In [42]: import pandas as pd import matplotlib.pyplot as plt from mlxtend.frequent\_patterns import apriori,association\_rules In [43]: movie\_data=pd.read\_csv('my\_movies.csv') movie\_data V5 Sixth Sense Gladiator LOTR1 Harry Potter1 Patriot LOTR2 Harry Potter2 LOTR Braveheart Green Mile **V1** V2 ٧3 V4 Out[43]: LOTR2 Sixth Sense LOTR1 Harry Potter1 Green Mile 1 0 1 0 0 1 Gladiator Patriot Braveheart NaN 0 0 0 NaN 2 LOTR1 LOTR2 0 0 0 0 0 0 0 0 NaN 1 1 NaN NaN Sixth Sense Gladiator Patriot NaN NaN Gladiator 0 0 0 0 0 0 0 Sixth Sense NaN 1 1 1 Patriot NaN Gladiator Sixth Sense Patriot NaN NaN 0 0 Harry Potter1 Harry Potter2 NaN NaN 0 0 0 0 0 1 0 0 0 NaN 1 Gladiator Patriot NaN NaN NaN 1 0 1 Gladiator 0 0 0 0 0 0 0 Patriot Sixth Sense NaN NaN 1 1 1 Sixth Sense LOTR Gladiator Green Mile NaN In [44]: movie\_data.shape Out[44]: (10, 15) In [45]: movie\_data.info() <class 'pandas.core.frame.DataFrame'> RangeIndex: 10 entries, 0 to 9 Data columns (total 15 columns): Column Non-Null Count \_\_\_\_\_ -----0 V1 10 non-null object 10 non-null object 1 V2 ٧3 7 non-null object 3 V4 2 non-null object 4 V5 1 non-null object Sixth Sense 5 10 non-null int64 Gladiator 10 non-null int64 LOTR1 10 non-null int64 Harry Potter1 10 non-null 8 int64 10 non-null int64 Patriot L0TR2 10 non-null int64 11 Harry Potter2 10 non-null int64 10 non-null int64 12 LOTR 10 non-null int64 13 Braveheart 14 Green Mile 10 non-null int64 dtypes: int64(10), object(5) memory usage: 1.3+ KB In [46]: movie=movie\_data.iloc[:,5:] movie Sixth Sense Gladiator LOTR1 Harry Potter1 Patriot LOTR2 Harry Potter2 LOTR Braveheart Green Mile Out[46]: 0 0 0 0 0 1 1 0 1 1 1 0 0 0 0 2 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 6 0 1 0 0 0 0 0 1 0 0 8 1 0 1 0 0 0 0 0 0 0 1 Apriori Algorithm Association rules with 10% Support and 70% Confidence In [47]: #With 10% support frequent\_itemsets=apriori(movie, min\_support=0.1, use\_colnames=True) frequent\_itemsets Out[47]: support itemsets 0 (Sixth Sense) 1 0.7 (Gladiator) 2 0.2 (LOTR1) 3 0.2 (Harry Potter1) 4 0.6 (Patriot) 5 0.2 (LOTR2) (Harry Potter2) 6 0.1 7 0.1 (LOTR) 8 (Braveheart) 9 0.2 (Green Mile) 10 0.5 (Gladiator, Sixth Sense) (LOTR1, Sixth Sense) 11 0.1 12 0.1 (Harry Potter1, Sixth Sense) 13 0.4 (Patriot, Sixth Sense) 14 0.1 (LOTR2, Sixth Sense) 15 0.1 (LOTR, Sixth Sense) 16 (Green Mile, Sixth Sense) 0.6 (Patriot, Gladiator) 17 18 0.1 (Gladiator, LOTR) 19 0.1 (Gladiator, Braveheart) 20 0.1 (Gladiator, Green Mile) 21 0.1 (LOTR1, Harry Potter1) (LOTR1, LOTR2) 22 0.2 23 0.1 (LOTR1, Green Mile) (LOTR2, Harry Potter1) 24 (Harry Potter2, Harry Potter1) 25 0.1 (Harry Potter1, Green Mile) 26 0.1 27 0.1 (Patriot, Braveheart) 