HAIMLC701 AI & ML in Healthcare

2.0		**AI, ML, Deep Learning and Data Mining Methods for Healthcare**	10
	2.1	Knowledge discovery and Data Mining, ML, Multi classifier Decision Fusion, Ensemble Learning, Meta-Learning and other Abstract Methods.	
	2.2	Evolutionary Algorithms, Illustrative Medical Application-Multiagent Infectious Disease Propagation and Outbreak Prediction, Automated Amblyopia Screening System etc.	
	2.3	Computational Intelligence Techniques, Deep Learning, Unsupervised learning, dimensionality reduction algorithms.	

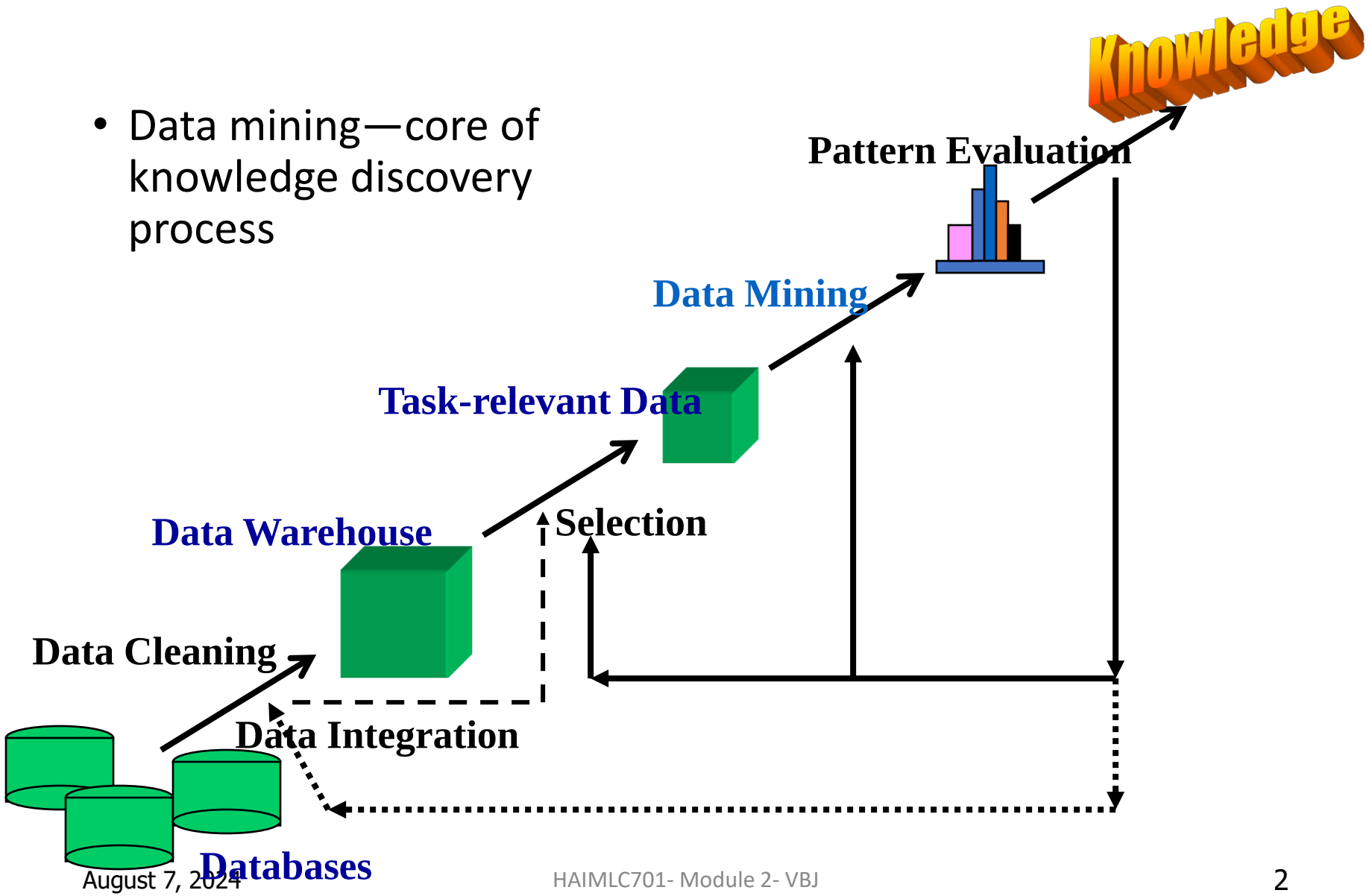
CO Mapped: CO2-Apply advanced AI and Computational Intelligence techniques for Healthcare Problems (L3)

Textbook to refer:

1. Arjun Panesar, "Machine Learning and AI for Healthcare", A Press.
2. Arvin Agah, "Medical applications of Artificial Systems ", CRC Press

