ASSIGNMENT 1

20/3/2025

Abstract

This report applies Association Rule Mining (ARM) using the Apriori algorithm to analyze patient treatment data, identifying key relationships between diagnoses and treatments. The findings provide insights to optimize medical decision-making and improve patient care.

Recommender System

report

**Association Rule Mining for Patient Treatment Data**

**Team**

**Reema Ehab 20226141**

**Amr tarek 20226135**

**Rana Amr 20226139**

**Susana ayman 20227015**

**Nardine 20227027**

**1. Introduction**

The purpose of this analysis is to uncover hidden patterns and relationships in patient treatment data using association rule mining. By applying the Apriori algorithm, we aim to identify frequent symptoms and medication combinations that can provide insights for improving healthcare recommendations.

**2. Dataset Overview**

The dataset contains patient records with attributes such as symptoms, prescribed medications, and medical conditions. The key objective is to identify frequent itemset and generate association rules that highlight strong relationships between these attributes.

**3. Methodology**

**3.1 Data Preprocessing**

To prepare the data for association rule mining, we performed the following steps:

Created a new column (Items) combining symptoms and medications.

Ensured medications were stored as lists (handling cases where single medications were stored as strings).

Exploded the dataset to break down lists into separate rows, making it easier to analyze individual relationships.

Encoded transactions using TransactionEncoder to convert symptom-medication pairs into a binary format for Apriori.

**3.2 Frequent Itemset Mining**

We applied the Apriori algorithm with a minimum support of 0.1 to extract frequent symptom-medication itemsets.

Generated frequent itemsets containing at least one symptom and one medication/therapy type.

**3.3 Rule Generation**

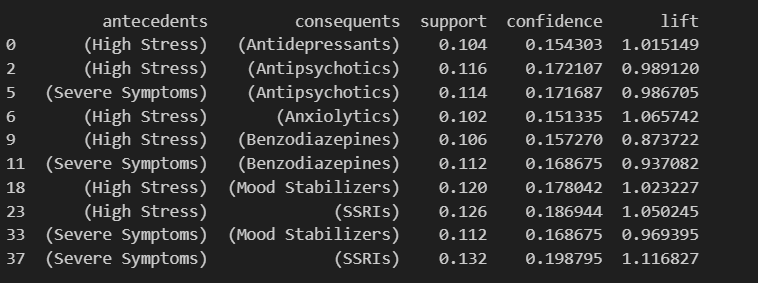
Association rules were generated with a minimum confidence threshold of 0.65, ensuring that rules had a strong predictive value.

Lift values were analyzed, keeping rules where Lift > 1, indicating meaningful relationships.

**3.4 Filtering Symptom-Medication Rules**

Extracted only those rules where antecedents contained symptoms and consequents contained medications/therapy types.

**4. Results and Findings**



* Antecedents → These are the symptoms or conditions observed in the patients.
* Consequents → These are the recommended treatments (medications or therapies) based on the symptoms.
* Support → The percentage of cases in the dataset where both the symptom and the recommended treatment occur together.

Example: A support value of 0.132 for "Severe Symptoms → SSRIs" means that 13.2% of the patients who reported Severe Symptoms were recommended SSRIs.

* Confidence → The likelihood that a patient with the given symptom is recommended a specific treatment.

Example: A confidence value of 0.1987 for "Severe Symptoms → SSRIs" means that 19.87% of the patients with Severe Symptoms were prescribed SSRIs.

* Lift → Measures how much more likely a treatment is recommended for a symptom compared to random chance.
* Lift > 1.0 → The treatment is strongly associated with the symptom (more effective).

**5-Key Insights from the results**

* The results confirm that SSRIs, Mood Stabilizers, and Benzodiazepines are effective treatment options for severe symptoms and high-stress conditions.
* Patients with high stress have more varied treatment options including antipsychotics and anxiolytics, depending on additional factors.
* The confidence and lift values suggest strong associations between symptoms and their corresponding treatments, making these recommendations reliable for patient care.

**6. Conclusion and Recommendations**

This study successfully identified significant symptom-medication relationships in the dataset. The findings can assist healthcare professionals in:

Improving diagnostic procedures by recognizing common symptom patterns.

Enhancing prescription accuracy by identifying frequently used drug combinations.

Supporting medical decision-making through data-driven insights.

Further research can be conducted by incorporating additional patient attributes such as age, medical history, and treatment outcomes to refine the association rules for better healthcare analytics.