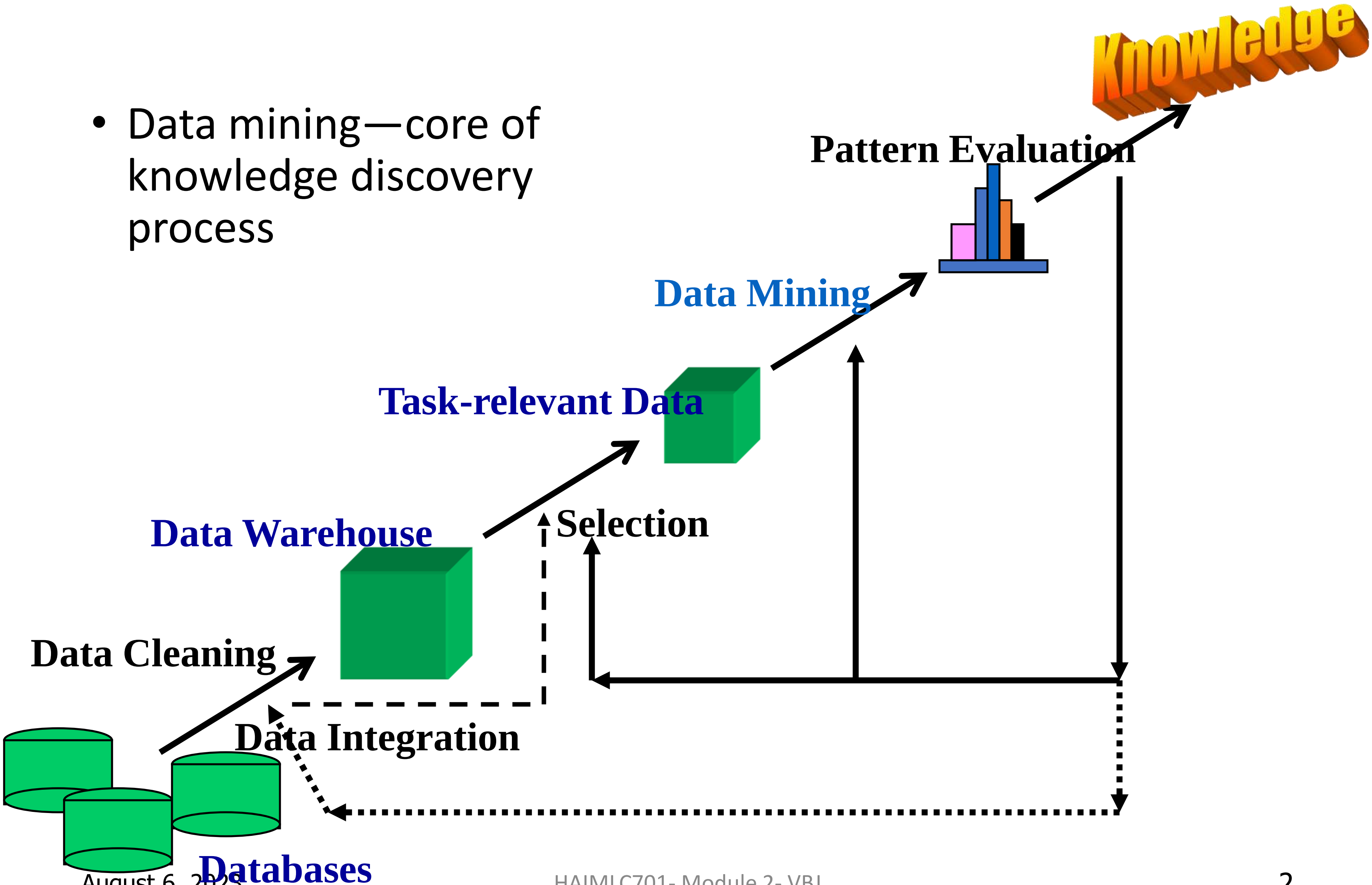
HAIMLC701 AI & ML in Healthcare

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.	