28 0.1 (LOTR2, Green Mile) (LOTR, Green Mile) 29 0.1 30 (Patriot, Gladiator, Sixth Sense) 31 0.1 (Gladiator, LOTR, Sixth Sense) 32 (Green Mile, Gladiator, Sixth Sense) 33 0.1 (LOTR1, Harry Potter1, Sixth Sense) 34 0.1 (LOTR1, LOTR2, Sixth Sense) 35 0.1 (Green Mile, LOTR1, Sixth Sense) (LOTR2, Harry Potter1, Sixth Sense) 36 0.1 37 0.1 (Green Mile, Harry Potter1, Sixth Sense) 38 0.1 (Green Mile, LOTR2, Sixth Sense) 39 0.1 (Green Mile, LOTR, Sixth Sense) 40 (Patriot, Gladiator, Braveheart) 41 0.1 (Gladiator, LOTR, Green Mile) 42 0.1 (LOTR1, LOTR2, Harry Potter1) (LOTR1, Harry Potter1, Green Mile) 43 0.1 (LOTR1, LOTR2, Green Mile) 44 0.1 45 0.1 (LOTR2, Harry Potter1, Green Mile) 46 0.1 (Green Mile, Gladiator, LOTR, Sixth Sense) (LOTR1, LOTR2, Harry Potter1, Sixth Sense) 47 0.1 48 (Green Mile, LOTR1, Harry Potter1, Sixth Sense) 49 (Green Mile, LOTR1, LOTR2, Sixth Sense) 0.1 (Green Mile, LOTR2, Harry Potter1, Sixth Sense) 50 0.1 (LOTR1, LOTR2, Harry Potter1, Green Mile) 51 0.1 (Harry Potter1, Green Mile, LOTR2, LOTR1, Sixt... 52 In [48]: # With 70% confidence rules=association\_rules(frequent\_itemsets, metric='lift', min\_threshold=0.7) Out[48]: antecedents consequents antecedent support consequent support support confidence lift leverage conviction 0 (Gladiator) (Sixth Sense) 0.7 0.6 0.714286 1.190476 0.08 1.40 0.7 0.5 0.833333 1.190476 0.08 1.80 (Sixth Sense) (Gladiator) 0.6 (LOTR1) (Sixth Sense) 0.6 0.500000 0.833333 -0.02 0.80 (Sixth Sense) 0.166667 0.833333 0.6 0.2 0.96 (LOTR1) 0.1 -0.02 4 (Harry Potter1) (Sixth Sense) 0.2 0.6 0.1 0.500000 0.833333 -0.02 0.80 ... 245 (Harry Potter1) (LOTR1, LOTR2, Green Mile, Sixth Sense) 0.2 0.1 0.500000 5.000000 0.08 1.80 1.80 (LOTR1, LOTR2, Harry Potter1, Sixth Sense) 0.2 0.1 0.1 0.500000 5.000000 0.08 246 (Green Mile) 247 (LOTR2) (LOTR1, Harry Potter1, Green Mile, Sixth Sense) 0.2 0.1 0.500000 5.000000 0.08 1.80 5.000000 0.500000 0.2 0.1 0.1 0.08 1.80 248 (LOTR1) (LOTR2, Harry Potter1, Green Mile, Sixth Sense) (Sixth Sense) (LOTR1, LOTR2, Harry Potter1, Green Mile) 0.6 0.1 0.166667 1.666667 0.04 1.08 250 rows × 9 columns In [49]: rules[rules.lift>1] antecedents consequents antecedent support consequent support support confidence lift leverage conviction Out[49]: (Gladiator) (Sixth Sense) 0.7 0.714286 1.190476 0.08 1.40 0.833333 1.190476 0.08 1.80 (Sixth Sense) 0.6 0.7 (Gladiator) (Patriot) (Sixth Sense) 0.6 0.6 0.666667 1.111111 0.04 1.20 0.666667 1.111111 0.6 0.04 1.20 (Sixth Sense) (Patriot) 0.6 0.4 10 (LOTR) (Sixth Sense) 0.6 1.000000 1.666667 245 (Harry Potter1) (LOTR1, LOTR2, Green Mile, Sixth Sense) 0.2 0.1 0.500000 5.000000 0.08 1.80 (LOTR1, LOTR2, Harry Potter1, Sixth Sense) 