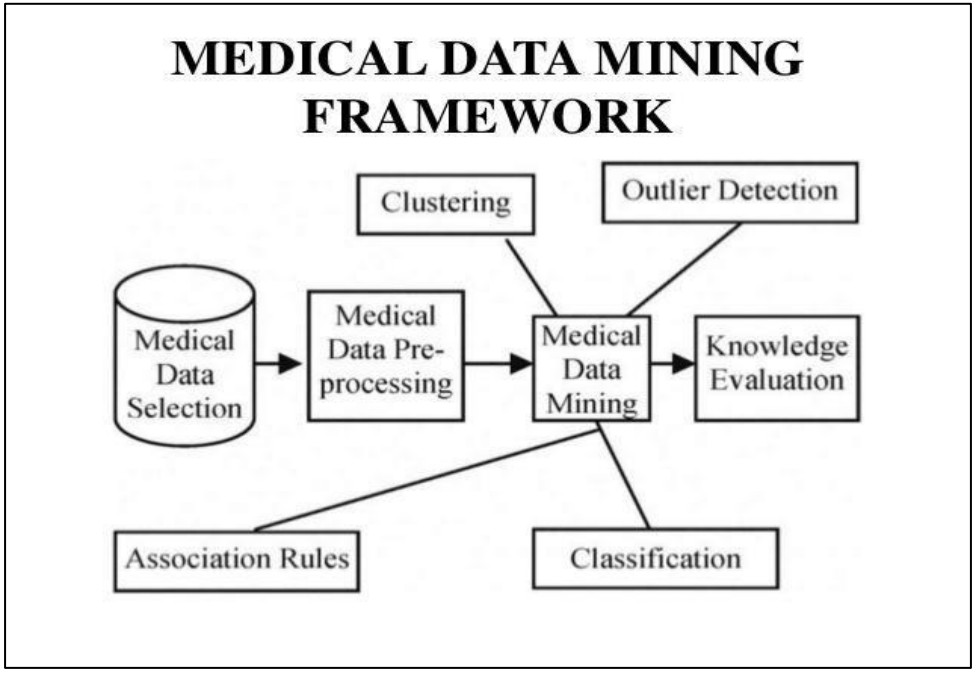
Knowledge Discovery (KDD) Process

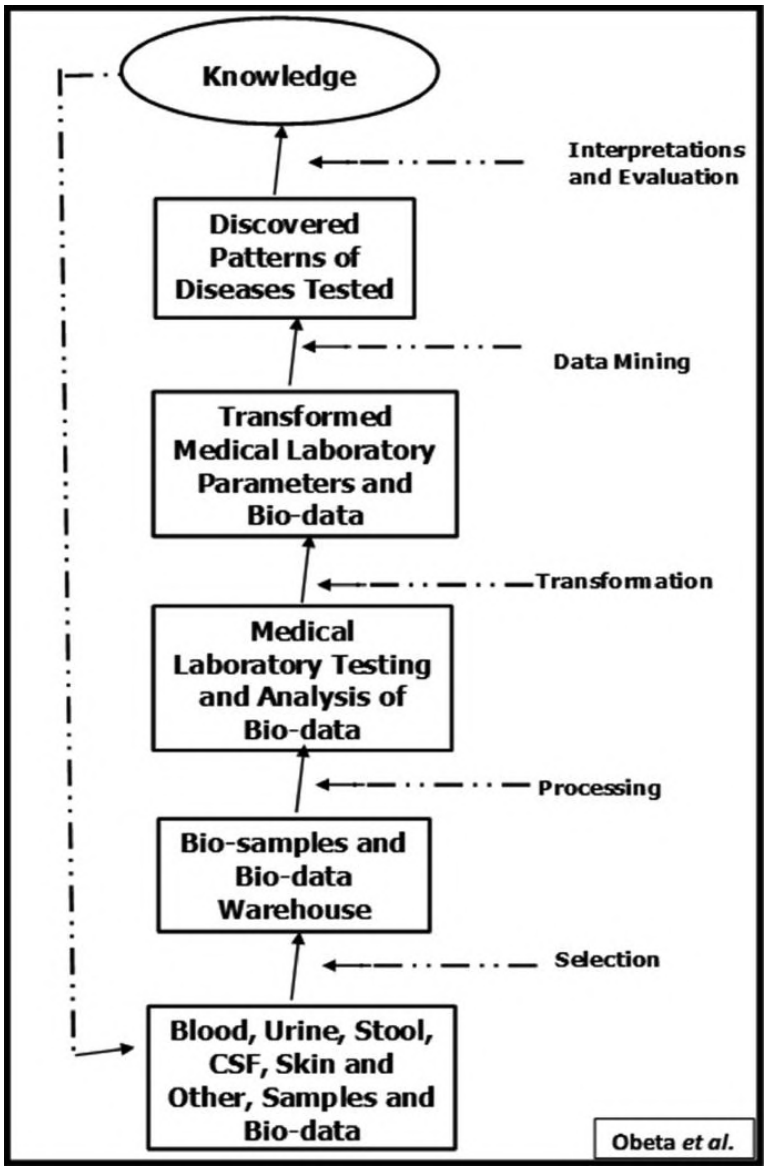
- Data mining—core of knowledge discovery process



Knowledge Discovery and Data Mining in Healthcare

- Data mining (knowledge discovery from data)
  - Extraction of interesting (non-trivial, implicit, previously unknown and potentially useful) patterns or knowledge from huge amount of data
  - Data mining: a misnomer?



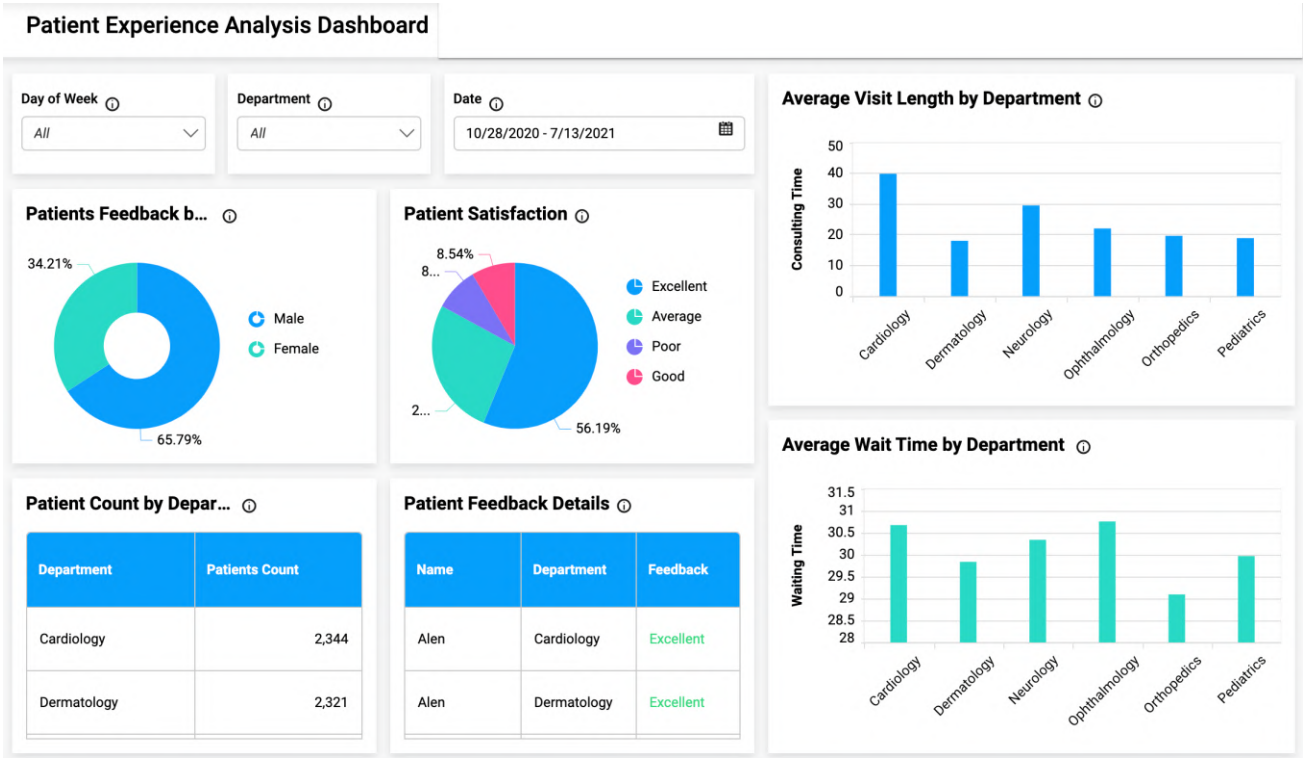


Data Mining Techniques in Healthcare

- Medical data has a lot of information buried within it that will reveal patterns relating to successes and failures in clinical operations
- Association or relationship analysis
  - if a group of patients with specific symptoms is steadily associated with certain prescribed medications they acquire in pharmacies during a preset season, pharmacists can use this information to manage their stock
- Sequence analysis
  - consequential flows of facts or events all patients of the group also share similar lifestyles, chronic diseases, or other health features. With this knowledge, physicians can offer preventive care
- Classification
  - Cases can be compared with each other to be verified as falling within a certain class, to identify differences and apply necessary algorithms and protocols, or to screen out and readdress unmatching data

Data Mining Techniques in Healthcare

- Visualization
  - Building different charts and graphs, such as Gantt charts, pie charts, bubble charts, treemaps, scatter plots, density map- identify trends, patterns, spikes, and declines in certain healthcare parameters or events.
- Clustering
- Forecasting and predicting