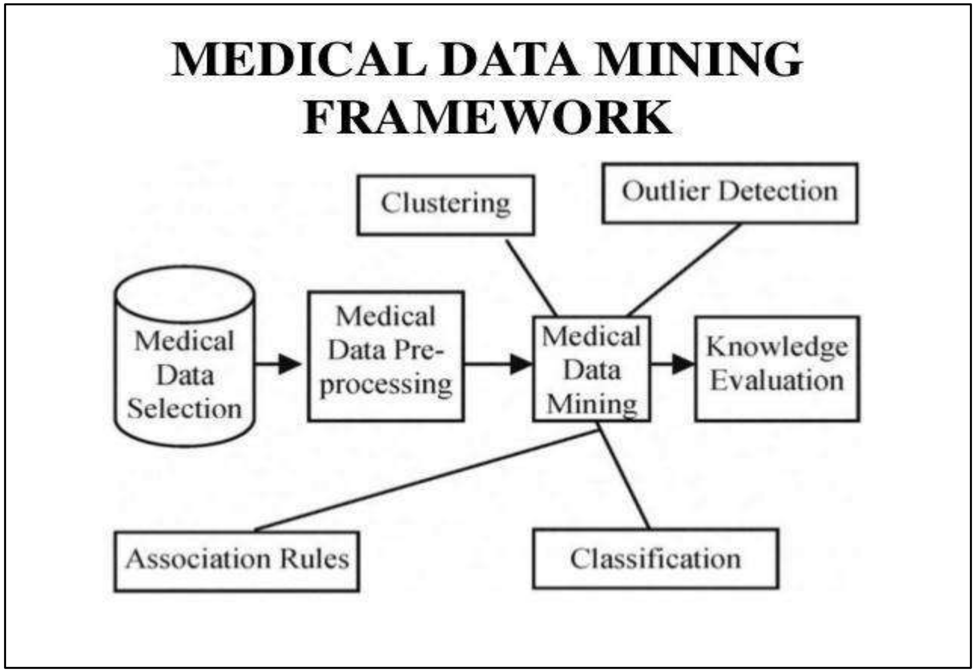
Knowledge Discovery (KDD) Process

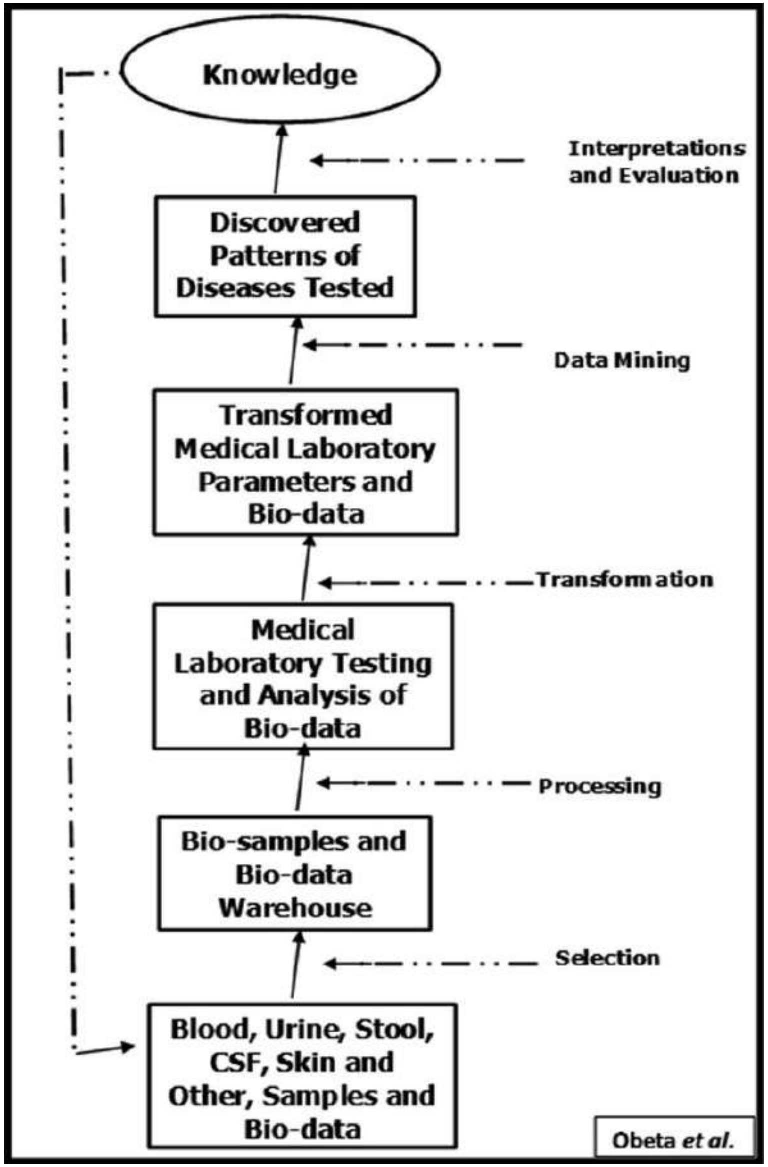
- Data mining—core of knowledge discovery process



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?



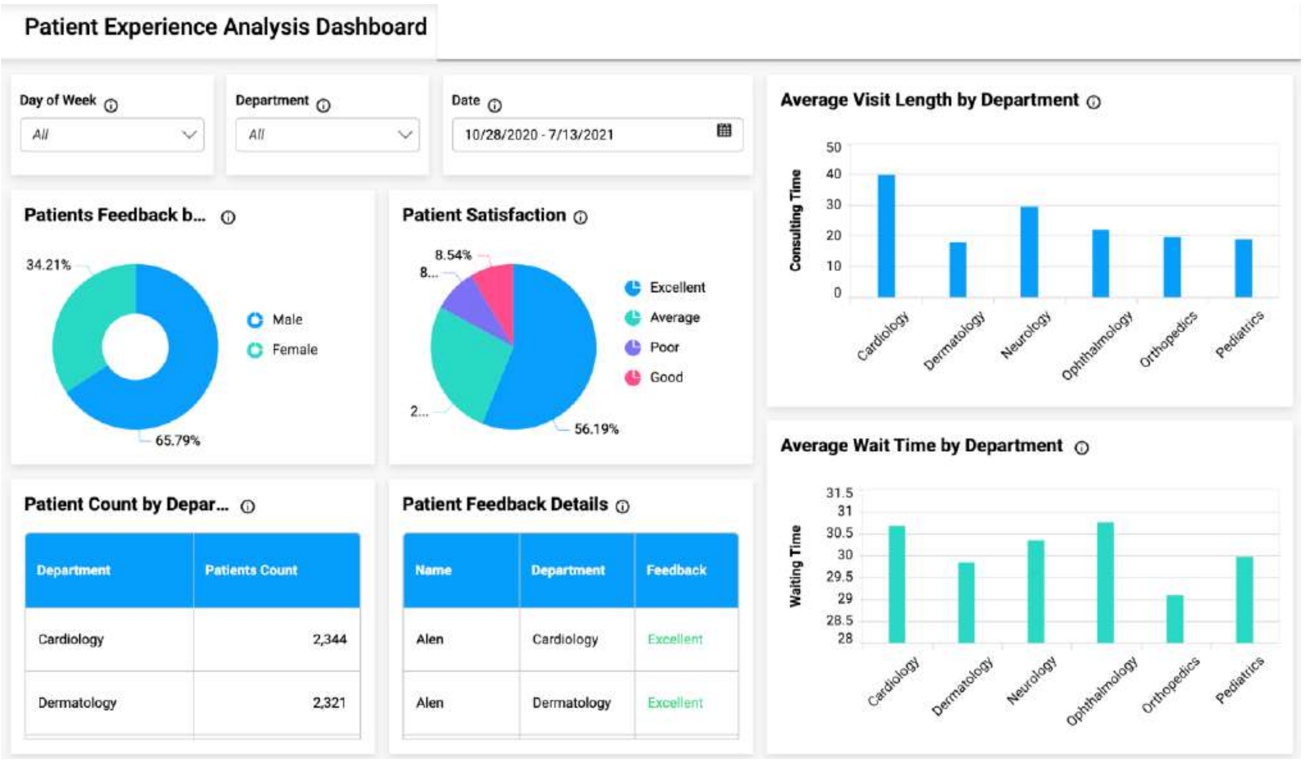


Data Mining Techniques in Healthcare

- 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 eventsall 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



