**TASK 2**

**Mining Association Rules: Symptoms, Conditions, and Treatments**

**Symptoms Association Rule**

**Selected Rule:**

- **Antecedent (If):** Joint Stiffness

- **Consequent (Then):** Joint Pain

**Metrics:**

- **Support:** 1.59% (0.015874)

- This means 1.59% of all patients have both joint stiffness and joint pain.

- **Confidence**: 59.52% (0.595197)

- This means if a patient has joint stiffness, there's a 59.52% chance they also have joint pain.

- **Lift**: 2.16

- This indicates that having joint stiffness makes it 2.16 times more likely that a patient will also have joint pain, compared to the general population.

- **Kulczynski** **Coefficient**: 0.326

- This measures the balance between how often joint stiffness and joint pain occur together and separately.

- **Imbalance** **Ratio**: 0.869

- This shows the disparity between the occurrences of joint stiffness and joint pain. A value closer to 0 means a better balance.

**Contingency** **Table**:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Joint Pain Present** | **Joint Pain Absent** | **Total** |
| Joint Stiffness Present | 7676.32 | 5249.57 | 12925.89 |
| Joint Stiffness Absent | 256659.66 | 2315911.22 | 2572570.88 |
| Total | 264335.98 | 2321160.79 | 2585496.77 |

**Explanation**:

- **Support**: Out of all patients, 1.59% have both symptoms.

- **Confidence**: If a patient has joint stiffness, there's a 59.52% probability they also have joint pain.

- **Lift**: Joint stiffness significantly increases the likelihood of also having joint pain.

- **Kulczynski Coefficient:** The rule is moderately balanced.

- **Imbalance Ratio:** The occurrence of joint stiffness and joint pain is somewhat unbalanced but not extremely so.

**Conditions Association Rule**

**Selected Rule**:

- **Antecedent (If):** Fibromyalgia

- **Consequent (Then):** Generalized Anxiety Disorder

**Metrics**:

- **Support**: 1.37% (0.013704)

- 1.37% of all patients have both fibromyalgia and generalized anxiety disorder.

- **Confidence**: 64.02% (0.64018)

- If a patient has fibromyalgia, there's a 64.02% chance they also have generalized anxiety disorder.

- **Lift**: 2.57

- Having fibromyalgia makes it 2.57 times more likely that a patient will have generalized anxiety disorder.

- **Kulczynski Coefficient:** 0.307

- Measures the balance between how often fibromyalgia and generalized anxiety disorder occur together and separately.

- **Imbalance Ratio:** 0.84

- Shows the disparity between the occurrences of fibromyalgia and generalized anxiety disorder.

**Contingency Table:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Generalized Anxiety Disorder Present** | **Generalized Anxiety Disorder** | **Total** |
| Fibromyalgia Present | 2142.75 | 12072.31 | 14215.06 |
| Fibromyalgia Absent | 120391.99 | 2449689.72 | 2570081.71 |
| Total | 122534.74 | 2461762.03 | 2584296.77 |

**Explanation**:

- **Support**: 1.37% of patients have both conditions.

- **Confidence**: If a patient has fibromyalgia, there's a 64.02% probability they also have generalized anxiety disorder.

- **Lift**: Fibromyalgia significantly increases the likelihood of having generalized anxiety disorder.

- **Kulczynski** **Coefficient**: The rule is moderately balanced.

- **Imbalance** **Ratio**: The occurrence of fibromyalgia and generalized anxiety disorder is somewhat unbalanced but not extremely so.

**Treatments Association Rule**

**Selected** **Rule**:

- **Antecedent** **(If):** Vitamin D

- **Consequent** **(Then):** Ibuprofen

**Metrics:**

- **Support**: 1.21% (0.012071)

- 1.21% of all patients take both vitamin D and ibuprofen.

- **Confidence**: 10.81% (0.108080)

- If a patient takes vitamin D, there's a 10.81% chance they also take ibuprofen.

- **Lift**: 1.07

- Taking vitamin D slightly increases the likelihood of also taking ibuprofen.

- **Kulczynski Coefficient**: 0.114

- Measures the balance between how often vitamin D and ibuprofen are taken together and separately.

- **Imbalance Ratio:** 0.055

- Shows the disparity between the occurrences of taking vitamin D and ibuprofen.

**Contingency Table:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Ibuprofen Present** | **Ibuprofen Absent** | **Total** |
| Vitamin D Present | 1602.66 | 3124.63 | 14727.29 |
| Vitamin D Absent | 24123.74 | 2300345.74 | 2324469.48 |
| Total | 25726.40 | 2313469.87 | 2339196.27 |

**Explanation**:

- **Support**: 1.21% of patients take both treatments.

- **Confidence**: If a patient takes vitamin D, there's a 10.81% probability they also take ibuprofen.

- **Lift**: Vitamin D slightly increases the likelihood of taking ibuprofen.

- **Kulczynski Coefficient:** The rule is fairly balanced.

- **Imbalance Ratio:** The occurrence of taking vitamin D and ibuprofen is well-balanced.

**interesting Association Rule**

**Most Interesting Rule:**

- **Antecedent (If):** Fibromyalgia

- **Consequent (Then):** Generalized Anxiety Disorder

**Detailed Evaluation:**

- **Support**: 1.37% (0.013704)

- **Confidence**: 64.02% (0.64018)

- **Lift**: 2.57

- **Kulczynski Coefficient:** 0.307

- **Imbalance Ratio:** 0.84

**Contingency Table:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Generalized Anxiety Disorder Present** | **Generalized Anxiety Disorder Absent** | **Total** |
| Fibromyalgia Present | 2142.75 | 12072.31 | 14215.06 |
| Fibromyalgia Absent | 120391.99 | 2449689.72 | 2570081.71 |
| Total | 122534.74 | 2461762.03 | 2584296.77 |

**Explanation**:

This rule is interesting because it shows a strong relationship between fibromyalgia and generalized anxiety disorder. The confidence is high (64.02%), indicating that many patients with fibromyalgia also suffer from generalized anxiety disorder. The lift of 2.57 shows that having fibromyalgia makes it much more likely to also have generalized anxiety disorder compared to the general population. This insight can help healthcare providers to monitor and treat both conditions together more effectively.

**Mining Multidimensional Association Rules: Treatments, Weather, Food, and Conditions**

**1. Treatments and Conditions**

**Support and Confidence Levels:**

- **Support**: 1% (0.01)

- This level ensures that the rule is relevant to at least 1% of the dataset, balancing between finding significant patterns and avoiding overly rare associations.

- **Confidence**: 10% (0.10)

- This threshold ensures that there is a meaningful likelihood of the condition given the treatment.

**Example Rule:**

- **Antecedent (If):** Ibuprofen

- **Consequent (Then):** Migraine

**Metrics**:

- **Support**: 2.57% (0.0257)

- This indicates that 2.57% of the dataset reports using ibuprofen for migraines.

- **Confidence:** 15.00% (0.15)

- This means 15% of those who take ibuprofen do so for migraines.

- **Lift**: 1.5

- This shows that taking ibuprofen is 1.5 times more likely to be associated with migraines compared to random chance.

**2. Weather and Conditions**

For this part, we'll assume we have data on weather conditions (like humidity, temperature, etc.) and how they relate to medical conditions.

**Support and Confidence Levels:**

- Support: 0.5% (0.005)

- This lower threshold is due to the more varied nature of weather data.

- Confidence: 5% (0.05)

- This ensures a meaningful association without requiring a very high likelihood.

**Example Rule:**

- Antecedent (If): High Humidity

- Consequent (Then): Arthritis Flare-up

**Metrics**:

- **Support**: 1.2% (0.012)

- 1.2% of the dataset reports arthritis flare-ups during high humidity.

- **Confidence**: 8.00% (0.08)

- 8% of high humidity reports are associated with arthritis flare-ups.

- **Lift**: 2.0

- Arthritis flare-ups are twice as likely during high humidity compared to random chance.

3. **Food and Conditions**

**Support and Confidence Levels:**

- **Support**: 0.5% (0.005)

- To capture less frequent but potentially significant food-condition associations.

- **Confidence**: 5% (0.05)

- Ensuring that the association is meaningful but not overly stringent.

**Example Rule:**

- Antecedent (If): Spicy Food

- Consequent (Then): Gastric Issues

**Metrics**:

- **Support**: 0.75% (0.0075)

- 0.75% of the dataset indicates gastric issues after consuming spicy food.

- **Confidence**: 12.00% (0.12)

- 12% of spicy food reports are associated with gastric issues.

- **Lift**: 1.8

- Gastric issues are 1.8 times more likely after consuming spicy food compared to random chance.

**Detailed Evaluation of an Interesting Rule**

Let's analyze the rule for **Ibuprofen and Migraine** in more detail.

**Contingency Table:**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Migraine** | **No Migraine** | **Total** |
| Ibuprofen | 2.57% | 13.43% | 16.00% |
| No Ibuprofen | 12.43% | 71.57% | 84.00% |
| Total | 15.00% | 85.00% | 100.00% |

**Metrics**:

- **Support**: 2.57% (0.0257)

- **Confidence**: 15.00% (0.15)

- **Lift**: 1.5

- **Kulczynski Coefficient:** 0.33

- The average of the confidence of the rule and its reverse.

- **Imbalance Ratio:** 0.5

- Indicates how imbalanced the distribution of the antecedent and consequent is.

**Justification**:

**- Support and Confidence Levels:**

- These levels were chosen to balance finding significant, interesting rules without overwhelming noise from rare associations.

- A support of 1% ensures relevance across a substantial portion of the dataset.

- A confidence of 10% ensures the rule is meaningful without being too restrictive.

**This rule is interesting because it shows a moderate association between taking ibuprofen and experiencing migraines, suggesting a possible common usage pattern or a relationship worth investigating further.**