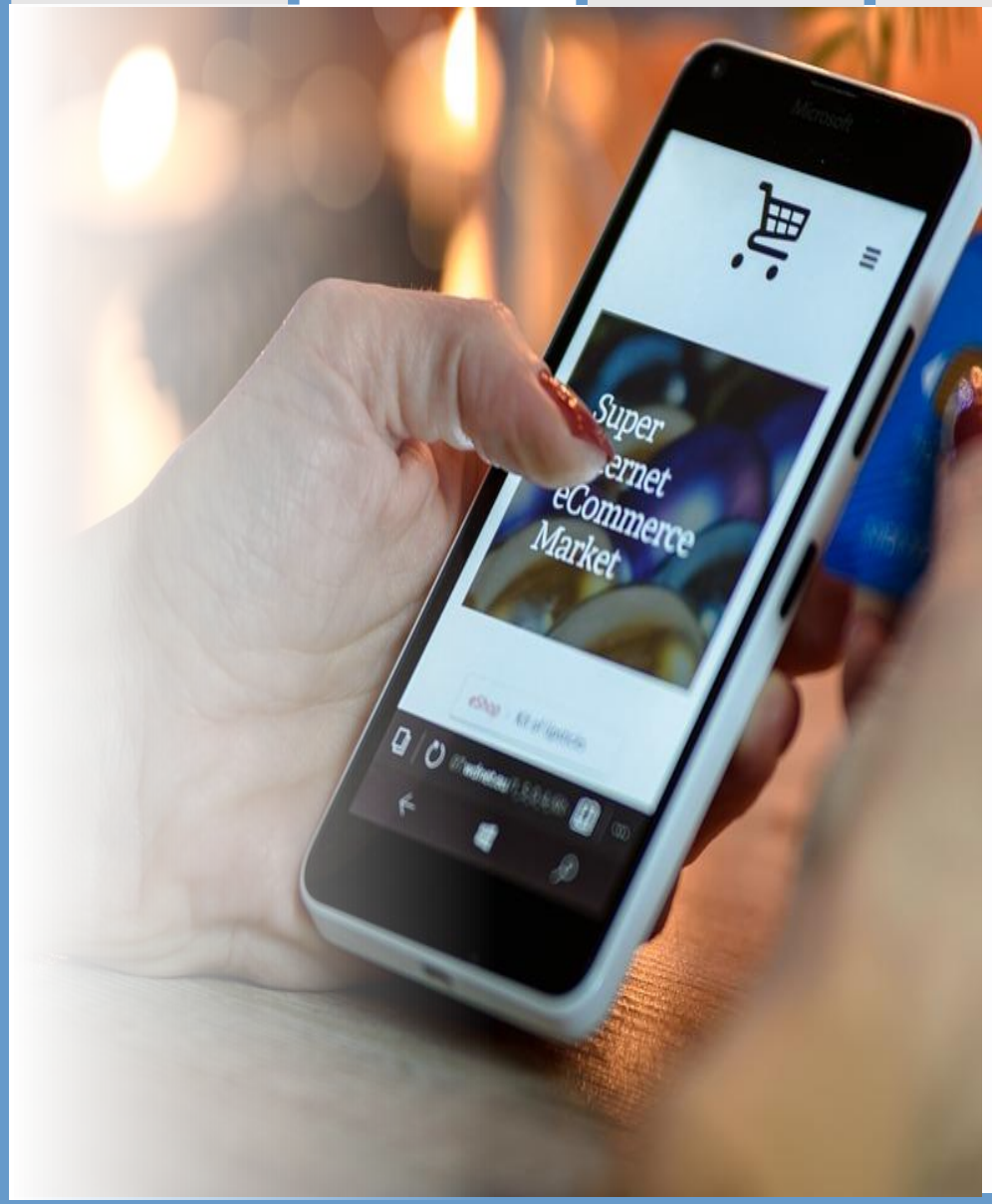
02

"고객 만족은 데이터에서 시작된다."

· L사의 고객 구매데이터 **320만건**

02

"고객 만족은 데이터에서 시작된다."



고객 구매 정보 list

ID
영수증 번호
업종
상품 소분류 코드
점포코드
구매일자
구매시간
구매금액
구매수량

02

"고객 만족은 데이터에서 시작된다."

쇼핑 업종 상품 분류

업종
상품 소분류 코드
소분류명
중분류명
대분류명

02

"고객 만족은 데이터에서 시작된다."

고객구매데이터 320만건 X 상품 소 분류 코드 3000건

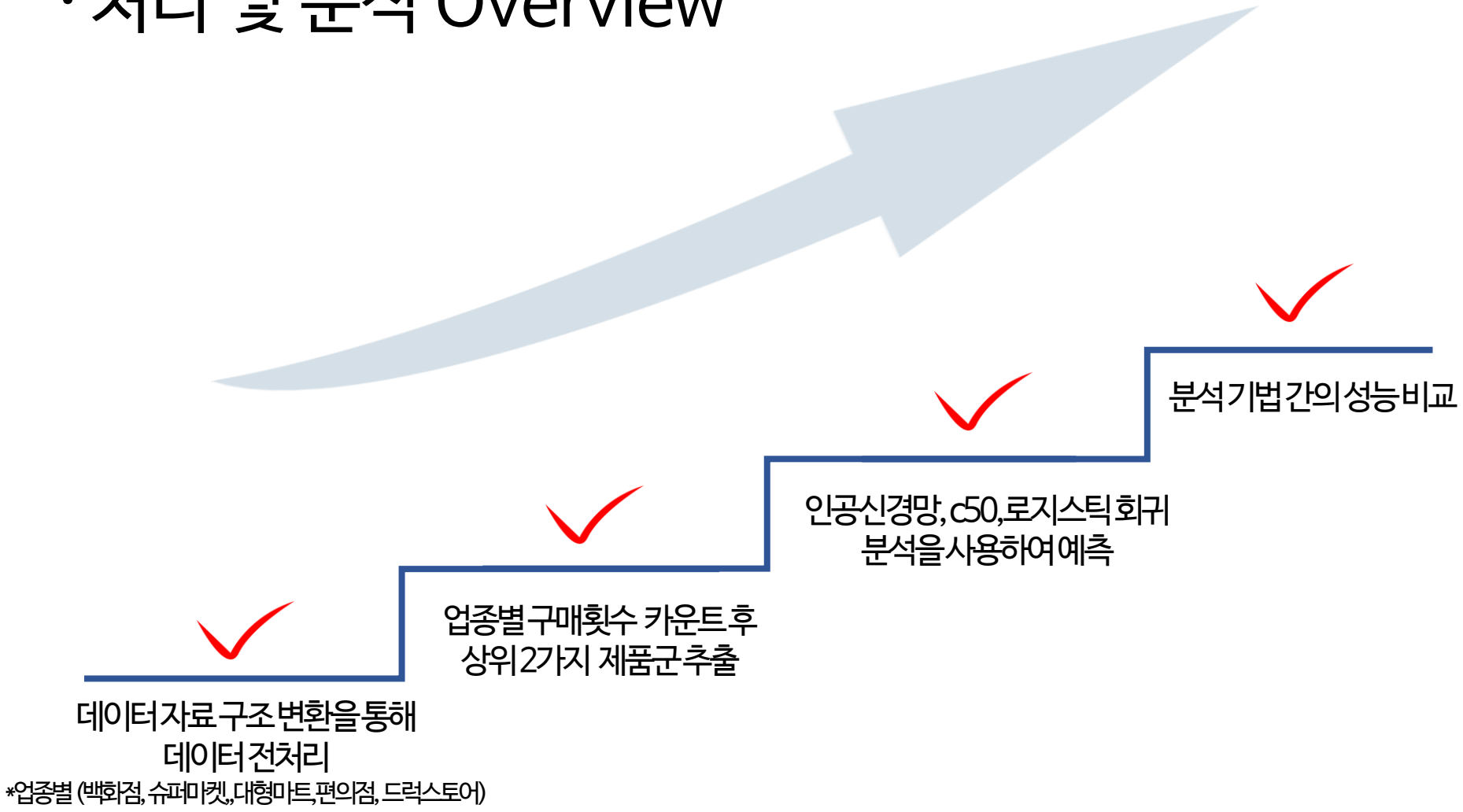
```
ID,RCT_NO,BIZ_UNIT,PD_S_C,BR_CDE_DT,DE_HR,BUY_AM,BUY_CT
04008,002108,A01,0215,0002,20150216,13,59600,2
06379,002109,A01,0075,0029,20150213,11,35000,1
06379,002109,A01,0149,0004,20150115,10,85000,1
08002,002110,A01,0138,0010,20151220,10,25000,1
08002,002110,A01,0138,0010,20151220,10,21000,1
08002,002110,A01,0558,0004,20150815,10,79200,1
07252,002111,A01,0013,0029,20150716,10,5400,1
05072,002112,A01,0223,0002,20150626,12,158000,1
05072,002112,A01,0216,0002,20150204,12,39000,1
05072,002112,A01,0121,0002,20150911,11,3000,1
05072,002112,A01,0121,0002,20150911,11,30000,2
10040,004082,A01,0421,0037,20151012,14,87000,1
10040,004082,A01,0421,0037,20151012,14,87000,1
13895,004083,A01,0532,0001,20151223,12,450000,1
09314,004084,A01,0113,0002,20150512,11,38700,1
15356,004085,A01,0143,0029,20150829,14,210000,1
15356,004085,A01,0439,0029,20151115,13,68000,1
15356,004085,A01,0531,0025,20150411,14,2198000,1
11502,004086,A01,0208,0022,20150102,14,142020,1
18739,004087,A01,0558,0029,20150621,11,80000,1
18739,004087,A01,0165,0047,20150501,10,227000,1
08283,004088,A01,0107,0001,20150911,11,12800,1
17230,004089,A01,0421,0040,20150921,16,108800,1
10214,004090,A01,0153,0027,20151025,11,17000,1
10214,004090,A01,0488,0016,20150203,10,53100,1
10214,004090,A01,0145,0016,20150203,10,94000,1
10214,004090,A01,0145,0016,20151124,12,24000,1
14040,004091,A01,0514,0007,20150427,13,45000,1
09321,004092,A01,0501,0041,20151113,11,27600,1
09321,004092,A01,0191,0041,20151113,12,175000,1
```