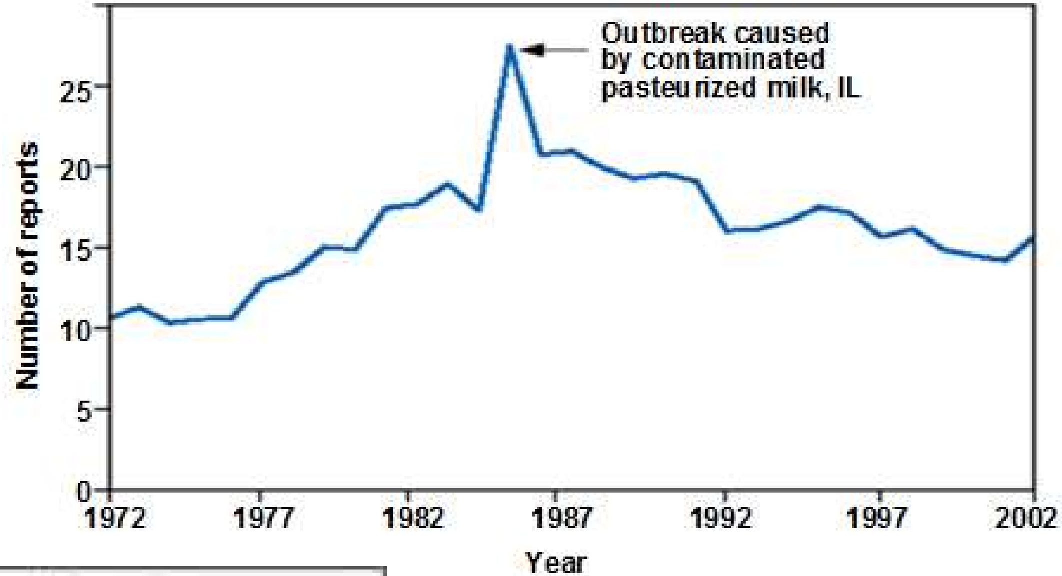
Challenges and Issues

- 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

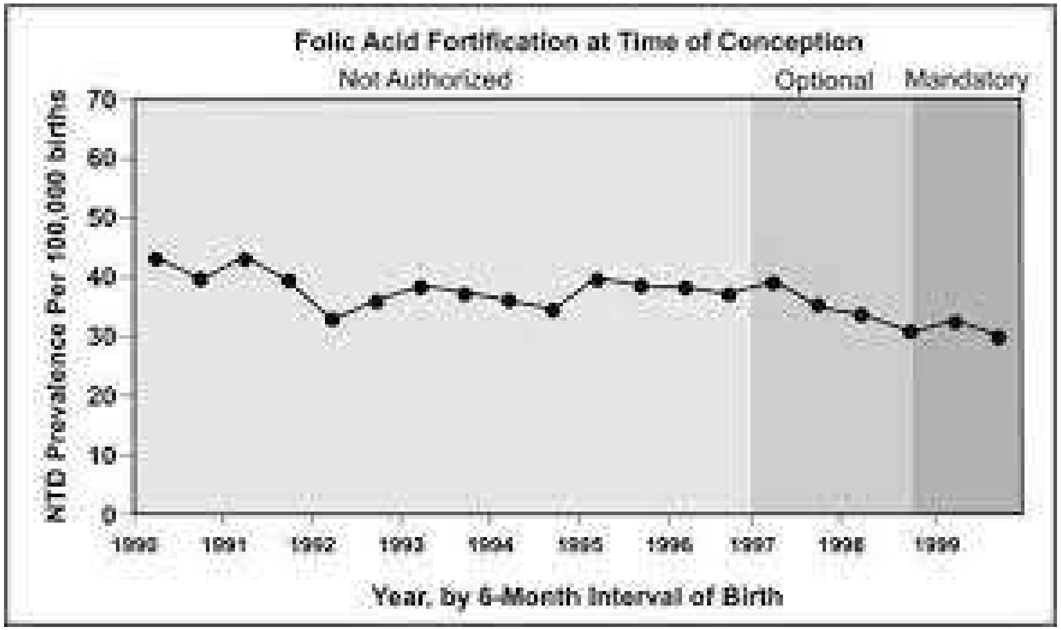
Data Mining in Healthcare: Examples

- 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

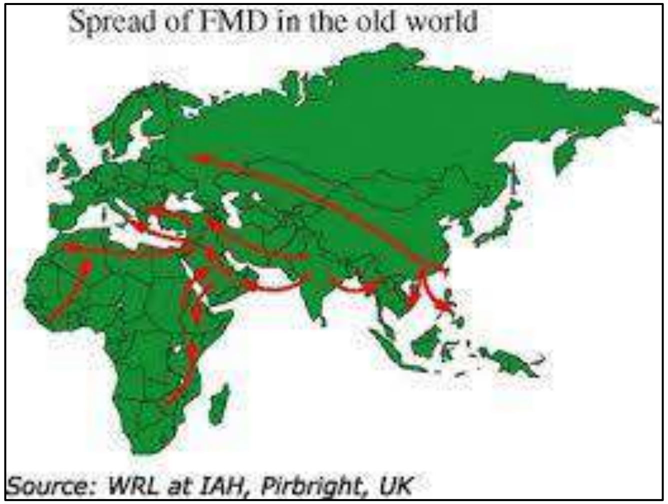
Epidemiology is the study and analysis of the distribution, patterns and determinants of health and disease conditions in a defined population.



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
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 and helpful approaches in the context of 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 in the healthcare database while identifying interesting phenomena and developing 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
  - Suggest new approaches and tweak existing healthcare plans and medical protocols to increase treatment efficiency and bring better outcomes

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, which offers the following capabilities…
  - 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 Therefore AI makers should address this risk and minimize biases at every stage to ensure it doesn't adversely impact 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

- The prediction of diabetes
  - 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

- Predicting liver disease
  - The liver plays a leading function in metabolism
  - It is vulnerable to diseases like chronic hepatitis, liver cancer, and cirrhosis.
  - It is a very hard task to effectively predict liver disease using enormous amounts of medical data; however, there have already been some significant shifts in this area
  - ML algorithms : classification and clustering are making the difference here. The Liver Disorders Dataset or the Indian Liver Patient Dataset (ILPD) could be used for this task

Machine Leaning Applications in Healthcare

- Finding the best cure
  - 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

- Making diagnoses via image analysis
  - Microsoft is revolutionizing healthcare data analysis with its InnerEye project.
  - This startup 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
  - Machine Learning in Medicine is making great progress.
  - 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
  - This is a very interesting area to observe
  - Giving tips on your daily activities to prevent cancer?
  - That's exactly what an application from Somatix, a B2B2C-based company, is doing
  - 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