0.08 1.80 (Green Mile) 0.2 0.1 0.1 0.500000 5.000000 246 247 (LOTR2) (LOTR1, Harry Potter1, Green Mile, Sixth Sense) 0.2 0.1 0.1 0.500000 5.000000 0.08 1.80 0.08 1.80 248 (LOTR1) (LOTR2, Harry Potter1, Green Mile, Sixth Sense) 0.2 0.1 0.1 0.500000 5.000000 (Sixth Sense) (LOTR1, LOTR2, Harry Potter1, Green Mile) 0.6 0.1 0.1 0.166667 1.666667 0.04 1.08 236 rows × 9 columns In [50]: # visualization of obtained rule plt.scatter(rules['support'], rules['confidence']) plt.xlabel('support') plt.ylabel('confidence') plt.show() 1.0 0.8 confidence 0.4 0.2 0.1 0.2 0.3 0.4 0.5 0.6 1. Association rules with 5 % Support and 90% confidence In [51]: # with 5% support frequent\_itemsets2=apriori(movie, min\_support=0.05, use\_colnames=True) frequent\_itemsets2 Out[51]: support itemsets 0 0.6 (Sixth Sense) 1 0.7 (Gladiator) 2 0.2 (LOTR1) (Harry Potter1) 0.2 4 0.6 (Patriot) 0.2 (LOTR2) 6 (Harry Potter2) 0.1 7 0.1 (LOTR) 8 0.1 (Braveheart) 9 0.2 (Green Mile) 10 (Gladiator, Sixth Sense) 0.5 11 0.1 (LOTR1, Sixth Sense) (Harry Potter1, Sixth Sense) 12 0.1 13 0.4 (Patriot, Sixth Sense) 0.1 (LOTR2, Sixth Sense) 14 (LOTR, Sixth Sense) 15 0.1 16 0.2 (Green Mile, Sixth Sense) 17 0.6 (Patriot, Gladiator) 18 0.1 (Gladiator, LOTR) 19 0.1 (Gladiator, Braveheart) (Gladiator, Green Mile) 20 0.1 21 0.1 (LOTR1, Harry Potter1) (LOTR1, LOTR2) 22 0.2 (LOTR1, Green Mile) 23 0.1 24 0.1 (LOTR2, Harry Potter1) 25 (Harry Potter2, Harry Potter1) 0.1 (Harry Potter1, Green Mile) 26 0.1 27 0.1 (Patriot, Braveheart) 28 0.1 (LOTR2, Green Mile) (LOTR, Green Mile) 29 0.1 (Patriot, Gladiator, Sixth Sense) 30 0.4 31 (Gladiator, LOTR, Sixth Sense) 0.1 32 0.1 (Green Mile, Gladiator, Sixth Sense) 33 (LOTR1, Harry Potter1, Sixth Sense) 0.1 34 0.1 (LOTR1, LOTR2, Sixth Sense) 35 (Green Mile, LOTR1, Sixth Sense) 0.1 (LOTR2, Harry Potter1, Sixth Sense) 36 0.1 37 0.1 (Green Mile, Harry Potter1, Sixth Sense) (Green Mile, LOTR2, Sixth Sense) 38 0.1 39 0.1 (Green Mile, LOTR, Sixth Sense) (Patriot, Gladiator, Braveheart) 40 0.1 41 (Gladiator, LOTR, Green Mile) 0.1 (LOTR1, LOTR2, Harry Potter1) 42 0.1 43 (LOTR1, Harry Potter1, Green Mile) 0.1 44 0.1 (LOTR1, LOTR2, Green Mile) 45 (LOTR2, Harry Potter1, Green Mile) 0.1 (Green Mile, Gladiator, LOTR, Sixth Sense) 46 0.1 (LOTR1, LOTR2, Harry Potter1, Sixth Sense) 47 0.1 (Green Mile, LOTR1, Harry Potter1, Sixth Sense) 48 0.1 49 0.1 (Green Mile, LOTR1, LOTR2, Sixth Sense) (Green Mile, LOTR2, Harry Potter1, Sixth Sense) 50 51 (LOTR1, LOTR2, Harry Potter1, Green Mile) 52 0.1 (Harry Potter1, Green Mile, LOTR2, LOTR1, Sixt... In [52]: # With 90% confidence rules=association\_rules(frequent\_itemsets, metric='lift', min\_threshold=0.9) rules consequents