BIZ_UNIT		PD_S_C	PD_S_NM	PD_M_NM	PD_H_NM
A04	341	단행본서적(직배)	특수점서적	서적/음반	
A04	1	삼각김밥	삼각김밥	미반	
A04	2	The커진삼각김밥	삼각김밥	미반	
A04	3	말이김밥	김밥	미반	
A04	4	트레이김밥	김밥	미반	
A04	5	초밥	김밥	미반	
A04	6	도시락	도시락	미반	
A04	7	미니도시락	도시락	미반	
A04	8	기타	도시락	미반	
A04	9	국/찌개도시락	도시락	미반	
A04	10	떠먹는주먹밥	용기형주먹밥	미반	
A04	11	동그란주먹밥	동그란주먹밥/사각주먹밥		미반
A04	12	사각주먹밥	동그란주먹밥/사각주먹밥		미반
A04	13	안주/간식류	안주/간식류	미반	
A04	14	조리면	조리면	미반	
A04	15	삼각샌드	샌드위치	조리빵	
A04	16	토스트샌드	샌드위치	조리빵	
A04	17	햄버거	햄버거	조리빵	
A04	18	잉글리쉬머핀	머핀/베이글	조리빵	
A04	19	베이글	머핀/베이글	조리빵	
A04	20	롤샌드	롤샌드/핫도그	조리빵	
A04	21	핫도그	롤샌드/핫도그	조리빵	
A04	22	조리샐러드	샐러드	간식	
A04	23	조리면	조리면	간식	
A04	24	냉장피자	냉장피자	간식	
A04	25	기타	기타	간식	
A04	26	일반케익	케익	디저트	
A04	27	조각케익	케익	디저트	
A04	28	프리미엄케익	케익	디저트	
A04	29	떡	떡	디저트	

03

"빅데이터는 정제부터 시작된다."

· 처리 및 분석 Overview



03

"빅데이터는 정제부터 시작된다."

Sqldf패키지를 활용한 데이터 분리

```
library(sqldf)
purchase_A04 <- sqldf("SELECT ID,RCT_NO,PD_S_C FROM purchase WHERE BIZ_UNIT==\'A04\'")
product_A04 <- sqldf("SELECT PD_S_C,PD_H_NM FROM product WHERE BIZ_UNIT==\'A04\'")
purchase_product_A04 <- merge(purchase_A04,product_A04,by="PD_S_C")

purchase_A01 <- sqldf("SELECT ID,RCT_NO,PD_S_C FROM purchase WHERE BIZ_UNIT==\'A01\'")
product_A01 <- sqldf("SELECT PD_S_C,PD_H_NM FROM product WHERE BIZ_UNIT==\'A01\'")
purchase_product_A01 <- merge(purchase_A01,product_A01,by="PD_S_C")


purchase_A02 <- sqldf("SELECT ID,RCT_NO,PD_S_C FROM purchase WHERE BIZ_UNIT==\'A02\'")
product_A02 <- sqldf("SELECT PD_S_C,PD_H_NM FROM product WHERE BIZ_UNIT==\'A02\'")
purchase_product_A02 <- merge(purchase_A02,product_A02,by="PD_S_C")

purchase_A03 <- sqldf("SELECT ID,RCT_NO,PD_S_C FROM purchase WHERE BIZ_UNIT==\'A03\'")
product_A03 <- sqldf("SELECT PD_S_C,PD_H_NM FROM product WHERE BIZ_UNIT==\'A03\'")
purchase_product_A03 <- merge(purchase_A03,product_A03,by="PD_S_C")

purchase_A05 <- sqldf("SELECT ID,RCT_NO,PD_S_C FROM purchase WHERE BIZ_UNIT==\'A05\'")
product_A05 <- sqldf("SELECT PD_S_C,PD_H_NM FROM product WHERE BIZ_UNIT==\'A05\'")
purchase_product_A05 <- merge(purchase_A05,product_A05,by="PD_S_C")
```

Sqldf패키지를 활용

↓

두 테이블을 업종별로 나누기

↓

상품소분류 코드로 join

↓

업종별로 데이터 분리

03

"빅데이터는 정제부터 시작된다."

업종 순위 파악을 위한 자료 구조 변환

표 3. 입력 데이터의 형태

N \ W	1	2	3	4	5	.....	17	18	19	20	21
1	0	0	0	0	1	.....	0	0	0	0	0
2	0	0	1	0	0	.....	0	0	0	0	0
3	0	0	0	0	1	.....	0	0	0	0	0
:	:	:	:	:	:	.....	:	:	:	:	:
1334	0	0	0	0	0	.....	0	0	0	0	0

자료구조 변환 처리 순서

업종별 영수증 번호(N),
구매수량 대분류명(W) 추출

↓

품을 산것은 1,
사지 않은것은 0으로할당

↓

매트릭스 형태로 변환

03

"빅데이터는 정제부터 시작된다."