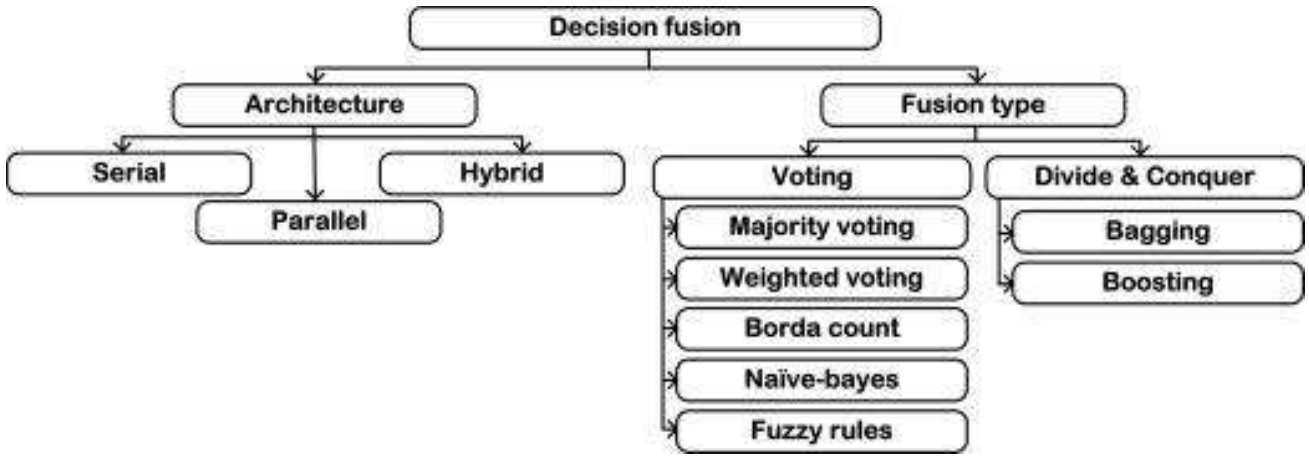
Machine Leaning Applications in Healthcare

- Medical research and clinical trial improvement
  - The technology will also lower the number of data-based 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 of that and make conclusions on epidemic outbreaks all over the world.
  - Dangerous diseases could be nipped in the bud before they could actually cause massive damage.
- Leveraging crowdsourced medical data
- Artificial Intelligence Surgery
  - Automatic suturing/Surgical workflow modeling
  - Improvement of robotic surgical materials/ Surgical skill evaluation
  - When a medical procedure is conducted, the robot will fetch instruments for the doctor with its robotic hands. This kind of practice lowers surgical complications by 50% and about decreases the time the patient stays in the operating room by about 20%.

Multi classifier Decision Fusion

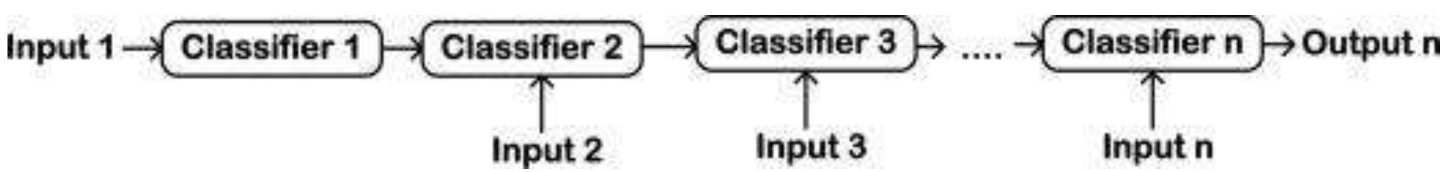
- Combine the decisions taken by different classifiers to common consensus that is better than individual decisions to enhance the performance of the classification task
- The decisions taken by a classifier are of three types:
- measurement level: this type of decision involves the classifier returning 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



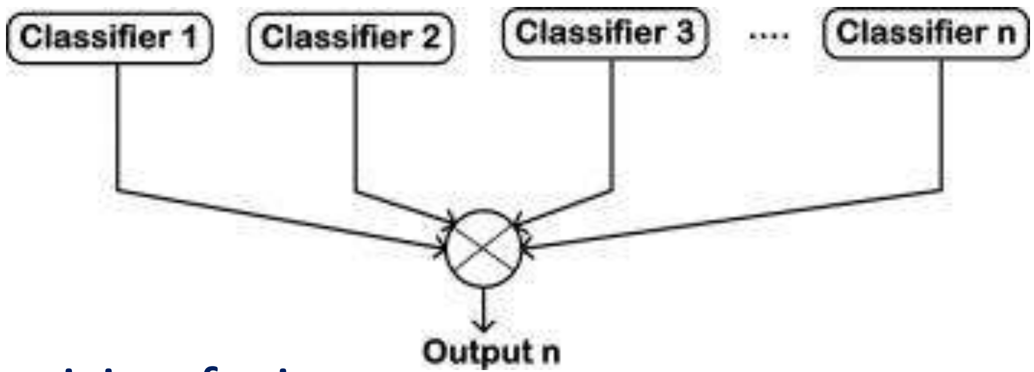
on the basis of the fusion architecture used;

- Serial decision fusion

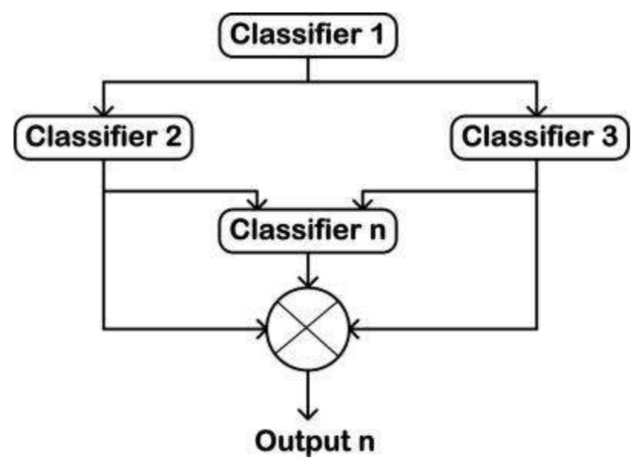


Decision Fusion Techniques

- Parallel decision fusion



- Hybrid decision fusion:



