Assignment 4

Can Binay

2025-10-27

Topic : Multi-Resistance Antibiotic Susceptibility

Summary

Multi-Resistance Antibiotic Susceptibility dataset is about bacterias tested against various antibiotics. It includes patient observations. There are informations such as age, gender, health risks, and antibiotics resistance of bacteria. It looks into multi-drug resistance (MDR), where there is resistance to three or more groups of drugs. It also research links between risks like diabetes or inpatient stays and resistance.

The dataset has been structured to allow researchers to study:

- Patterns of multi-drug resistance (MDR) across bacterial species
- The relationship between patient risk factors (e.g., diabetes, hypertension, hospitalization history) and resistance
- Epidemiological trends of resistance across antibiotic families

The dataset includes 15+ antibiotics spanning 4 major families, representing commonly prescribed agents in clinical practice. Results are coded primarily as:

- R → Resistant
- S → Susceptible

with occasional alternative encodings such as i, Intermediate, r, s, NA, missing.

β-lactams (Beta-lactams)

Target bacterial cell wall synthesis by inhibiting penicillin-binding proteins (PBPs).

- AMX/AMP → Amoxicillin / Ampicillin (aminopenicillins)
- AMC → Amoxicillin + Clavulanic Acid (β-lactam/β-lactamase inhibitor)
- CZ → Cefazolin (first-generation cephalosporin)
- FOX → Cefoxitin (second-generation cephamycin)
- CTX/CRO → Cefotaxime / Ceftriaxone (third-generation cephalosporins)
- IPM → Imipenem (a carbapenem, considered a last-resort treatment)

Aminoglycosides (Aminosides)

Bind to the 30S ribosomal subunit, inhibiting protein synthesis.

- GEN → Gentamicin
- AN → Amikacin (broader spectrum, often effective against gentamicin-resistant strains)

Quinolones / Fluoroquinolones

Inhibit bacterial DNA gyrase and topoisomerase IV, interfering with DNA replication.

- Acide nalidixique → Nalidixic Acid (first-generation quinolone, Gram-negative coverage)
- OFX → Ofloxacin (fluoroquinolone, broad spectrum)
- CIP → Ciprofloxacin (fluoroquinolone, widely used for urinary tract infections)

Other Important Agents

Include drugs from various classes used for Gram-positive, Gram-negative, or multidrug-resistant infections.

- C → Chloramphenicol (broad-spectrum, rarely used due to toxicity but still tested)
- Co-trimoxazole → Combination of Trimethoprim + Sulfamethoxazole (folate pathway inhibitor)
- Furanes → Nitrofurantoin (commonly used for urinary tract infections)
- Colistine → Colistin (polymyxin class, last-line therapy for carbapenem-resistant organisms)

Identification & Demographics

- ID → Unique strain identifier (e.g., S290)
- Souche → Bacterial species name (e.g., Escherichia coli, Klebsiella pneumoniae, Pseudomonas aeruginosa)
- Age → Patient age (numeric)
- Gender → Encoded as M / F

Risk Factors

- Diabetes → (Yes, No, Y, N, True, 1, ?)
- Hypertension → (Yes, No, missing, ?)
- Hospital_before → (Yes, No) (prior hospitalization within the last 6 months)
- Infection_Freq → Frequency of recurrent infections

(Never → no recurrence, Rarely → <2 times per year, Regularly → <6 times per year, Often → >=6 times per year)

Antibiotic Susceptibility Test (AST) Results

Each column corresponds to one antibiotic (e.g., GEN, CIP, IPM)

Results: (R, S, i, Intermediate, NA, missing)

Dataset Description

Multi-Resistance Antibiotic Susceptibility

This dataset provides a comprehensive collection of bacterial isolates tested against a wide panel of antibiotics. Each row represents a unique bacterial strain isolated from a patient sample, along with associated demographic information, clinical risk factors, and antibiotic susceptibility test results.

Rows: ~10,000 isolates

Columns: ~25 features (including demographics, risk factors, and AST results)

Coverage: Multiple bacterial species, diverse antibiotics, heterogeneous patients

This dataset taken from Kaggle. Link in below.

Link: https://www.kaggle.com/datasets/adilimadeddinehosni/multi-resistance-antibiotic-susceptibility/data