*업종별 데이터 전처리

백화점 상품 구매횟수 카운트

	가전	가정	기타	남성의류	명품	미분류분상품.(팽한 됨.	스포츠	식품	아동	여성의류	잡화
1	0	0	0	1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	1	0
3	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	1	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	1	0
6	0	0	0	0	0	0	0	1	0	0	0
7	0	0	0	1	0	0	0	0	0	0	0
8	0	1	0	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1	0
11	0	0	0	0	0	0	1	0	0	0	0
12	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	1	0
14	0	0	0	0	0	0	1	0	0	0	0
15	0	1	0	0	0	0	0	0	0	0	0
16	0	1	0	0	0	0	0	0	0	1	0
17	0	1	0	0	0	0	1	0	0	0	0
18	0	0	0	1	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	1	0	0	0
20	0	0	0	1	0	0	0	0	0	0	0
21	0	1	0	0	0	0	0	0	0	0	0

백화점 상품 중분류 기준으로 분류

백화점

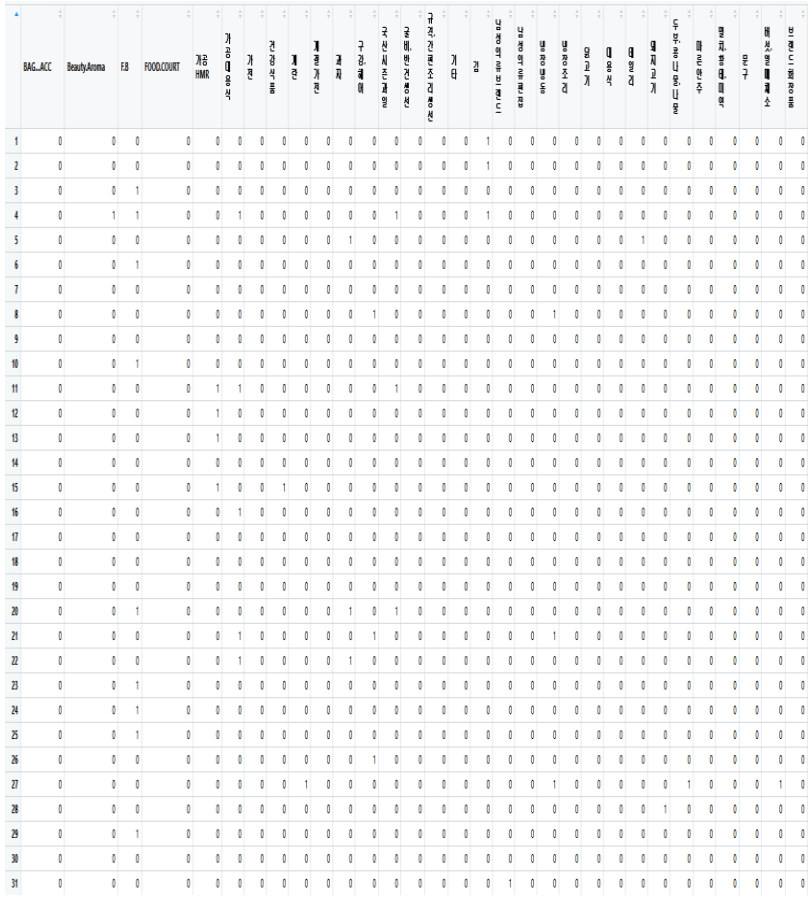
ID	카테고리	품목
1	식품	안주류 플라자 친환경올가 생필(부대).....
2	여성의료	Theory 비트윈 Taste5.1 아울렛 ...
3	잡화	ACC Bloom (1F) ACC Bloom (3F) 캐주얼화 파크랜드 ...
4	스포츠	아웃도아 수영복 휘트니스 가 방 운동용품....
5	가정	생화 이미용 수입주방 주방잡화 수저 스텐 ...
6	남성의류	ANDZ 엠비오 레노마 BON 코모도스퀘어 ...
7	아동	N/B 아동복 L/C 아동복 수입 아동복 고정행사(아동복) ..
8	기타	펫샵 금고 SM TOWN 멘즈셀렉샵 상품군미지정,,,
9	명품	수입잡화 수입의류 직수입 잡화멀티샵 명품잡화 ....
10	가전	국산A/V 국산냉장고.세탁기 수입냉장고 수입세탁기 ...

03

"빅데이터는 정제부터 시작된다."

*업종별 데이터 전처리

대형마트 상품 구매횟수 카운트



대형마트 상품 중분류 기준으로 분류

대형마트

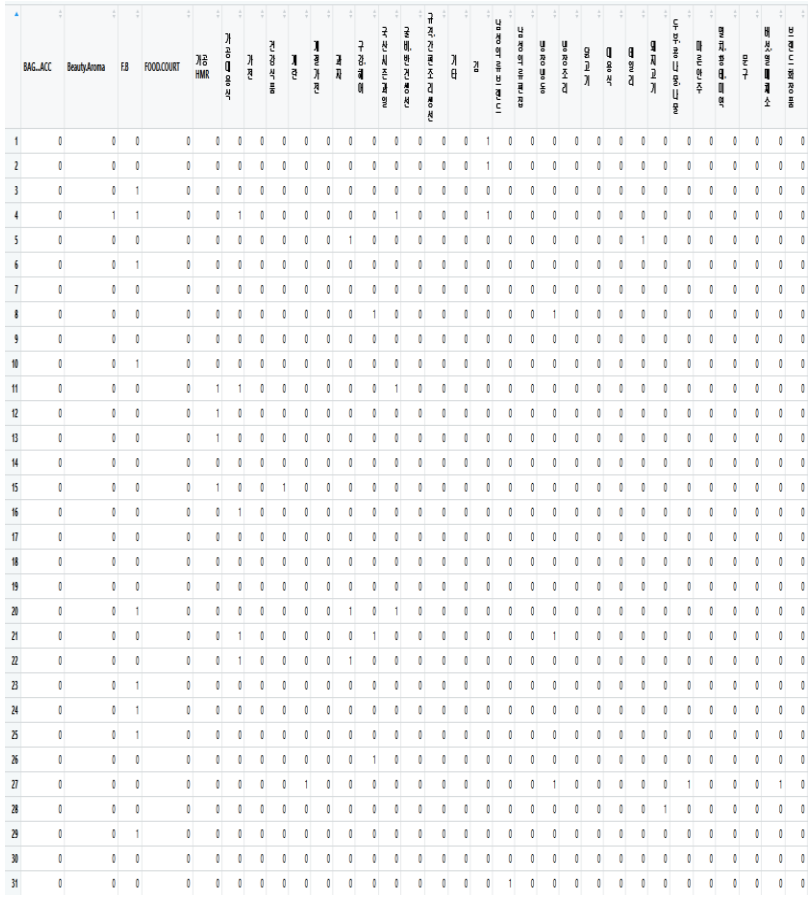
ID	카테고리	품목
2	데일리	기능성우유 친환경우유 어린이우유 멸균흰우유 과일맛우유 커피/초코우유
3	과자	옛날과자 유기농과자 프리미엄비스켓 프리미엄스낵 프리미엄초콜릿 프리
4	가공대용식	봉지라면 컵라면 국수 당면 수입면류 숙면 3분요리류 분말조리식 즉석 죽/
5	냉장냉동	어묵 맛살 포장김치 물만두 군만두 손찜만두 냉동튀김 냉동부침/구이 홈이
6	음료	수입과채혼합음료 수입탄산/이온음료 수입커피/수입혼합차음료 채소음료
7	즉석반찬	나물/부침 반찬류 젓갈류 즉석김치류 즉석국탕/볶음 제철반찬 김장김치
8	조미식품	케찹 마요네즈 간장 고추장 된장 쌈장 춘장 참기름 조미혼합세트 기타고급
9	주류	병소주 페트소주 국산맥주 수입맥주 레드와인 화이트와인 와인선물세트
10	침구	일반아동침구 캐릭터침구 디자인엔 룸베이 메종디오르 바니루이스 브랜드
11	국산시즌과일	PB감귤 감귤기타 노지감귤 레드향 천혜향 하우스감귤 한라봉 황금향 단감
12	열대과일	망고스틴 메론기타 아보카도 용과 허니듀메론 PB바나나 고당도바나나 기타
13	위생용품	시니어용품 유아기저귀 온라인유아기저귀 유아세제 수유용품 유아용품 환
14	브랜드화장품	남성화장품 여성화장품 일반화장품 기타색조화장품 네일 립 베이스 블러셔
15	세제	분말세제 세탁비누 액체세제 세탁보조제 주거청소세제 식기세제 방충제 일
16	계란	메추리알 브랜드란 일반란 차별화계란 가공계란 가공메추리알
17	인스턴트	이유식 분유 이유식 과일잼 수산통조림 축산물통조림 통조림선물세트 스프
18	완구	유아용품 미취학 신생아/유아 아동방구미기 발육기 액세서리 유모차/카시
19	돼지고기	NB돼지고기 PB돼지고기 소포장팩 기능성웰빙돼지고기 품종차별화돼지고
20	가공HMR	냉동디저트HMR 수입냉동식사 냉동식사HMR 냉동면 냉동밥 냉장간편식 냉
21	문구	수정용품 접착용품 게시용품 데스크정리용품 OA용품 서식지 사무용지 오