Decision Fusion Techniques

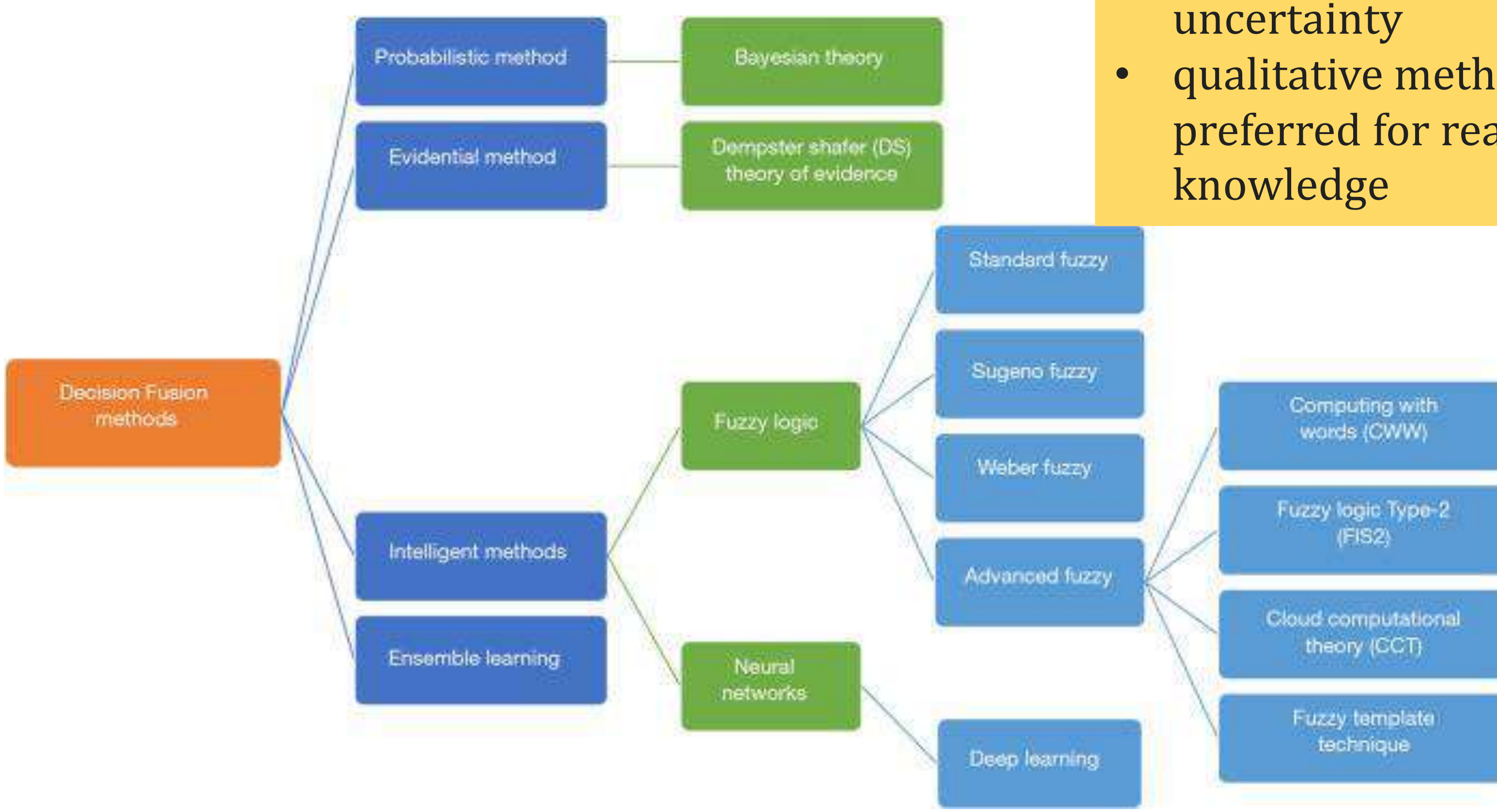
- On the basis of the fusion type, these techniques are of two types:
- **Voting-based:** In the voting-based decision fusion techniques, majority voting is the most popular and is widely used
- Some of the other techniques include weighted voting in which a weight to each classifier is attached and 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

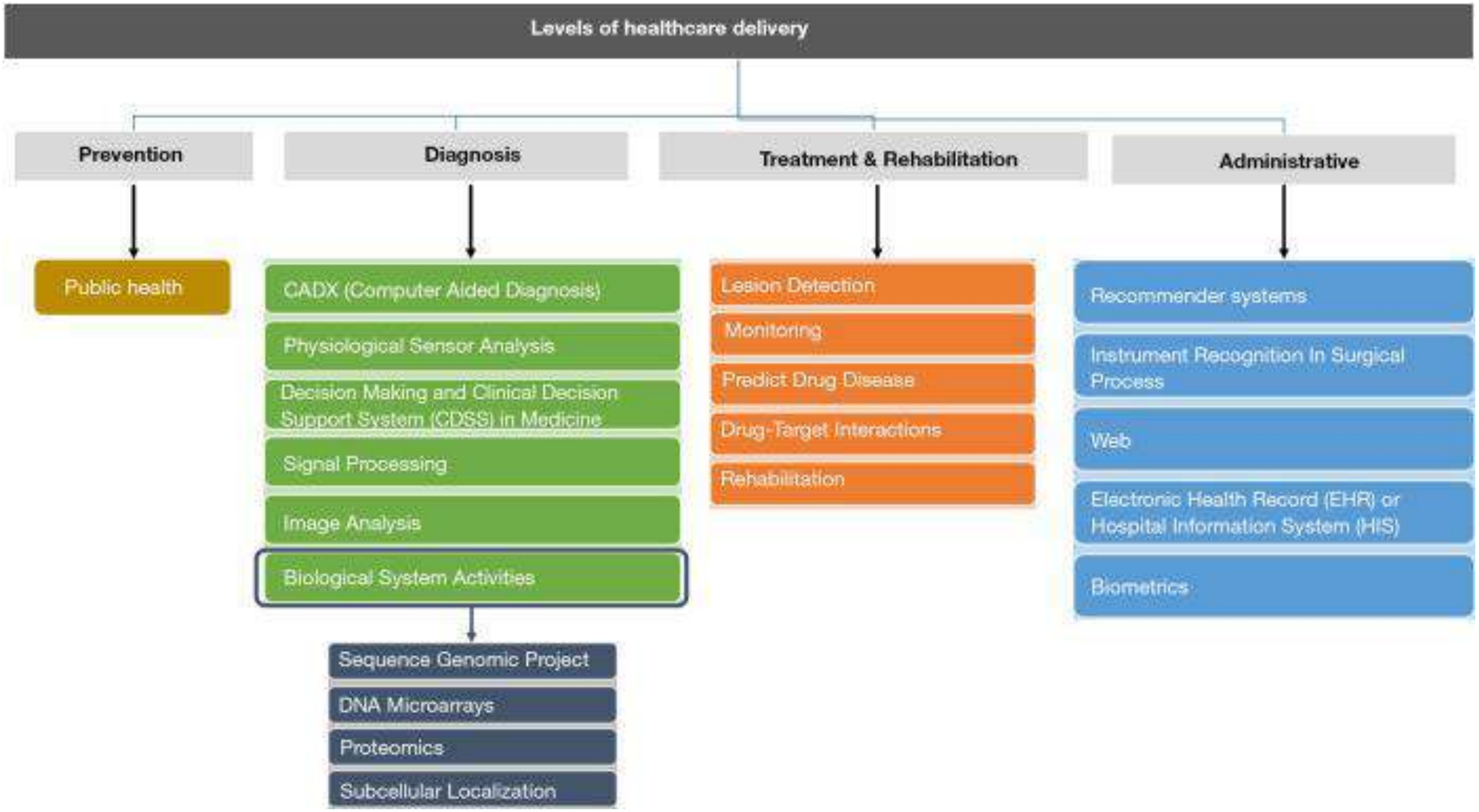
- Quantitative methods: evidence theory, probability theory, fuzzy sets, random sets and possibility theory.
- are the best means for displaying and managing uncertainty
- qualitative methods are preferred for reasoning and knowledge