antecedent support consequent support support confidence Out[52]: antecedents lift leverage conviction (Gladiator) 0.714286 1.190476 0 0.08 1.40 (Sixth Sense) 0.7 0.6 1 (Sixth Sense) (Gladiator) 0.6 0.7 0.833333 1.190476 0.08 1.80 0.666667 1.111111 (Patriot) (Sixth Sense) 0.6 0.6 0.4 0.04 1.20 0.6 0.666667 1.111111 0.04 1.20 (Sixth Sense) (Patriot) 0.6 4 (LOTR) 0.1 1.000000 1.666667 0.04 (Sixth Sense) 0.1 0.6 inf (Harry Potter1) (LOTR1, LOTR2, Green Mile, Sixth Sense) 0.2 0.1 0.500000 5.000000 0.08 1.80 233 0.1 0.2 0.1 0.500000 5.000000 0.08 1.80 234 (Green Mile) (LOTR1, LOTR2, Harry Potter1, Sixth Sense) 0.1 (LOTR2) (LOTR1, Harry Potter1, Green Mile, Sixth Sense) 0.500000 5.000000 235 0.2 0.08 1.80 0.1 0.1 0.500000 5.000000 0.08 1.80 236 (LOTR1) (LOTR2, Harry Potter1, Green Mile, Sixth Sense) 0.2 0.1 0.1 (LOTR1, LOTR2, Harry Potter1, Green Mile) 0.6 237 (Sixth Sense) 0.1 0.1 0.166667 1.666667 0.04 1.08 238 rows × 9 columns In [53]: rules[rules.lift>1] antecedents antecedent support consequent support support confidence lift leverage conviction Out[53]: 0 (Gladiator) (Sixth Sense) 0.7 0.6 0.5 0.714286 1.190476 0.08 1.40 (Sixth Sense) (Gladiator) 0.6 0.7 0.5 0.833333 1.190476 0.08 1.80 2 (Patriot) (Sixth Sense) 0.6 0.4 0.666667 1.111111 0.04 1.20 0.6 (Sixth Sense) 0.04 1.20 3 (Patriot) 0.6 0.6 0.4 0.666667 1.111111 4 (LOTR) (Sixth Sense) 0.1 0.6 0.1 1.000000 1.666667 0.04 inf (Harry Potter1) (LOTR1, LOTR2, Green Mile, Sixth Sense) 0.2 0.1 0.500000 5.000000 0.08 1.80 233 0.1 0.500000 5.000000 0.08 234 (Green Mile) (LOTR1, LOTR2, Harry Potter1, Sixth Sense) 0.2 0.1 0.1 1.80 235 (LOTR2) (LOTR1, Harry Potter1, Green Mile, Sixth Sense) 0.2 0.1 0.500000 5.000000 0.08 1.80 0.1 0.2 0.500000 5.000000 0.08 1.80 236 (LOTR1) (LOTR2, Harry Potter1, Green Mile, Sixth Sense) 0.1 0.1 237 (Sixth Sense) (LOTR1, LOTR2, Harry Potter1, Green Mile) 0.6 0.1 0.1 0.166667 1.666667 0.04 1.08 236 rows × 9 columns In [54]: plt.scatter(rules['support'], rules['confidence']) plt.xlabel('support') plt.ylabel('confidence') plt.show() 1.0 • 0.8 0.4 0.2 0.5 0.2 0.3 0.4 0.6 0.1 support Association rules for 20% support and 80 % confidence In [55]: # with 20% support frequent\_itemsets2=apriori(movie, min\_support=0.20, use\_colnames=True) frequent\_itemsets2 Out[55]: itemsets support 0.6 (Sixth Sense) 2 0.2 (LOTR1) 0.2 (Harry Potter1) 3 4 0.6 (Patriot) 5 0.2 (LOTR2) 6 0.2 (Green Mile) 7 0.5 (Gladiator, Sixth Sense) 8 0.4 (Patriot, Sixth Sense) 0.2 9 (Green Mile, Sixth Sense) 10 0.6 (Patriot, Gladiator) 11 (LOTR1, LOTR2) 0.2 12 0.4 (Patriot, Gladiator, Sixth Sense) In [39]: # With 80% confidence rules2=association\_rules(frequent\_itemsets2, metric='lift', min\_threshold=0.8) consequents