03

"빅데이터는 정제부터 시작된다."

*업종별 데이터 전처리

슈퍼마켓 상품 구매횟수 카운트



슈퍼마켓 상품 중분류 기준 분류

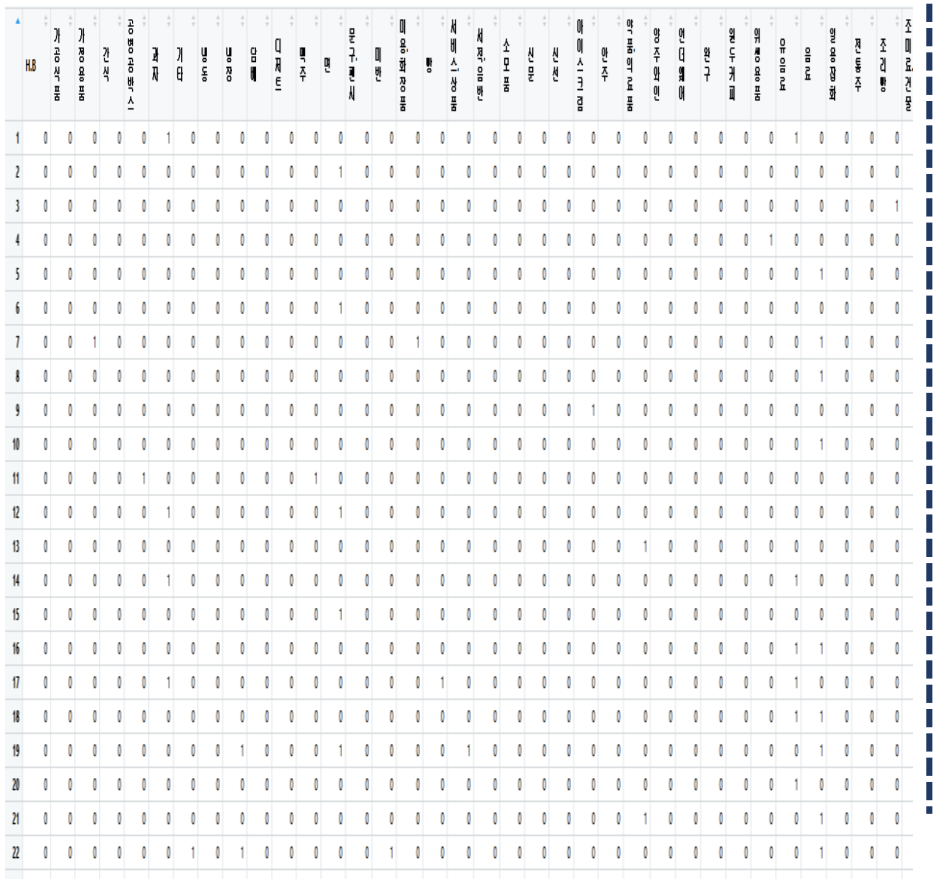
슈퍼마켓		
ID	카테고리	품목
1	채소	원예식물 가지 고추 건고추 ....엉
2	면과자	국수 당면 파스타류 냉면 쌈/또ㄸ
3	낙농	냉장과일야채음료 냉장곡물견과
4	생활잡화	기타이미용가전 기타남성의류세
5	과일	사과 배 감 귤류 유자/모과 복숭(
6	음료	생수 탄산수 사이다 콜라 혼합탄
7	주류	공병/공박스 일반소주 칵테일소ᄎ
8	일상용품	샴푸 린스/컨디셔너 샴푸/린스세
9	축산	한우안심 한우등심 한우채끝 닭윙
10	일반식품	일반비타민 건강보조식품세트 홈
11	냉장	김치류 단무지 햄 베이컨 소시지
12	냉동	얼음 냉동만두 바아이스크림 튜브
13	수산	삼치 국물용건멸치 볶음조림용건
14	조미료류	케첩 마요네즈 상온드레싱 양념소
15	조리식품	...샌드위치 햄버거 기타냉장조리
16	MS	푸딩/젤리 빵/케이크 절임반찬 조
17	곡물	일반미 기능성미 즉석도정미 보ㄹ

03

"빅데이터는 정제부터 시작된다."