Multimodal Fusion in Healthcare

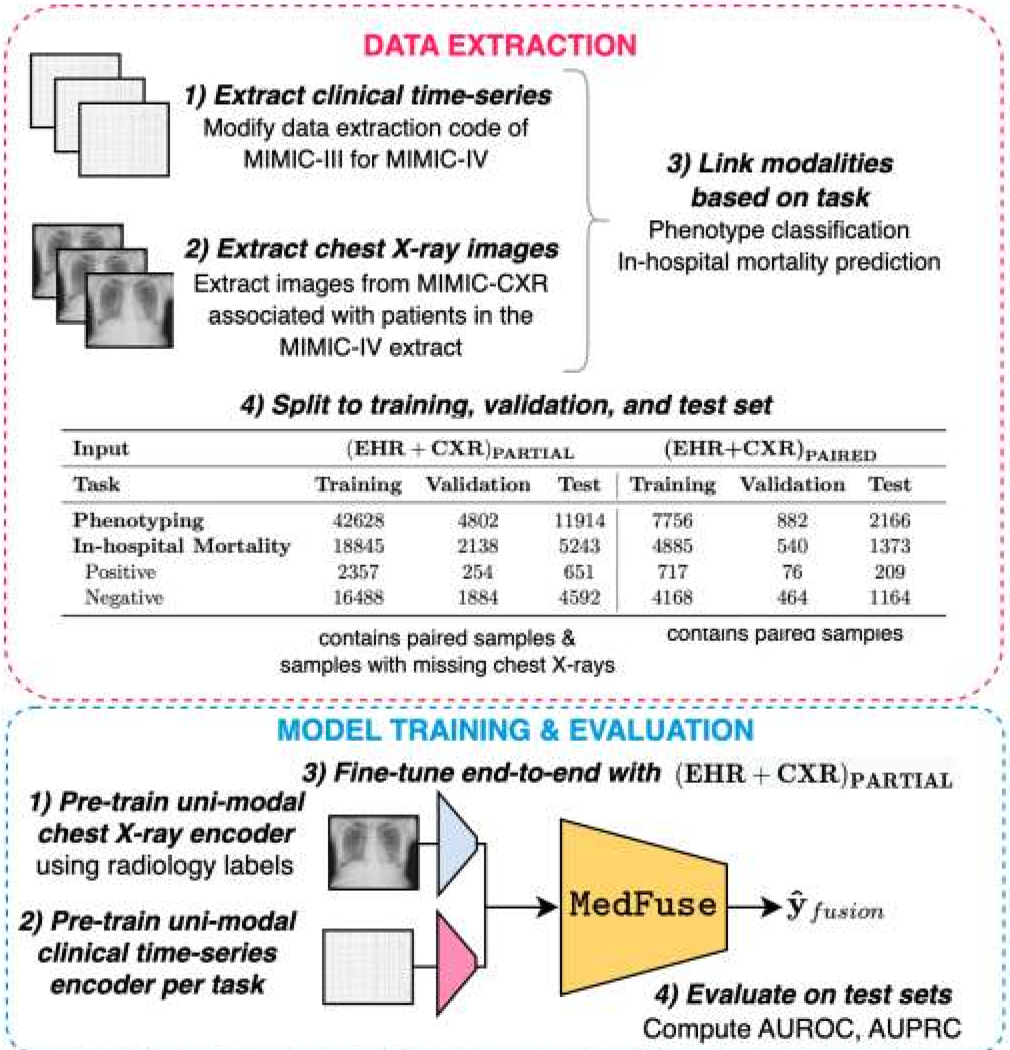
- Examples
  - Self driving cars
  - Healthcare
    - Alzheimer
    - Breast cancer
    - Skin cancer
- Demographics
- Medical history
- Lab values
- Imaging data

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



Input	$(\text{EHR}+\text{CXR})_{\text{PARTIAL}}$			$(\text{EHR}+\text{CXR})_{\text{PAIRED}}$		
Task	Training	Validation	Test	Training	Validation	Test
Phenotyping	42628	4802	11914	7756	882	2166
In-hospital Mortality	18845	2138	5243	4885	540	1373
Positive	2357	254	651	717	76	209
Negative	16488	1884	4592	4168	464	1164

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.

Ensemble Learning

- A composite model, combines a series of low performing classifiers with the aim of creating an improved classifier
- Here, individual classifier vote and final prediction label returned that performs majority voting
- Offer more accuracy than individual or base classifier
- Can be parallelized by allocating each base learner to different-different machines
- Decrease variance using bagging
  - bias using a boosting and improve predictions using stacking

Ensemble Learning

- Use multiple learning algorithms (classifiers)
- Combine the decisions
- Can be more accurate than the individual classifiers
- Generate a group of base-learners
- Different learners use different
  – Algorithms
  – Hyperparameters
  – Representations (Modalities)
  – Training sets

Advantages

- Performance:
  - better predictions and achieve better performance than any single contributing model
- Robustness:
  - reduces the spread or dispersion of the predictions and model performance