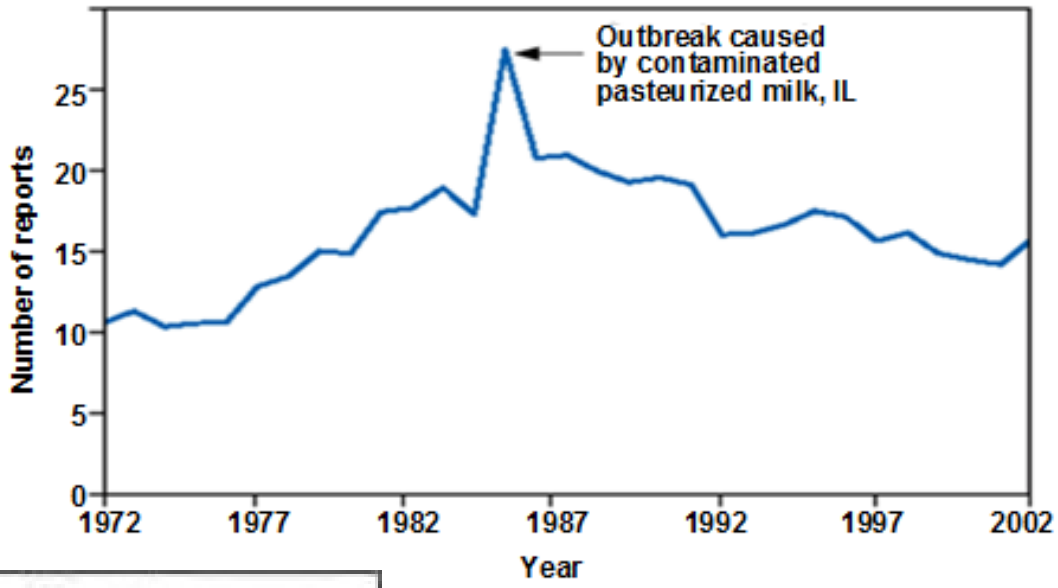
Challenges and Issues

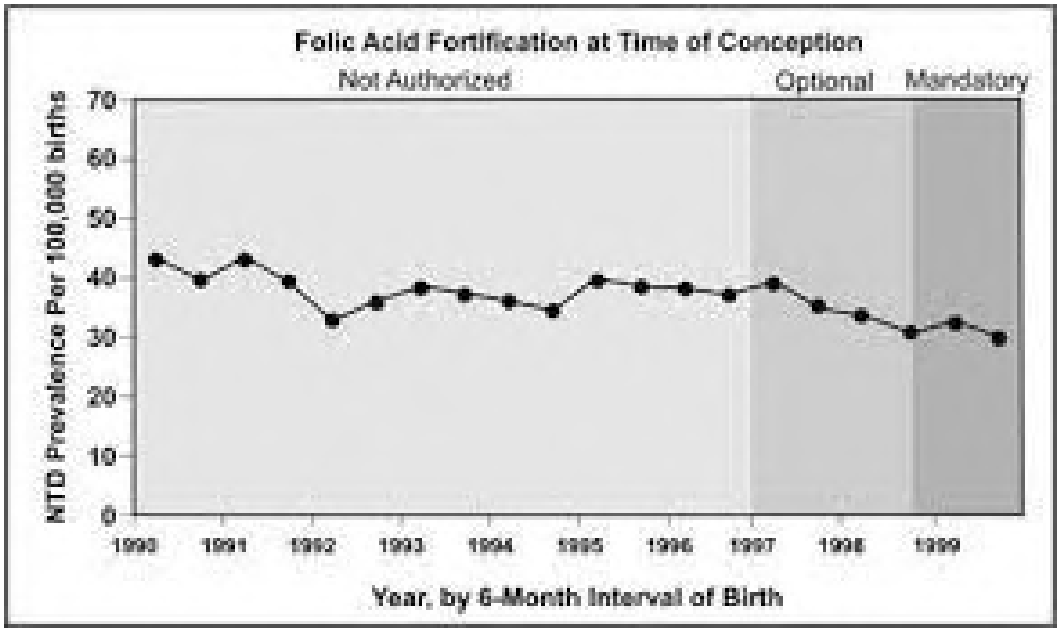
- Infrastructure
  - Lack of information technology sophistication and some historical clinician skepticism have hindered the ability to analyse data adequately
- Data
  - In the medical field, overemphasizing aggregate data can have dire consequences for a patient
- Quality Assurance
  - Since the quality of the data in the data warehouse affects the quality of the decisions being made, it is essential to use data quality management methods in the prototyping phase before building the warehouse
- Segmenting versus Sampling
  - Assume a health insurance company has a data warehouse and wants to find patterns for patient claims
- Privacy and Access
  - Removing patient identifiers from the database makes sense for organizations interested in looking at their data in the aggregate
- End Users
  - Doctors might be even less tolerant to forced change than most users since they're accustomed to a high degree of professional autonomy
- Inadequate tool support
- Scalability

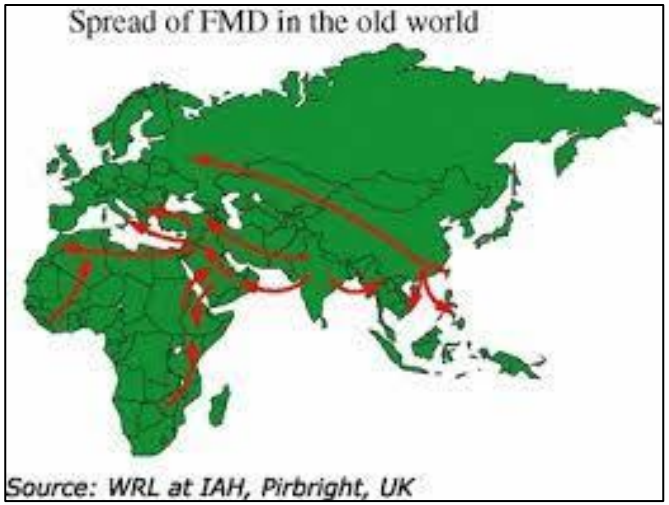
Data Mining in Healthcare: Examples

- Epidemiology patterns: discovery and prognosis
  - With the help of medical data mining, doctors and other healthcare specialists can…
    - Monitor the number of disease occurrences / how an epidemic or disease outbreak scales over a period of time
    - Mark considerable fluctuations in correlation with seasons, location, selected patient groups (i.e., gender, age, health features)
    - Predict changes in the epidemiologic situation and manage response according to prior experience and/or educated expectations
    - Study hidden trends, relationships, or patterns in epidemiological situations
    - Research and predict the spread of disease in terms of expected timelines, affected areas, risks, potential numbers of severe vs. mild cases, and epidemic endpoints

Epidemiology is the study and analysis of the distribution, patterns and determinants of health and disease conditions in a defined population.


Outbreak caused by contaminated pasteurized milk, IL
Number of reports
0
5
10
15
20
25
1972
1977
1982
1987
1992
1997
2002
Year


Folic Acid Fortification at Time of Conception
Not Authorized
Optional
Mandatory
NTD Prevalence Per 100,000 births
0
10
20
30
40
50
60
70
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
Year, by 6-Month Interval of Birth


Spread of FMD in the old world
Source: WRL at IAH, Pirbright, UK

Data Mining in Healthcare: Examples

- Personalized disease course and treatment forecast
- With clinical data mining, it's easy to…
  - Classify patients into separate groups and identify the most frequent and/or severe symptoms and complaints relevant to every specific group
  - Identify the most successful treatment protocols for every patient group
  - Correlate health events (like a chronic disease) with the frequency of specific symptoms and their severity
  - Forecast the development of a disease in an individual patient based upon their health conditions and medical background
  - Find and study new and significant relationships between health conditions, symptoms, treatment plans, medical methods, and outcomes

Data Mining in Healthcare: Examples

- Medical knowledge research and automatic diagnostics
- Software can be trained to automatically…
  - Discover and report all possible relationships between facts while identifying interesting phenomena and provide valuable conclusions
  - Recognize the most probable diagnoses in a specific patient and offer physicians individualized treatment approaches and recommendations
  - Process MRI scans in bulk to mine visual and technical data barely noticed by human physicians and help them quickly detect even the slightest signs of disease
  - Research DNA data on tumor segmentation and sequencing, and execute other DNA-related medical examinations and/or scientific studies
  - Suggest personalized patient insurance plans and custom health policies based on health risk