*업종별 데이터 전처리

편의점 상품 구매횟수 카운트



편의점 상품 중분류 기준 분류

편의점

ID	카테고리	품목
1	카테고리	목록
2	음료	콜라(소) 콜라(중) 콜라(대) 사이다(소) 사
3	유음료	흰우유(소) 흰우유(중) 흰우유(대) 기능
4	과자	크래커 쿠키 파이 스틱과자 기타 알뜰절
5	미반	삼각김밥 The커진삼각김밥 말이김밥 트
6	담배	&G BAT PMI JTI 시가 직수입
7	면	라면 비빔면 우동 라면(대형컵) 라면(소
8	아이스크림	바 콘 컵 모나카 펜슬 기타 카톤 멀티 바
9	맥주	병맥주(소) 병맥주(중) 병맥주(대) 캔맥주
10	냉장	소시지 햄 소형바 치킨 족발 기타 김치
11	빵	천연발효빵 페스트리 샌드빵 캐릭터빵
12	조리빵	삼각샌드 토스트샌드 햄버거 잉글리쉬더
13	전통주	병소주 팩소주 PET소주 탁주 청주 약주
14	일용잡화	AA사이즈 AAA사이즈 기타사이즈 3단우
15	공병공박스	공병공박스 소주공병 맥주공병
16	안주	오징어(소) 오징어(중) 대용량 쥐포 어육
17	위생용품	소형 중형 대형 팬티라이너 탐폰 4P 박
18	가공식품	즉석밥 즉석덮밥 죽 즉석컵밥 덮밥 소스
19	기타	봉투보증금
20	원두커피	HOT ICE HOT HOT
21	가정용품	에어졸 매트/훈증기 모기향 방충제 기피
22	언더웨어	팬티스타킹 판타롱스타킹 풋커버/앵클삭
23	디저트	일반케익 조각케익 프리미엄케익 떡 선
24	냉동	봉지만두 트레이만두 냉동면 냉동밥 냉
25	양주와인	위스키 기타양주 레드 화이트 샴페인 소
26	완구	일반완구 캡슐완구 스티커 폭죽 용품 파
27	소모품	슬러피원액 커피원액 기타판매용소모품
28	신문	종합일간지 스포츠지 경제지 영자신문

03

"빅데이터는 정제부터 시작된다."

*업종별 데이터 전처리

드러그스토어 상품 구매횟수 카운트

	가정	기초	기타.비상품.	색조	서비스상품	식품	잡화	퍼스널	향수	헬스
1	0	0	0	0	0	0	0	1	0	0
2	0	1	0	0	1	0	0	0	0	0
3	0	0	0	0	0	1	0	1	0	0
4	0	0	0	0	0	1	0	0	0	0
5	0	0	0	0	0	0	1	1	0	0
6	0	1	0	0	0	0	1	0	0	0
7	0	0	0	0	0	0	0	1	0	0
8	0	0	0	0	0	1	0	0	0	0
9	0	1	0	0	0	0	1	0	0	0
10	0	0	1	0	1	1	0	0	0	0
11	0	0	0	0	0	1	0	0	0	0
12	0	1	0	0	0	0	0	1	0	0
13	1	0	0	0	0	0	0	1	0	0
14	0	0	0	1	0	0	0	0	0	0
15	0	0	0	0	0	1	0	0	0	0
16	0	0	0	0	0	0	0	1	0	0
17	0	1	0	0	0	0	0	0	0	0
18	0	1	0	0	0	0	0	1	0	0
19	0	1	0	0	0	0	0	1	0	0
20	0	1	0	0	0	0	0	1	0	0
21	0	0	0	0	0	0	1	0	0	0
22	0	0	0	0	0	0	0	0	1	0
23	0	0	0	0	0	1	0	0	0	0
24	0	0	0	0	0	0	0	0	0	1

드러그스토어 상품 중분류 기준으로 분류

대형마트

ID	카테고리	품목
2	퍼스널	샴푸 린스/컨디셔너 트리트먼트/팩 헤어에센스 헤어스프
3	식품	기타견과류 일반시리얼 즉석스프 기타레토르트 일반껌
4	기초	남성용면도기/날 쉐이빙폼/젤 클래식면도용품 페이셜클
5	잡화	네일케어도구 면봉/화장솜 미용거울 건전지 보안용품 기
6	색조	메이크업베이스/프라이머 BB/파운데이션/컴팩트류 블러
7	서비스상품	봉투보증금 생활잡화균일가
8	가정	롤티슈 각티슈/미용티슈 물티슈 방향제 차량용방향/제취
9	향수	남성향수 여성향수

03

"데이터 분석을 통한 가치발굴"

*데이터 분석기법1 - 분류 분석

STEP 01	STEP 02	STEP 03
" 업종별 상위 품목변수 2개를Y1,Y2로 지정 나머지는 독립변수 "	" 분류분석방법론 진행 (로지스틱 회귀분석,인공신경망,C50) "	" 성능비교 "

03

"빅데이터는 정제부터 시작된다."

*분석기법 1 -로지스틱 회귀분석

*예시) 업종 - A04 편의점

로지스틱 회귀분석의 각 변수중요도

```
Coefficients:
                Values  Std. Err.
(Intercept)  0.7074263 0.01656276
H.B         -1.0727686 0.08453069
가공식품     -0.7255059 0.07710847
가정용품     -0.2814300 0.10183330
간식         -0.2040485 0.15602827
공병공박스   -0.0178530 0.11256966
과자         -0.7738967 0.02271275
기타          0.9059787 0.08431891
냉동          0.1611866 0.12821046
냉장         -0.3836303 0.04110812
담배          0.2537848 0.03160475
디저트       -0.3002972 0.13028921
맥주         -1.6284102 0.04364099
면           -0.8775638 0.03802133
문구.팬시    -1.1654111 0.15130053
미반         -1.0016197 0.03313135
미용.화장품 -1.1858532 0.12223849
빵           -0.6214686 0.05097134
서비스.상품  0.7273431 0.21161122
서적.음반    -0.7355133 0.43916838
소모품        0.6804771 0.22208044
신문         -1.3843762 0.41519853
신선         -0.9657245 0.11568494
아이스크림   -1.5476746 0.04005548
약품.의료품 -0.8498921 0.07957747
양주와인     -0.6035265 0.14151963
언더웨어     -1.6561531 0.12878446
완구         -0.4123764 0.14677250
원두커피     -2.2364288 0.11788759
위생용품     -1.0002906 0.07345925
유음료       -1.5844683 0.02217097
```

상위 품목 간의 상관관계 분석(편의점)

```
> mean(predict_train_purchase_product_A04_1_fit_gm==test_purch
se_product_A04_1$as.factor.test_purchase_product_A04...32..)
[1] 0.7840638
> mean(predict_train_purchase_product_A04_2_fit_gm==test_purch
se_product_A04_2$as.factor.test_purchase_product_A04...31..)
[1] 0.8026002
>
```

-Y1 : 음료 정확도 78.4%
-Y2 : 유.음료 80%
*높은 상관관계

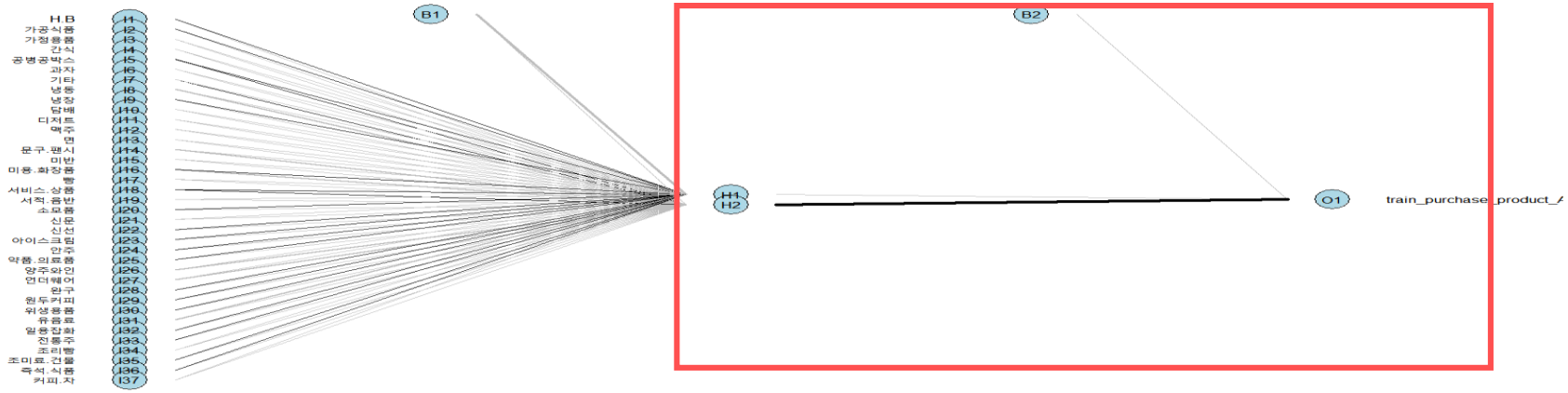
03

"빅데이터는 정제부터 시작된다."