Model Error

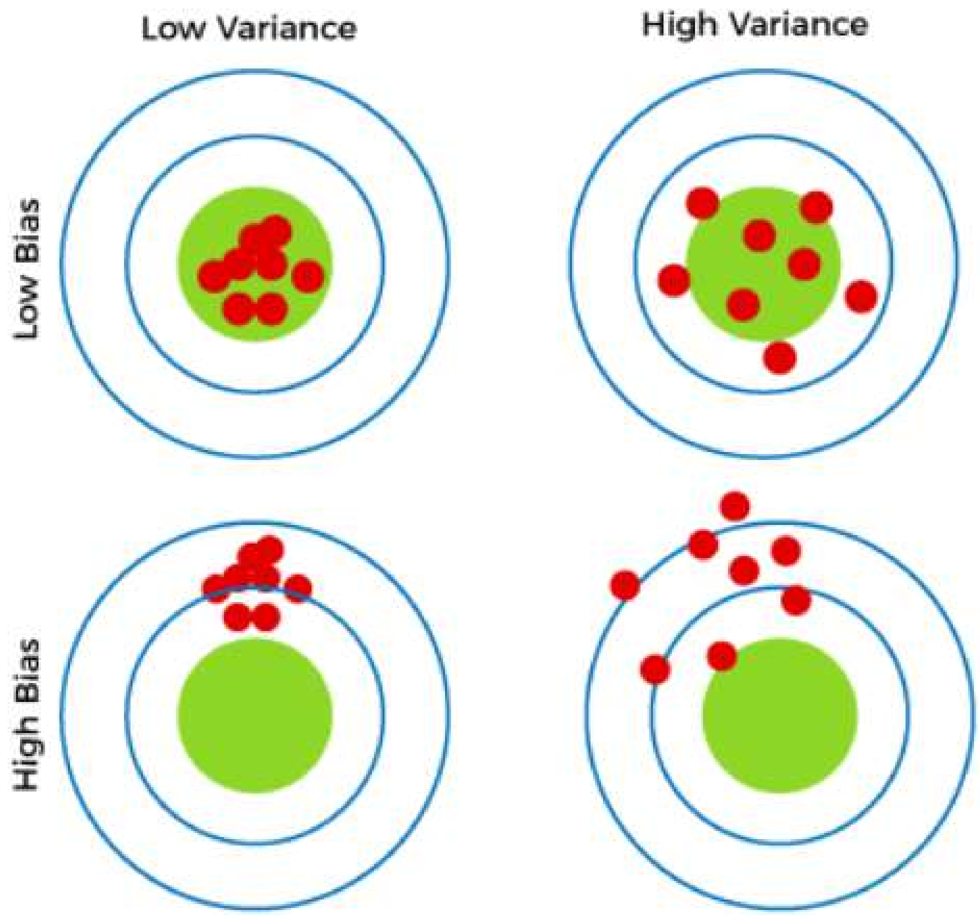
- The error emerging from any machine model

$$Err(x) = \left(E[\hat{f}(x)] - f(x)\right)^2 + E\left[\hat{f}(x) - E[\hat{f}(x)]\right]^2 + \sigma_e^2$$

$$Err(x) = \mathrm{Bias}^2 + \mathrm{Variance} + \mathrm{Irreducible\ Error}$$

- Bias:
  - Assumptions made by a model to make a function easier to learn/ error rate of the training data
  - When the error rate has a high value- High Bias(underfit)
- Variance:
  - The difference between the error rate of training data and testing data
  - If the difference is high- high variance(overfit)

Bias and Variance



Ensemble Learning

- Linear algorithm -high bias, as it makes them learn fast
    - **Linear Regression, Linear Discriminant Analysis and Logistic Regression.**
- Low bias **are Decision Trees, k-Nearest Neighbours and Support Vector Machines**

Ways to reduce High Bias:

- Increase the input features as the model is underfitted
- Decrease the regularization term
- Use more complex models, such as including some polynomial features(feature engineering)

- Ways to Reduce High Variance:

- Reduce the input features or number of parameters as a model is overfitted
- Do not use a much complex model
- Increase the training data
- Increase the Regularization term

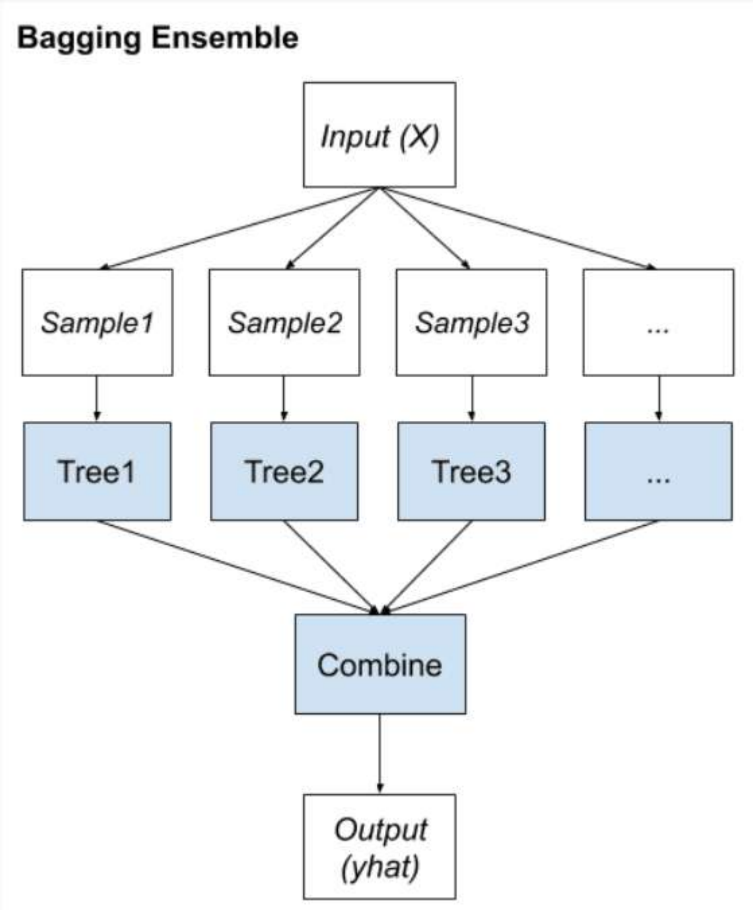
Regularization

- Form of regression, that regularizes or shrinks the coefficient estimates towards zero
- ***Discourages learning a more complex or flexible model, so as to avoid the risk of overfitting***
- The fitting procedure involves a loss function, known as residual sum of squares or RSS
- The coefficients are chosen, such that they minimize this loss function
- y= β0+β1x1+β2x2+β3x3+⋯+βnxn +b
- X1, X2, …Xn are the features for Y.
- β0,β1,…..βn are the weights or magnitude attached to the features, respectively. Here β0 represents the bias of the model, and b represents the intercept

$$\sum_{i=1}^{M}(y_i - y'_i)^2 = \sum_{i=1}^{M}\left(y_i - \sum_{j=0}^{n} \beta_j * Xij\right)^2$$