Data Mining in Healthcare: Examples

- Pharmacy and hospital management insights
- Hospital resources and pharmacy stock-management applications can be enhanced with the help of data mining
  - Identify seasonal spikes or declines in patient symptoms and drug prescriptions
  - Dig into a pharmacy's CRM system or Hospital Information Management System (HIMS) to classify, cluster, visualize, and analyze current customer data
  - Predict future demand for beds, medication stock, workforce, and various resources found in hospitals, pharmacies, and other medical institutions
  - Use accurate insights to manage pharmacy stock and hospital beds to be reserved prior to seasonal disease outbreaks
  - Correlate seasonal epidemics and environmental changes with the way in which risk is spread among different patient groups
  - Integrate data-mining tools into hospital management apps to provide clinical staff with access to medical evidence-supported insights

Data Mining in Healthcare: Examples

- Dietary pattern exploration
- Studies suggest that certain products can lead to the development of chronic disease and even cancer
  - Dig into database to find the right patients and form focus groups of patients to be supervised under dietary pattern research in medical organization
  - Control their meals and collect data
  - Once a sufficient volume of data is accumulated, it can be mined by data analysts to discover relationships and patterns
  - Employ self-reporting apps for patients
  - People can report their meals and access nutrition plans or suggestions from the clinic, which can follow them and provide notifications

Machine learning in healthcare

- To draw insights from large medical data sets to enhance clinicians' decision-making, improve patient outcomes, automate healthcare professionals’ daily workflows, accelerate medical research
- Opportunities for Machine Learning in Healthcare
  1. Automating clinical tasks
  2. Providing clinical support
  3. Expanding clinical capacities

Benefits of Machine Learning in Healthcare

- Identifying and Diagnosing Critical Diseases
- Faster Data Collection
- New drug discovery and development
- Communications
- Reading and analyzing medical documents and data
- Remote assistance in treating patients
- Clinical Trials

Ethics of Employing Machine Learning in Healthcare

- Informed Consent to Use
  - kind of data collected and the possible limitations of using AI
- Safety and Transparency
  - need to ensure these systems' safety and reliability and be transparent about them to minimize harm
- Algorithmic Fairness and Biases
  - AI system is only as reliable and effective as it is trained, hence AI makers should address this risk and minimize biases to ensure the effectiveness of healthcare solutions
- Data Privacy
  - must have adequate information about the collection and processing of their data to adhere to the fundamental privacy rights of human beings

Machine Leaning Applications in Healthcare

- Diabetes prediction
  - Diabetes mostly damages the kidneys, the heart, and nerves
  - ML detects very early, saving lives
  - Classification algorithms like KNN, Decision Tree, and Naive Bayes could be a basis to build a system that predicts diabetes
  - Naive Bayes is the most efficient among them in terms of performance and computation time

- Liver disease prediction
  - The liver plays a leading function in metabolism
  - It is vulnerable to diseases like chronic hepatitis, liver cancer, and cirrhosis.
  - ML algorithms : classification and clustering on The Liver Disorders Dataset or the Indian Liver Patient Dataset (ILPD) could be used for this task

Machine Leaning Applications in Healthcare

- Personalised Drug
  - Microsoft is using AI-based technology in its Project Hanover, which aims to find personalized drug combinations to cure Acute Myeloid Leukemia
  - A type of cancer of the blood and bone marrow with excess immature white blood cells

- Diagnoses via image analysis
  - Microsoft is revolutionizing healthcare data analysis with its InnerEye project
  - uses Computer Vision to process medical images to make a diagnosis
  - The **InnerEye**-DeepLearning toolkit enables researchers to use Azure Machine Learning to train and evaluate models in hours, that would otherwise take days

Machine Leaning Applications in Healthcare

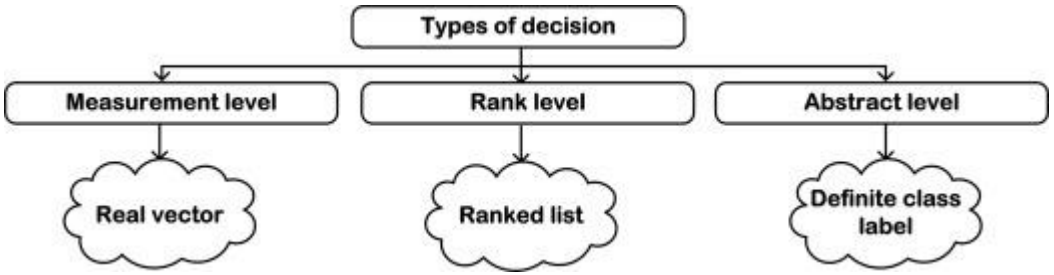
- Personalizing treatment
  - IBM Watson Oncology(software as a service, is a cognitive computing decision support system) is a distinctive leader in this area by providing numerous treatment plans that first analyze a patient's medical history
- Adjusting behavior
  - Provides tips on your daily activities to prevent diseases
  - An application from Somatix, a B2B2C-based company
  - This application keeps track of the unconscious activities we do every day and alerts us to those that might be dangerous from the long-term perspective

Machine Leaning Applications in Healthcare

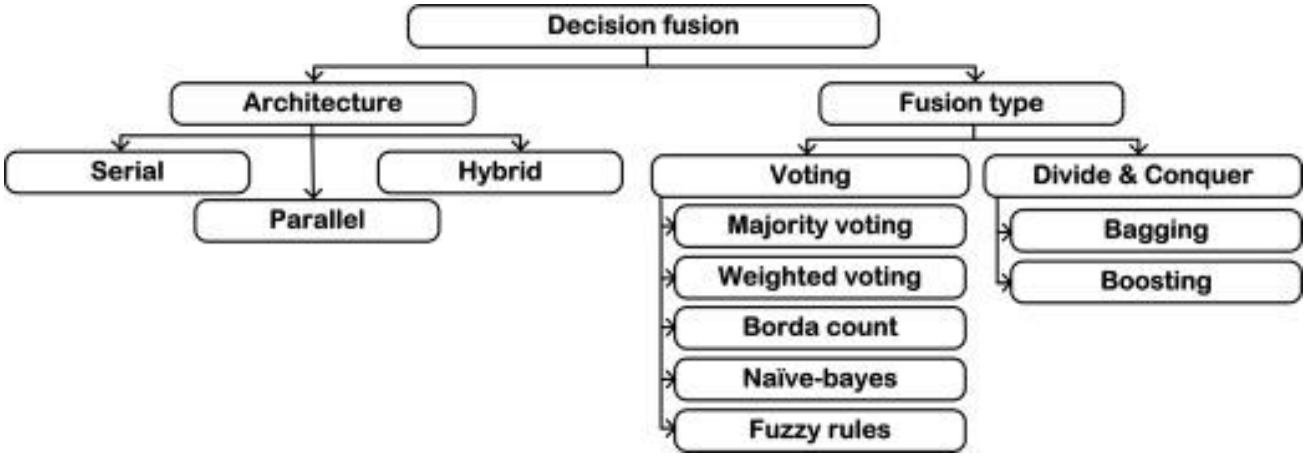
- Medical research and clinical trial improvement
  - The technology reduce errors and could suggest the best sample sizes to be tested
- Epidemic control
  - experts have access to information from satellites, social media trends, news websites, and video streams
  - Neural networks could process all and make conclusions on epidemic outbreaks all over the world
  - Dangerous diseases could be nipped in the bud before they could cause massive damage
- Leveraging crowdsourced medical data
- Artificial Intelligence Surgery
  - Automatic suturing/Surgical workflow modeling
  - Improvement of robotic surgical materials/ Surgical skill evaluation
  - Robotic arm fetches the instruments during surgery lowers complications by 50% and about decreases the time the patient stays in the operating room by about 20%.

Multi classifier Decision Fusion

- Combine the decisions taken by different classifiers to common consensus that is better than individual decisions to enhance the performance of the classification task
- The decisions taken by a classifier are of three types:
- measurement level: this type of decision involves the classifier returning a real valued vector
- rank level: involves the classifiers to return an ordered sequence of classes
- abstract level: this is the most widely applied type of decision where the classifiers return a single class label as the decision

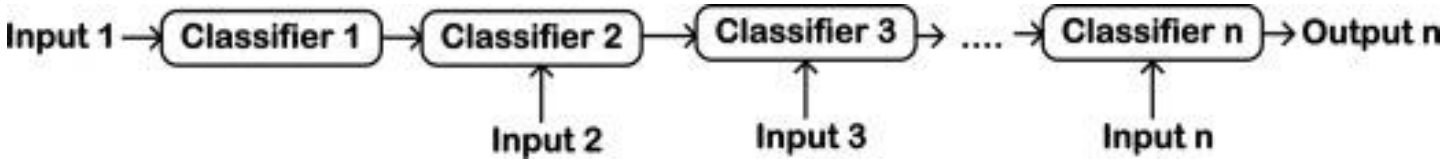


Decision Fusion Techniques



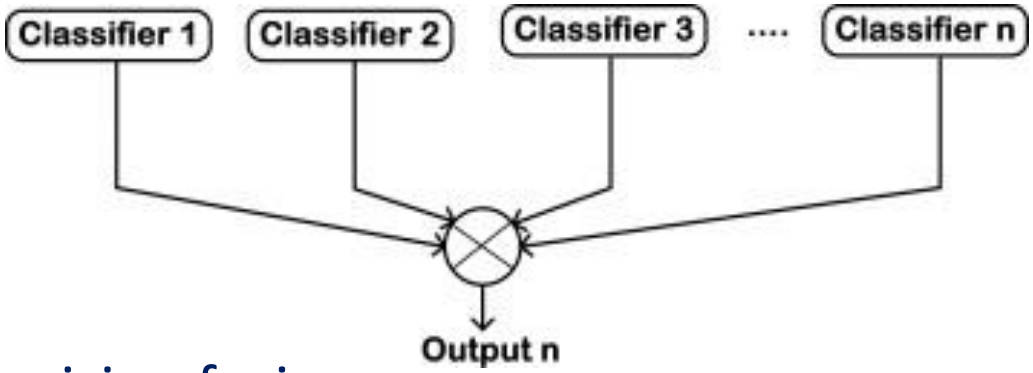
on the basis of the fusion architecture used;

- Serial decision fusion

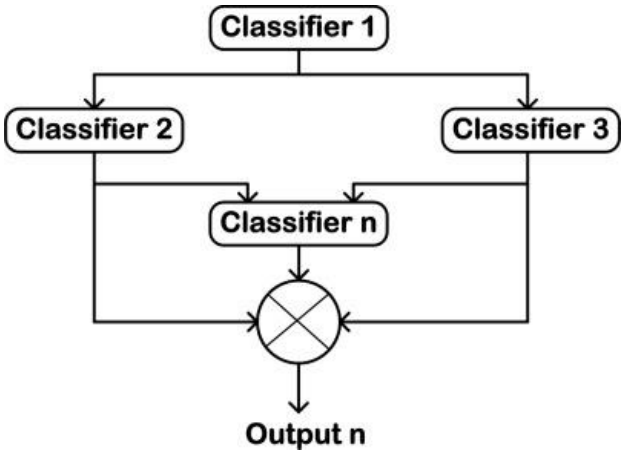


Decision Fusion Techniques

- Parallel decision fusion

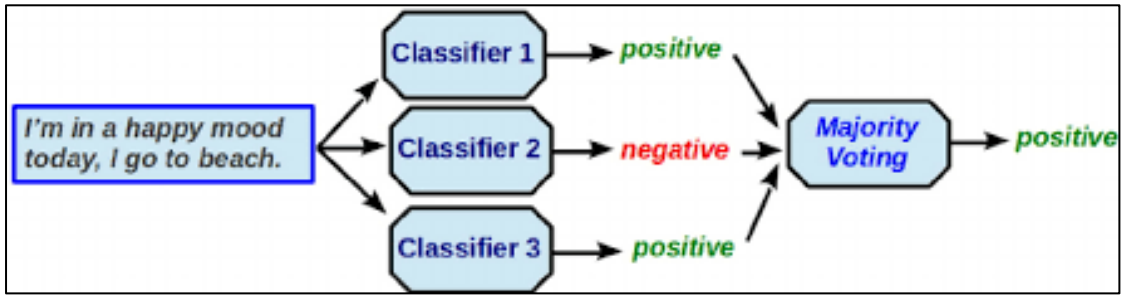


- Hybrid decision fusion:



Voting Techniques

- Majority Voting(hard voting)

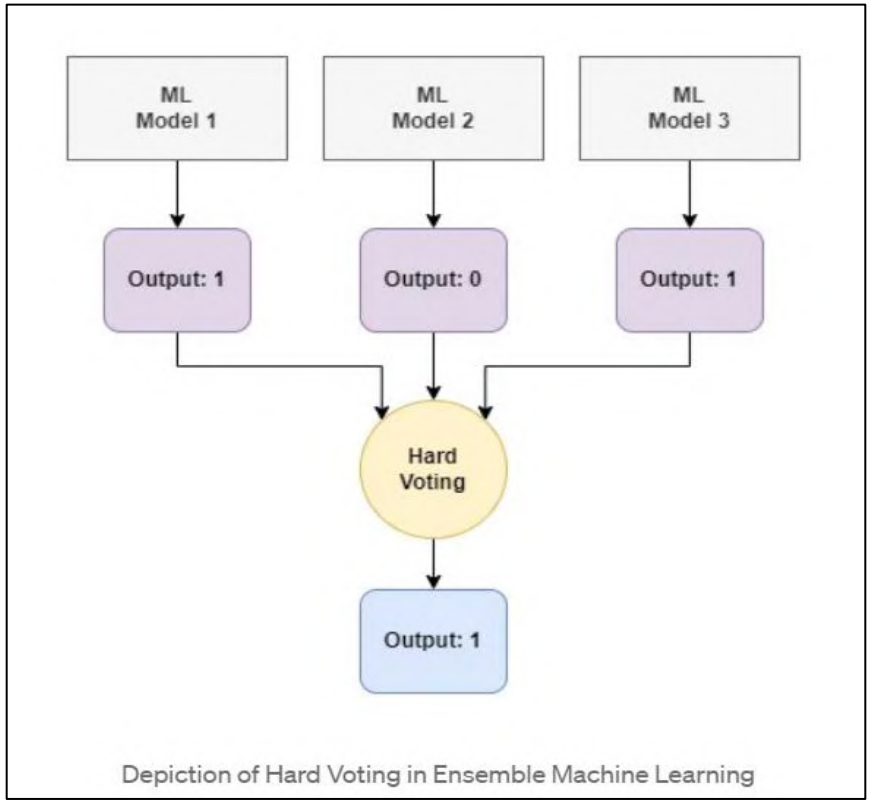


every individual classifier votes for a class, and the majority wins

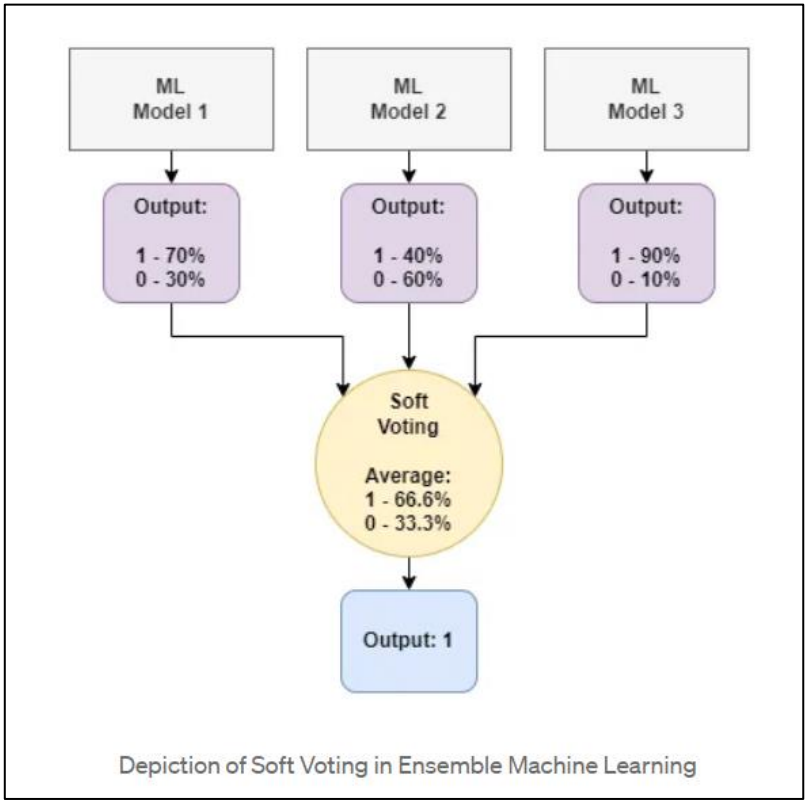
- Weighted voting (soft voting)
- Each classifier assigns a probability to each class, and the ensemble's prediction is the class with the highest total probability

- Borda count
- Naïve Bayes
- Fuzzy rule

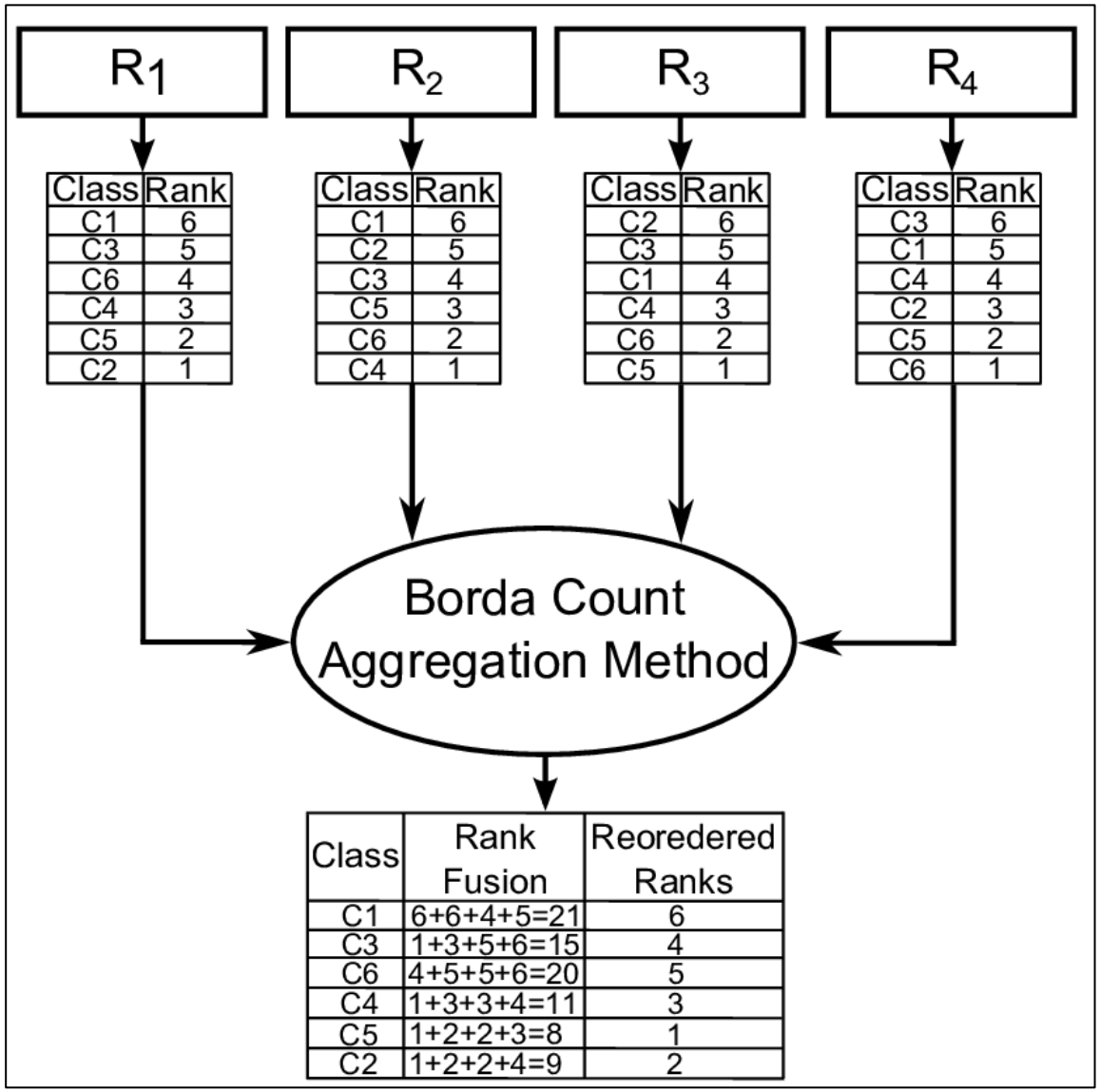
Voting Techniques



Depiction of Hard Voting in Ensemble Machine Learning



Depiction of Soft Voting in Ensemble Machine Learning


R1
R2
R3
R4
Class Rank
C1 6
C3 5
C6 4
C4 3
C5 2
C2 1
Class Rank
C1 6
C2 5
C3 4
C5 3
C6 2
C4 1
Class Rank
C2 6
C3 5
C1 4
C4 3
C6 2
C5 1
Class Rank
C3 6
C1 5
C4 4
C2 3
C5 2
C6 1
Borda Count Aggregation Method
Class | Rank Fusion | Reoredered Ranks
C1 | 6+6+4+5=21 | 6
C3 | 1+3+5+6=15 | 4
C6 | 4+5+5+6=20 | 5
C4 | 1+3+3+4=11 | 3
C5 | 1+2+2+3=8 | 1
C2 | 1+2+2+4=9 | 2

Decision Fusion Techniques

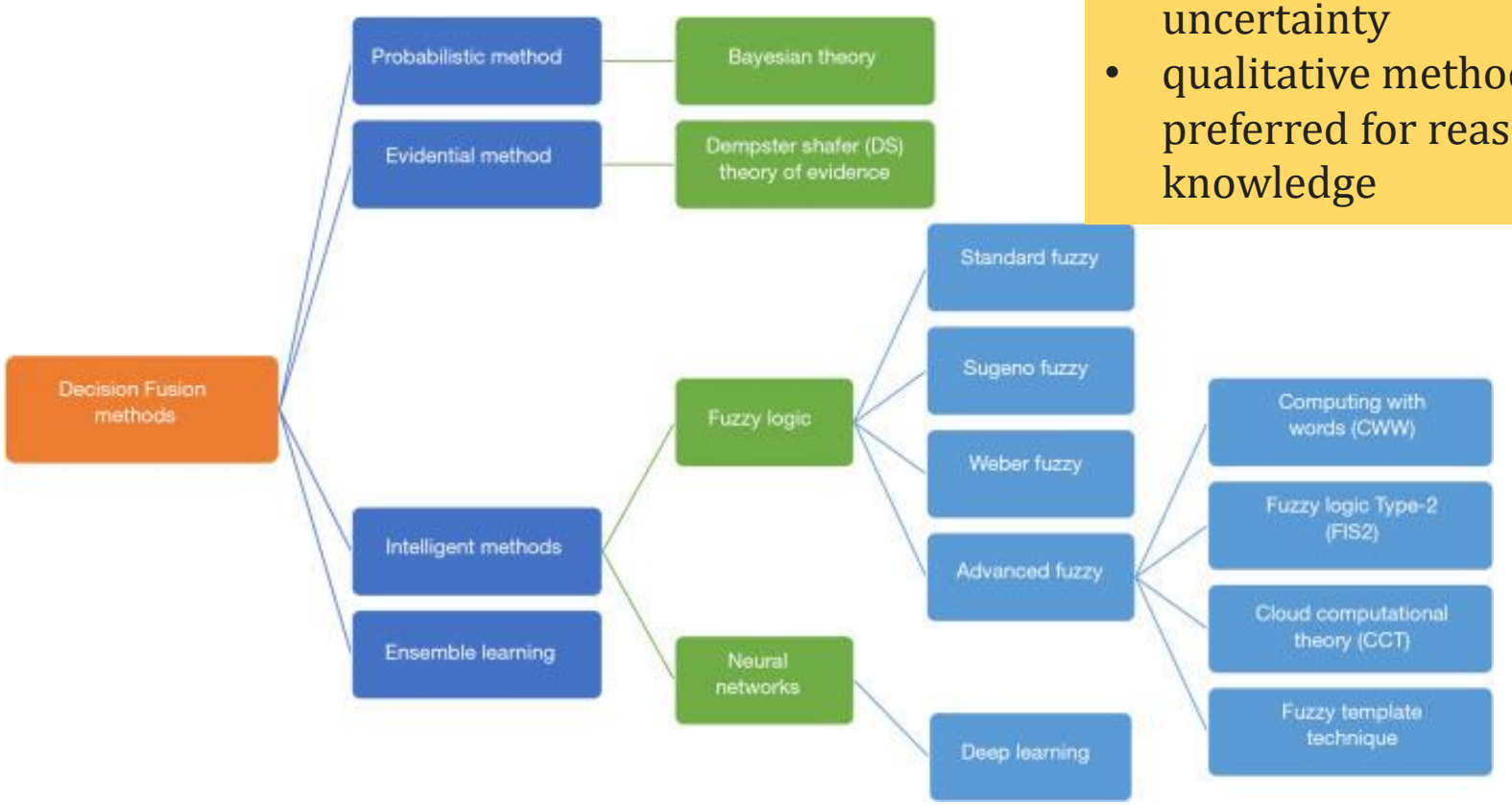
- On the basis of the fusion type, these techniques are of two types:
- **Voting-based:** In the voting-based decision fusion techniques, majority voting is the most popular and is widely used
- Some of the other techniques include weighted voting in which a weight to each classifier is attached and then decision fusion is performed
- Other voting techniques are probability-based, such as fuzzy rules, Naïve-Bayes, Dempster-Shafer theory etc.

Decision Fusion Techniques

- **Divide and conquer:**
- In this decision fusion technique, the dataset is divided into subsets of equal sizes, and then the classification is performed followed by decision fusion on the results of those smaller dataset classifications
- These divide and conquer methods include the concepts of bagging and boosting

Types of methods used in the decision fusion technique

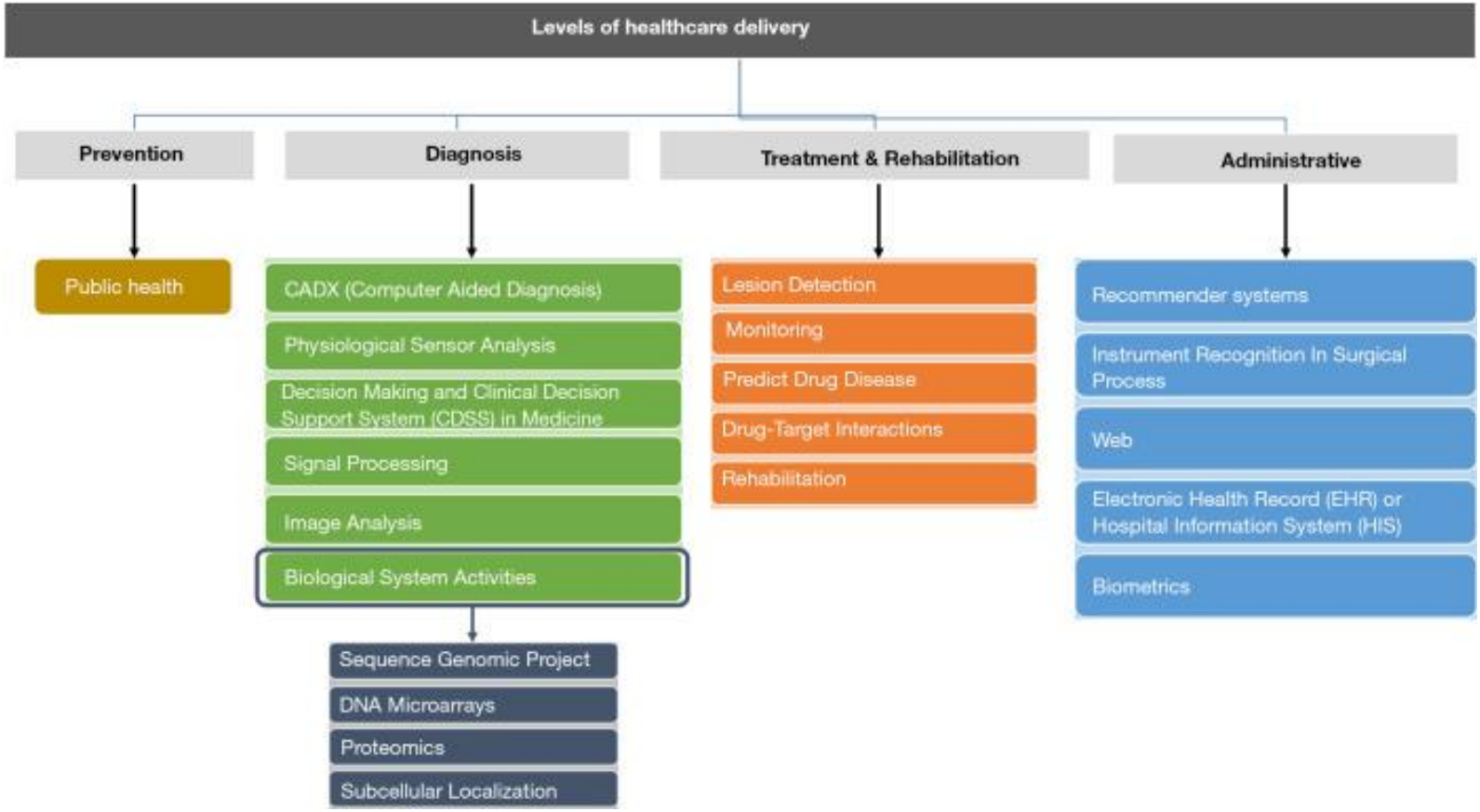
- Quantitative methods: evidence theory, probability theory, fuzzy sets, random sets and possibility theory.
- are the best means for displaying and managing uncertainty
- qualitative methods are preferred for reasoning and knowledge



Multimodal Fusion in Healthcare

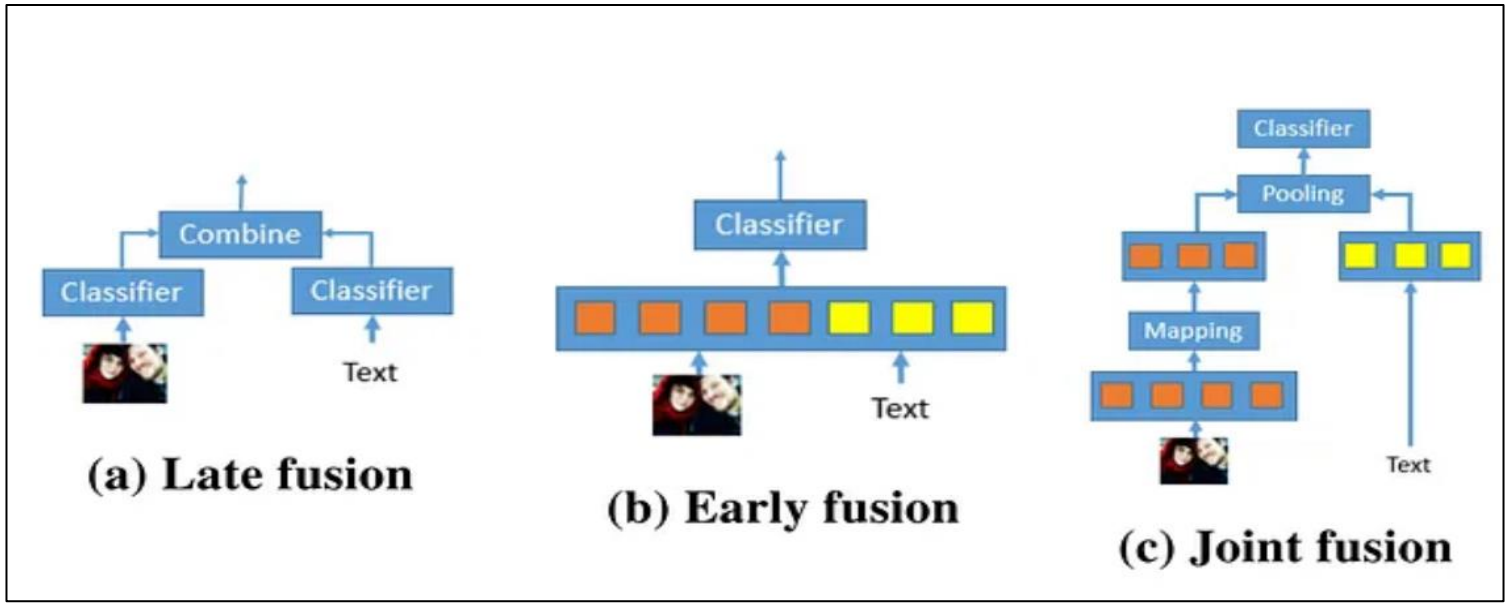
- Examples
  - Self driving cars
  - Healthcare
    - Alzheimer
    - Breast cancer
    - Skin cancer
- Demographics
- Medical history
- Lab values
- Imaging data

Applications of decision fusion technique according to levels of healthcare delivery

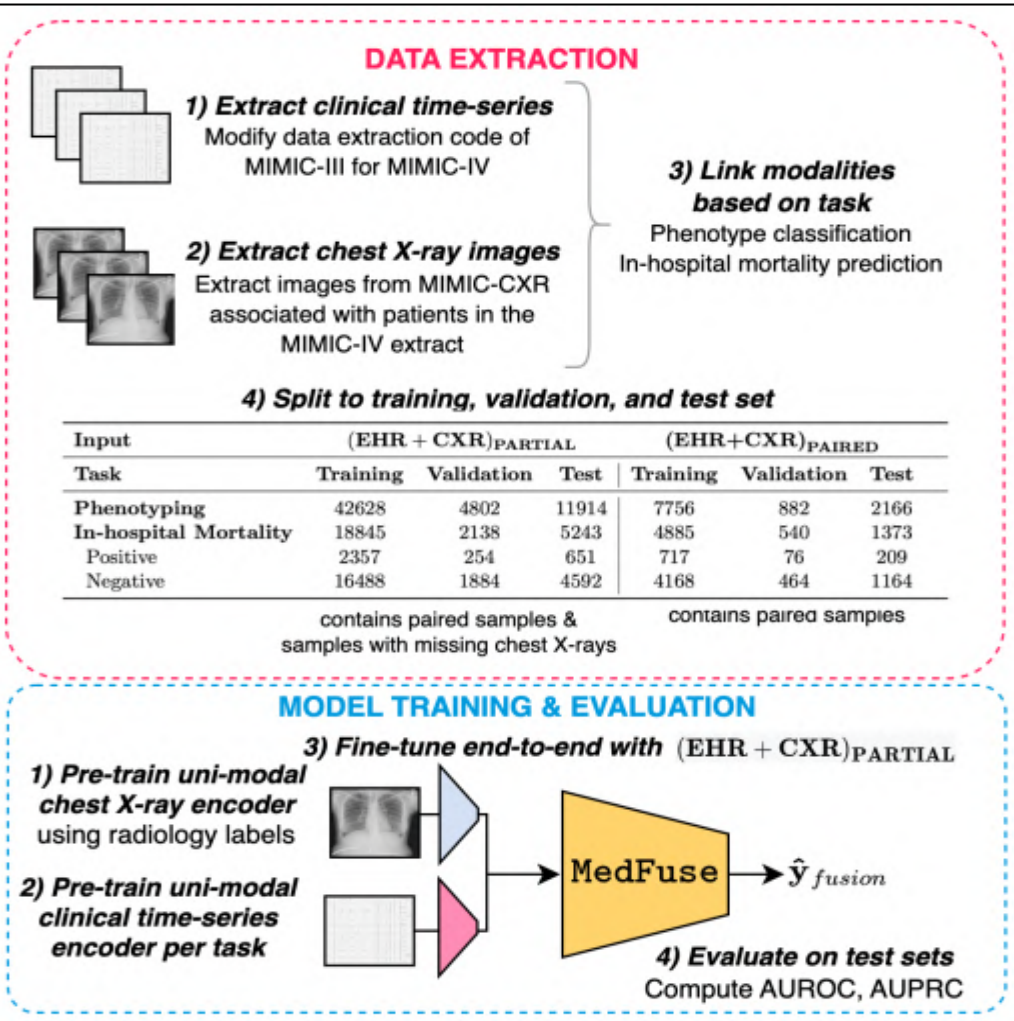


Fusion Techniques

- Early Fusion
  - Feature level fusion
- Joint Fusion
  - During learning
- Late Fusion
  - Averaging out the predictions



- From research paper



1: **Overview of the proposed work.** We first extract and link the datasets from MIMIC-IV and MIMIC-CXR based on the task definition (i.e., inhospital mortality prediction, or phenotype classification). The data splits of the training, validation, and test sets are summarized for each task, and the prevalence of positive and negative labels for in-hospital mortality is shown. Phenotype classification involves 25 labels as shown in Table 4.