*분석기법2 -인공신경망

*예시) 업종 - A04 편의점

=〉은닉층 개수 2개 지정,가중치 감소모수 5e-04, 반복횟수2000번




```
> mean(pre_model_nnet_A04_1==test_purchase_product_
A04_1$as.factor.test_purchase_product_A04...32..)
[1] 0.8051747
> mean(pre_model_nnet_A04_2==test_purchase_product_
A04_2$as.factor.test_purchase_product_A04...31..)
[1] 0.8094227
```

Y1 음료 정확도 80.5%,
Y2 유음료 81% 도출

*

03

"빅데이터는 정제부터 시작된다."

*분석기법3 - C50

*예시) 업종 - A04 편의점

의사결정 나무의 각 변수중요도

```
Attribute usage:

100.00%  유음료
 68.03%  맥주
 62.05%  아이스크림
 56.25%  원두커피
 55.35%  즉석.식품
 54.04%  문구.팬시
 53.14%  일용잡화
 52.83%  약품.의료품
 51.32%  면
 45.76%  과자
 32.59%  언더웨어
 32.12%  H.B
 31.29%  미반
 27.14%  가공식품
 26.44%  미용.화장품
 26.08%  위생용품
 25.14%  전통주
 24.23%  커피.차
 23.71%  신선
 23.40%  냉장
 21.84%  빵
 20.25%  양주와인
 20.10%  조미료.건물
 19.96%  가정용품
 19.63%  조리빵
 18.88%  안주
```

의사결정 트리로 본 분석결과(편의점)

```
Decision tree:

유음료 > 0:
:...문구.팬시 > 0: 1 (54/20)
:   문구.팬시 <= 0:
:   :...약품.의료품 <= 0: 0 (19677/4216)
:       약품.의료품 > 0:
:       :...면 <= 0: 0 (120/49)
:           면 > 0: 1 (15)
유음료 <= 0:
:...맥주 > 0: 0 (3716/614)
    맥주 <= 0:
    :...아이스크림 > 0: 0 (3609/675)
        아이스크림 <= 0:
        :...원두커피 > 0: 0 (558/60)
            원두커피 <= 0:
            :...즉석.식품 > 0: 0 (1370/233)
                즉석.식품 <= 0:
                :...일용잡화 > 0: 0 (1268/230)
                    일용잡화 <= 0:
                    :...면 > 0: 0 (3322/797)
                        면 <= 0:
                        :...과자 > 0:
                            :...기타 <= 0: 0 (8080/2740)
                            :   기타 > 0: 1 (106/44)
                            과자 <= 0:
                            :...언더웨어 > 0: 0 (289/41)
                                언더웨어 <= 0:
                                :...H.B > 0: 0 (514/92)
                                    H.B <= 0:
                                    :...미반 > 0: 0 (2584/774)
```

Y1 음료 정확도 88.5%

Y2 유음료 87%

03

"데이터 분석을 통한 가치발굴"

*데이터 분석기법2 - 연관 분석

STEP 01

"

대분류와 영수증 아이디
연관 규칙 리스트 작성

"

STEP 02

"

작성된 리스트로
연관 규칙 및 패턴 분석

"

STEP 03

"

연관 규칙 결과 시각화

"

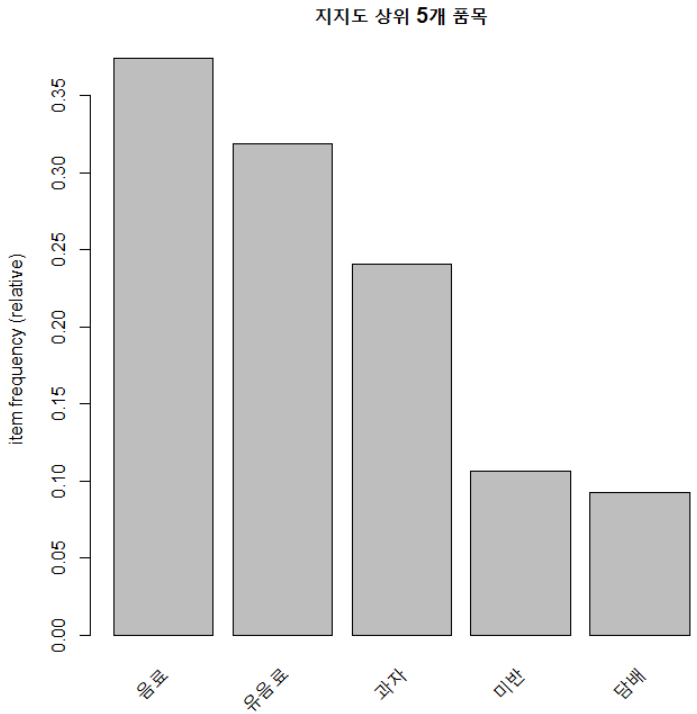
04

"빅데이터,보이는 만큼 활용한다."

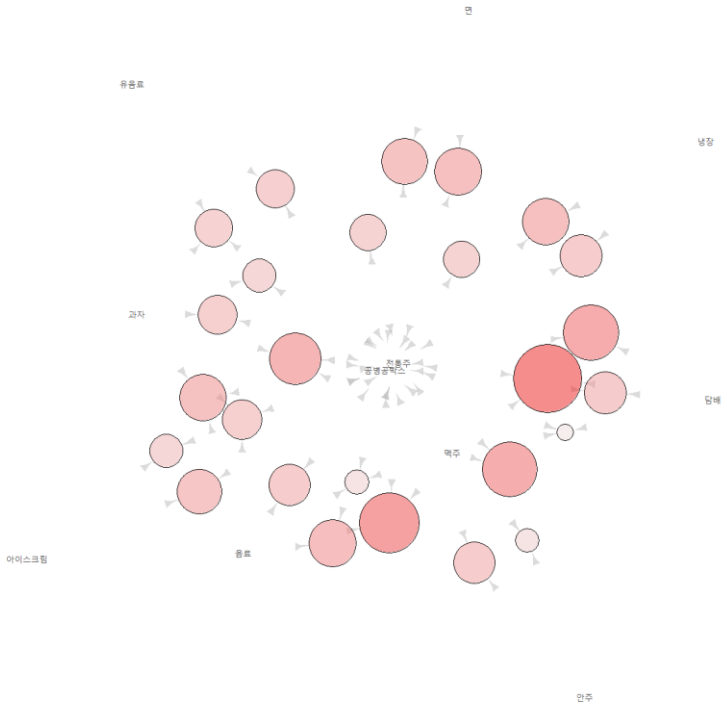
*데이터 분석기법2 - 연관 분석

*예시) 업종 - A04 편의점

지지도 상위 5개 품목



연관규칙 네트워크



04

"빅데이터,보이는 만큼 활용한다."

*데이터 분석기법2 - 연관 분석

*예시) 업종 - A04 편의점

연관 규칙 분석 결과

	lhs		rhs	support	confidence
[1]	{공병공박스}	=>	{전통주}	0.019193389	0.858376511
[2]	{전통주}	=>	{공병공박스}	0.019193389	0.734844751
[3]	{공병공박스,안주}	=>	{전통주}	0.001621976	0.782608696
[4]	{안주,전통주}	=>	{공병공박스}	0.001621976	0.759036145
[5]	{공병공박스,아이스크림}	=>	{전통주}	0.001107063	0.905263158
[6]	{아이스크림,전통주}	=>	{공병공박스}	0.001107063	0.716666667
[7]	{맥주,전통주}	=>	{공병공박스}	0.006449287	0.825370675
[8]	{공병공박스,냉장}	=>	{전통주}	0.002484456	0.889400922
[9]	{냉장,전통주}	=>	{공병공박스}	0.002484456	0.784552846
[10]	{공병공박스,면}	=>	{전통주}	0.002471583	0.918660287
[11]	{면,전통주}	=>	{공병공박스}	0.002471583	0.780487805
[12]	{공병공박스,담배}	=>	{전통주}	0.003346936	0.890410959
[13]	{담배,전통주}	=>	{공병공박스}	0.003346936	0.828025478
[14]	{공병공박스,과자}	=>	{전통주}	0.004737201	0.838268793
[15]	{과자,전통주}	=>	{공병공박스}	0.004737201	0.746450304
[16]	{공병공박스,유음료}	=>	{전통주}	0.002793404	0.868
[17]	{공병공박스,음료}	=>	{전통주}	0.004750074	0.920199501
[18]	{음료,전통주}	=>	{공병공박스}	0.004750074	0.759259259
[19]	{공병공박스,담배,맥주}	=>	{전통주}	0.001042699	0.743119266
[20]	{담배,맥주,전통주}	=>	{공병공박스}	0.001042699	0.89010989
[21]	{과자,맥주,전통주}	=>	{공병공박스}	0.001789323	0.808139535
[22]	{공병공박스,맥주,음료}	=>	{전통주}	0.001531867	0.78807947
[23]	{맥주,음료,전통주}	=>	{공병공박스}	0.001531867	0.85
[24]	{공병공박스,과자,유음료}	=>	{전통주}	0.001132809	0.862745098
[25]	{공병공박스,과자,음료}	=>	{전통주}	0.001647722	0.914285714
[26]	{과자,음료,전통주}	=>	{공병공박스}	0.001647722	0.748538012

LHS->RHS 패턴

SUPPORT 지지도
전체에서 LHS를 구매 후 RHS를 구매한 거래 비율

CONFIDENCE 신뢰도
A가 포함된 거래중 B도 포함한 거래 비율

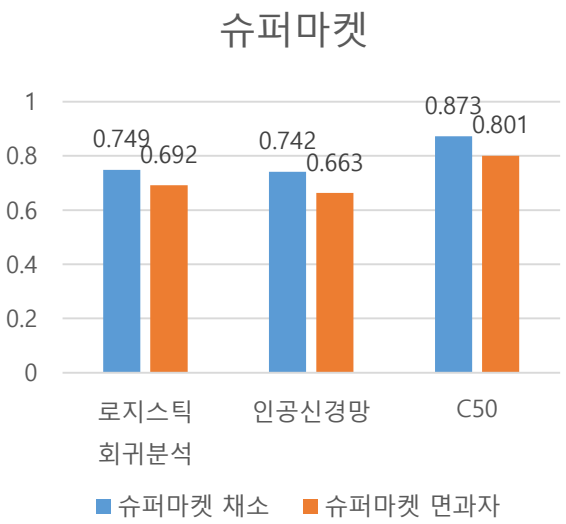
LIFT 향상도
LHS와 RHS 사이의 독립성
(1이하 음의 상관 1이상 양의 상관 1 독립)

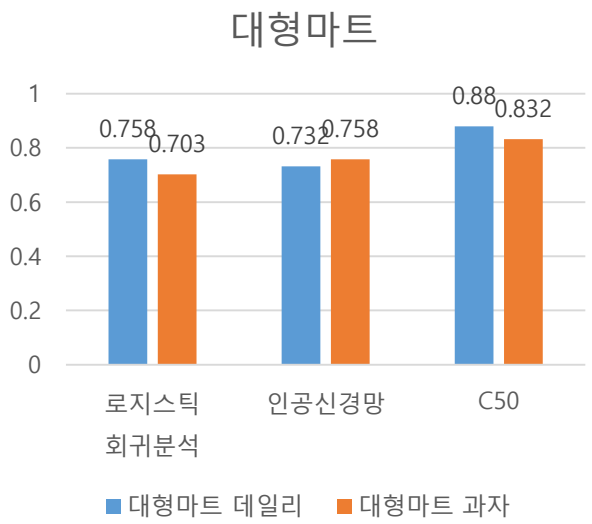
COUNT 패턴이 나타난 횟수

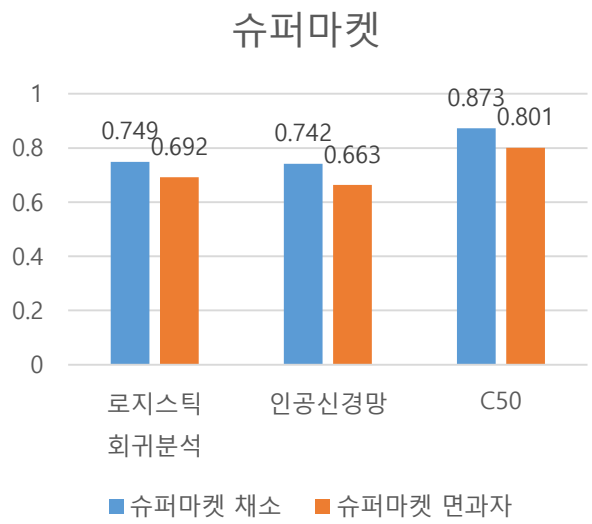
04

"빅데이터,보이는 만큼 활용한다."