antecedent support consequent support support confidence lift leverage conviction Out[39]: antecedents 0 (Gladiator) 0.7 0.6 0.5 0.714286 1.190476 0.08 1.4 (Sixth Sense) 0.833333 1.190476 1 0.7 0.5 1.8 (Sixth Sense) (Gladiator) 0.6 0.08 2 (Patriot) (Sixth Sense) 0.6 0.6 0.4 0.666667 1.111111 0.04 1.2 3 0.666667 1.111111 1.2 (Sixth Sense) 0.6 0.4 0.04 (Patriot) 0.6 4 (Green Mile) (Sixth Sense) 0.2 0.6 0.2 1.000000 1.666667 0.08 inf 5 (Sixth Sense) 0.2 0.2 0.333333 1.666667 1.2 (Green Mile) 0.6 0.08 6 (Patriot) (Gladiator) 0.6 0.7 0.6 1.000000 1.428571 0.18 inf 7 0.857143 1.428571 (Gladiator) (Patriot) 0.7 0.6 0.6 0.18 2.8 8 (LOTR1) (LOTR2) 0.2 0.2 0.2 1.000000 5.000000 0.16 inf 9 1.000000 5.000000 (LOTR2) (LOTR1) 0.2 0.2 0.2 0.16 inf 10 (Patriot, Gladiator) (Sixth Sense) 0.6 0.6 0.4 0.666667 1.111111 0.04 1.2 (Patriot, Sixth Sense) 0.7 1.000000 1.428571 0.12 inf 11 (Gladiator) 0.4 0.4 12 (Gladiator, Sixth Sense) (Patriot) 0.5 0.6 0.4 0.800000 1.333333 0.10 2.0 0.666667 1.333333 13 1.5 (Patriot) (Gladiator, Sixth Sense) 0.6 0.5 0.4 0.10 14 (Gladiator) (Patriot, Sixth Sense) 0.7 0.4 0.4 0.571429 1.428571 0.12 1.4 15 (Sixth Sense) (Patriot, Gladiator) 0.6 0.6 0.4 0.666667 1.111111 0.04 1.2 In [56] rules2[rules2.lift>1] consequents antecedent support consequent support support confidence Out[56]: antecedents lift leverage conviction 0 0.714286 1.190476 (Gladiator) (Sixth Sense) 0.7 0.6 0.5 0.08 1.4 1 (Gladiator) 0.6 0.7 0.5 0.833333 1.190476 0.08 1.8 (Sixth Sense) (Patriot) 2 0.6 0.4 0.666667 1.111111 0.04 1.2 (Sixth Sense) 0.6 3 (Sixth Sense) 0.6 0.6 0.4 0.666667 1.111111 0.04 1.2 (Patriot) (Green Mile) 1.000000 1.666667 4 0.2 0.6 0.2 0.08 inf (Sixth Sense) 0.333333 1.666667 5 (Sixth Sense) (Green Mile) 0.6 0.2 0.2 0.08 1.2 6 (Patriot) 0.7 1.000000 1.428571 (Gladiator) 0.6 0.6 0.18 inf 0.857143 1.428571 7 (Gladiator) (Patriot) 0.7 0.6 0.6 0.18 2.8 8 (LOTR1) (LOTR2) 0.2 0.2 1.000000 5.000000 0.2 0.16 inf 1.000000 5.000000 (LOTR2) (LOTR1) 0.2 0.2 10 (Patriot, Gladiator) (Sixth Sense) 0.6 0.6 0.666667 1.111111 0.04 1.2 1.000000 1.428571 inf 11 (Patriot, Sixth Sense) (Gladiator) 0.4 0.7 0.4 0.12 **12** (Gladiator, Sixth Sense) (Patriot) 0.5 0.800000 1.333333 0.10 2.0 0.6 0.4 0.666667 1.333333 13 (Patriot) (Gladiator, Sixth Sense) 0.6 0.5 0.4 0.10 1.5 0.571429 1.428571 14 (Gladiator) (Patriot, Sixth Sense) 0.7 0.4 0.4 0.12 1.4 0.6 15 (Sixth Sense) (Patriot, Gladiator) 0.6 0.4 0.666667 1.111111 0.04 1.2 In [57]: plt.scatter(rules2['support'], rules2['confidence']) plt.xlabel('support') plt.ylabel('confidence') plt.show() 1.0 0.9 0.8 0.7 0.6 0.5 0.4 0.3 0.25 0.30 0.35 0.40 0.45 0.50 0.55 0.60 In [ ]: