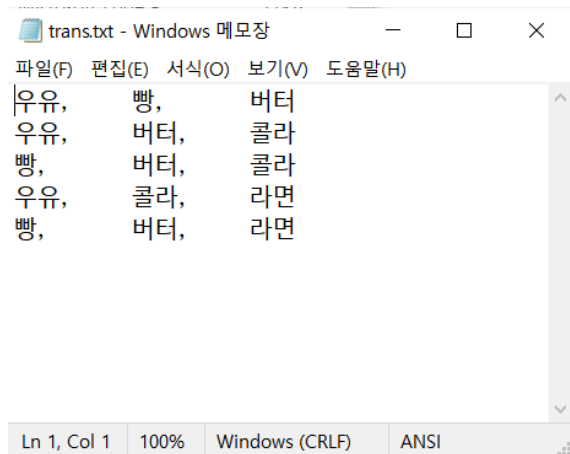


[6주차] 패턴탐색 과제

데이터테크전공 20173204 곽명빈



trans.txt

지지도, 신뢰도, 향상도는 다음과 같다.

지지도

$$A \rightarrow B = \frac{A \text{와 } B \text{를 포함한 거래수}}{\text{전체 거래수}}$$

신뢰도

$$A \rightarrow B = \frac{A \text{와 } B \text{를 포함한 거래수}}{A \text{를 포함한 거래수}}$$

향상도

$$A \rightarrow B = \frac{A \rightarrow B \text{의 신뢰도}}{B \text{를 포함한 거래비율}} = \frac{P(B|A)}{P(B)}$$

1. trans.txt의 거래에서 연관규칙 '우유 -> 콜라'에 대한 지지율, 신뢰도, 향상도를 구하라.

지지율

~~~~~

$$\text{우유} \rightarrow \text{콜라} = \frac{\text{우유와콜라를 포함한거래수}}{\text{전체거래수}} = \frac{2}{5} = 0.4$$

우유,콜라를 포함한 거래수 / 전체거래수 = 2/5 = 0.4

### 신뢰도

우유,콜라를 포함한 거래수 / 우유를 포함한 거래수 = 2/3 = 0.67

$$\text{우유} \rightarrow \text{콜라} = \frac{\text{우유와콜라를 포함한거래수}}{\text{우유를 포함한거래수}} = \frac{2}{3} = 0.67$$

### 향상도

A → B의 신뢰도 / B를포함한거래비율 = 0.67 / 0.6 = 1.12

$$\text{우유} \rightarrow \text{콜라} = \frac{\text{우유와콜라의 신뢰도}}{\text{우유를 포함한거래비율}} = \frac{0.67}{0.6} = 1.12$$

### R로 검증

```
rules <- apriori(dat, parameter = list(support = 0.4, confidence = 0.4))
inspect(rules)
```

```
> inspect(rules)
```

|      | lhs  | rhs     | support | confidence | coverage | lift      | count |
|------|------|---------|---------|------------|----------|-----------|-------|
| [1]  | {}   | => {라면} | 0.4     | 0.4000000  | 1.0      | 1.0000000 | 2     |
| [2]  | {}   | => {빵}  | 0.6     | 0.6000000  | 1.0      | 1.0000000 | 3     |
| [3]  | {}   | => {우유} | 0.6     | 0.6000000  | 1.0      | 1.0000000 | 3     |
| [4]  | {}   | => {콜라} | 0.6     | 0.6000000  | 1.0      | 1.0000000 | 3     |
| [5]  | {}   | => {버터} | 0.8     | 0.8000000  | 1.0      | 1.0000000 | 4     |
| [6]  | {빵}  | => {버터} | 0.6     | 1.0000000  | 0.6      | 1.2500000 | 3     |
| [7]  | {버터} | => {빵}  | 0.6     | 0.7500000  | 0.8      | 1.2500000 | 3     |
| [8]  | {우유} | => {콜라} | 0.4     | 0.6666667  | 0.6      | 1.1111111 | 2     |
| [9]  | {콜라} | => {우유} | 0.4     | 0.6666667  | 0.6      | 1.1111111 | 2     |
| [10] | {우유} | => {버터} | 0.4     | 0.6666667  | 0.6      | 0.8333333 | 2     |
| [11] | {버터} | => {우유} | 0.4     | 0.5000000  | 0.8      | 0.8333333 | 2     |
| [12] | {콜라} | => {버터} | 0.4     | 0.6666667  | 0.6      | 0.8333333 | 2     |
| [13] | {버터} | => {콜라} | 0.4     | 0.5000000  | 0.8      | 0.8333333 | 2     |

## 2. 또한 연관규칙 '빵 & 우유 -> 버터'에 대한 지지율, 신뢰도, 향상도를 구 하여라.

### 지지율

빵,우유, 버터를 포함한 거래수 / 전체거래수

$$\text{빵\&우유} \rightarrow \text{버터} = \frac{\text{빵,우유,버터를 포함한거래수}}{\text{전체거래수}} = \frac{1}{5} = 0.2$$

## 신뢰도

빵,우유,버터를 포함한 거래수 / 빵,우유를 포함한 거래수

$$\text{빵\&우유} \rightarrow \text{버터} = \frac{\text{빵, 우유, 버터를 포함한 거래수}}{\text{빵, 우유를 포함한 거래수}} = \frac{1}{1} = 1$$

## 향상도

A → B의 신뢰도 / B를 포함한 거래비율

$$\text{빵\&우유} \rightarrow \text{버터} = \frac{\text{빵\&우유} \rightarrow \text{버터의 신뢰도}}{\text{버터를 포함한 거래비율}} = \frac{1}{0.8} = 1.25$$

## R로 검증

```
rules <- apriori(dat)
inspect(rules)
```

```
> inspect(rules)
```

|     | lhs      | rhs     | support | confidence | coverage | lift     | count |
|-----|----------|---------|---------|------------|----------|----------|-------|
| [1] | {}       | => {버터} | 0.8     | 0.8        | 1.0      | 1.000000 | 4     |
| [2] | {빵}      | => {버터} | 0.6     | 1.0        | 0.6      | 1.250000 | 3     |
| [3] | {라면, 빵}  | => {버터} | 0.2     | 1.0        | 0.2      | 1.250000 | 1     |
| [4] | {라면, 버터} | => {빵}  | 0.2     | 1.0        | 0.2      | 1.666667 | 1     |
| [5] | {라면, 우유} | => {콜라} | 0.2     | 1.0        | 0.2      | 1.666667 | 1     |
| [6] | {라면, 콜라} | => {우유} | 0.2     | 1.0        | 0.2      | 1.666667 | 1     |
| [7] | {빵, 우유}  | => {버터} | 0.2     | 1.0        | 0.2      | 1.250000 | 1     |
| [8] | {빵, 콜라}  | => {버터} | 0.2     | 1.0        | 0.2      | 1.250000 | 1     |