분석 도구에 따른 성능 비교



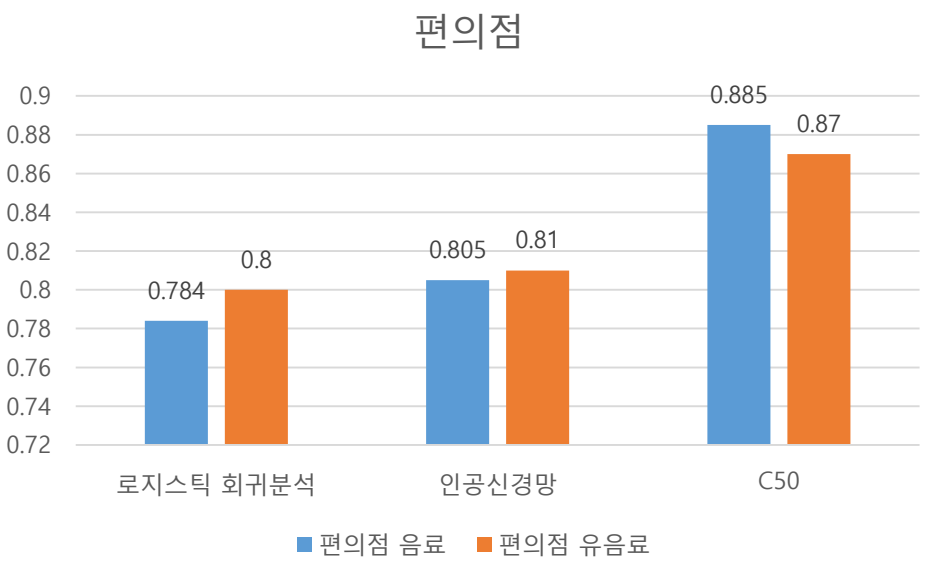


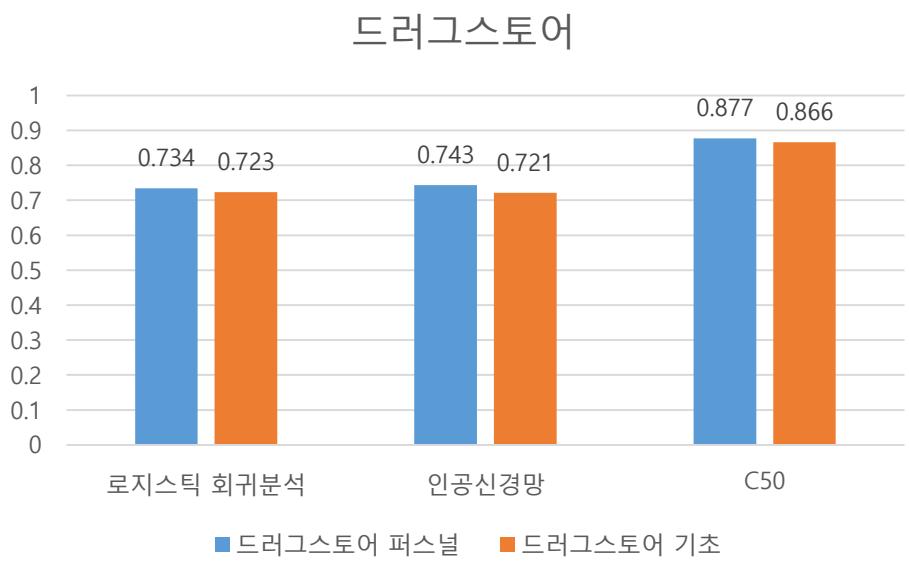


04

"빅데이터,보이는 만큼 활용한다."

분석 도구에 따른 성능 비교2



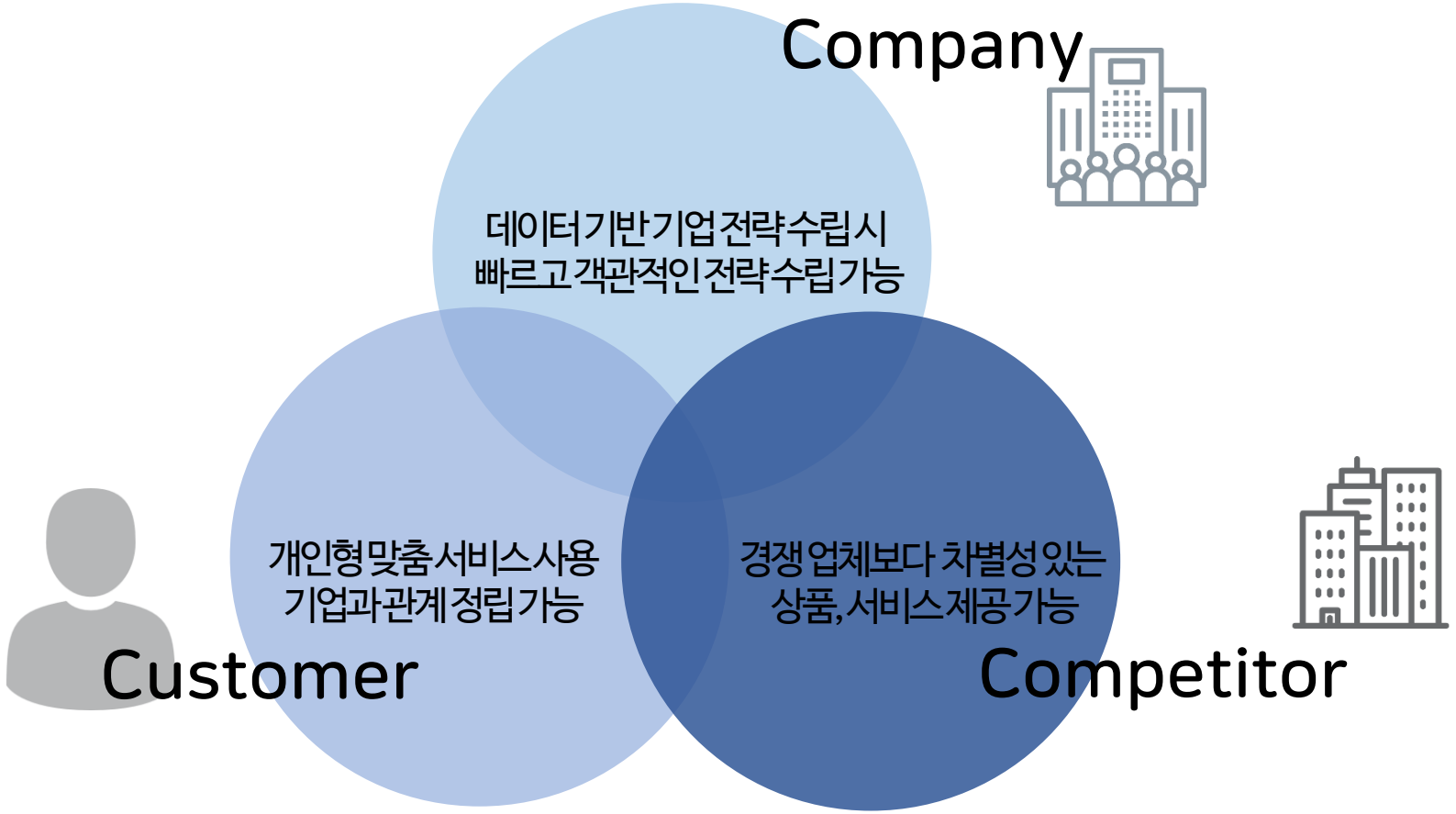


고객의 상품 구매의도를 예측하는 데 있어 데이터마이닝 기법을 적용하여 결과를 비교함으로써 가장 우수한 예측 정확도를 보여주는 것은 C50이다

05

"CRM 기업 승패를 좌우한다."

*서비스 활용 방안 - 3C



기업과 고객의 상생(相生)의 도구로 활용될 가능성 有

06

"빅데이터를 활용한 예측시스템"

· 고객과의 관계 정립(CRM)

분류 분석을 통한 고객 상품
구매 여부 예측

연관성 규칙 분석을 통한 고객 상품
구매 패턴 발견

특정 상품 구매 행동 결정 요인 파악 + 우수한 CRM 전략 수립

06

"빅데이터를 활용한 예측시스템"

·구매 의도 예측 및 추천 시스템

구매 행동 예측을 통한 신사업 발굴

1.고객의 구매행동을 정확하게 파악할 수 있는 능력 획득
2.개인 상품추천시스템을 기업이 보유하고 있는 경우 다양한 사업기회를 발굴

한 명의 고객을 위한 서비스 제공

개인화 상품 추천 시스템을 통해
고객의 신상정보와 상점에서의 구매 행위에 대한 정보를 바탕으로 고객 취향을 반영한 상품이나 서비스를 추천가능

Thank U ☺