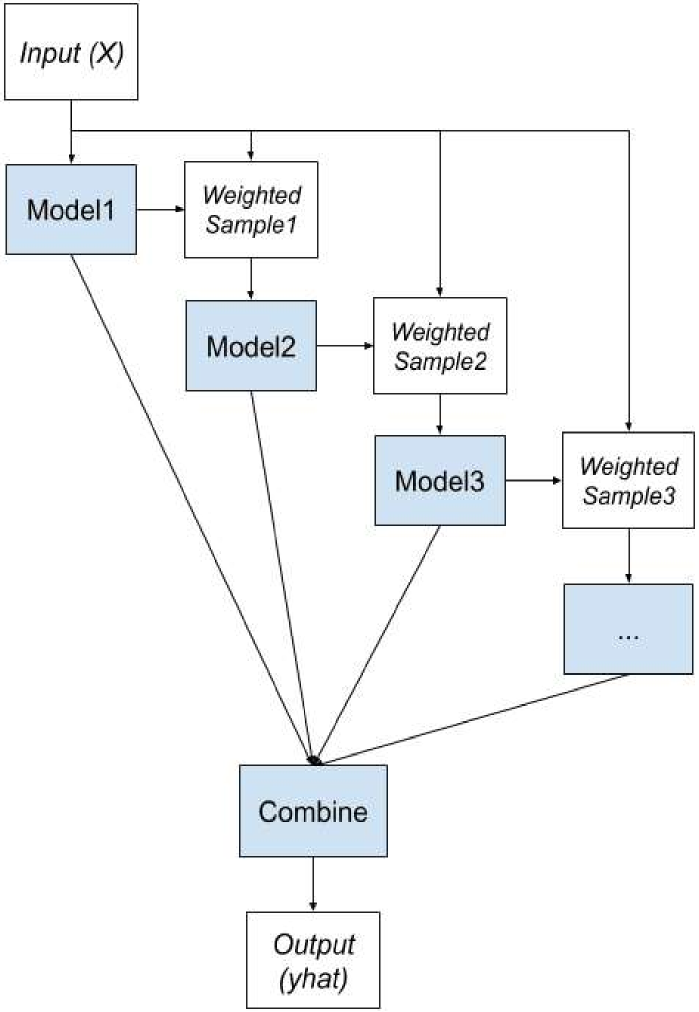
Ensemble Learning Methods

- Bagging
  - involves using a single machine learning algorithm, almost always an unpruned decision tree, and training each model on a different sample of the same training dataset
  - The predictions made by the ensemble members are then combined using simple statistics, such as voting or averaging
- Bootstrap samples of the training dataset
- Unpruned decision trees fit on each sample
- Simple voting or averaging of predictions
- Examples:
  - Bagged Decision Trees (canonical bagging)
  - Random Forest
  - Extra Trees



Ensemble Learning Methods

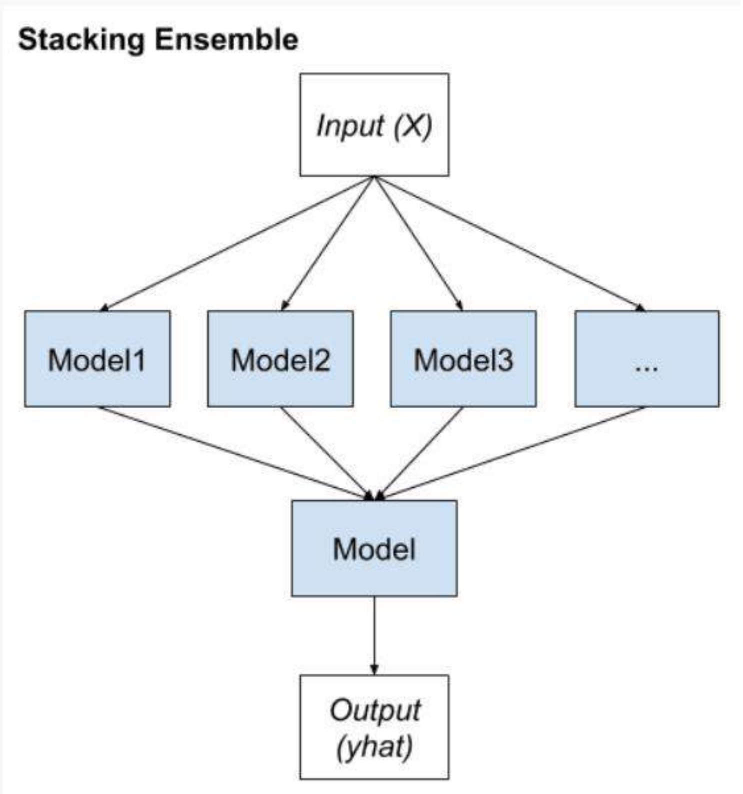
- Boosting
  - the training dataset for each subsequent classifier increasingly focuses on instances misclassified by previously generated classifiers
  - The models are fit and added to the ensemble sequentially such that the second model attempts to correct the predictions of the first model, the third corrects the second model, and so on.
  - develop a so-called “*strong-learner*” from many purpose-built “*weak-learners*.”
- Bias training data toward those examples that are hard to predict
- Iteratively add ensemble members to correct predictions of prior models
- Combine predictions using a weighted average of models



AdaBoost (canonical boosting)
Gradient Boosting Machines
Stochastic Gradient Boosting
(XGBoost and similar)

Ensemble Learning Methods

- Stacking
- a general procedure where a learner is trained to combine the individual learners
- Here, the individual learners are called the first-level learners, while the combiner is called the second-level learner, or meta-learner



- Unchanged training dataset
- Different machine learning algorithms for each ensemble member.
- Machine learning model to learn how to best combine predictions.

Meta Learning

- Subset of machine learning described as "learning to learn"
- Metalearning differs from base-learning in the scope of the level of adaptation
- Synonyms:
  - Selection of Algorithms, Ranking learning methods; Hyper-parameter Optimization; Adaptive learning; Self-adaptive systems; Dynamic selection of bias; Learning to learn
- Learning at the base-level is focused on accumulating experience on a specific task (e.g., credit rating, medical diagnosis, mine-rock discrimination, fraud detection, etc.)
- Learning at the metalevel is concerned with accumulating experience on the performance of multiple applications of a learning system

Meta Learning

- Focuses on the design and application of learning algorithms to acquire and use metaknowledge to assist machine learning users with the process of model selection
- It studies how to choose the right bias dynamically, and thus differs from base-level learning, where the bias is fixed or user-parameterized
- Important feature of self-adaptive systems- increase in efficiency through experience

Importance of Meta Learning

- Machine learning algorithms have some problems, such as
  - The need for large datasets for training
  - High operating costs due to many trials/experiments during the training phase
  - Experiments/trials take a long time to find the best model that performs best for a given data set

Structure of the Meta-learning System

- composed of two parts
  - Acquisition of metaknowledge from machine learning systems
  - Application of metaknowledge to new problems with the objective of identifying an optimal learning algorithm or technique
  - Example- if we are dealing with a classification task, metaknowledge can be used to select a suitable classifier for the new problem

Employing Metaknowledge to Select Machine Learning Algorithms

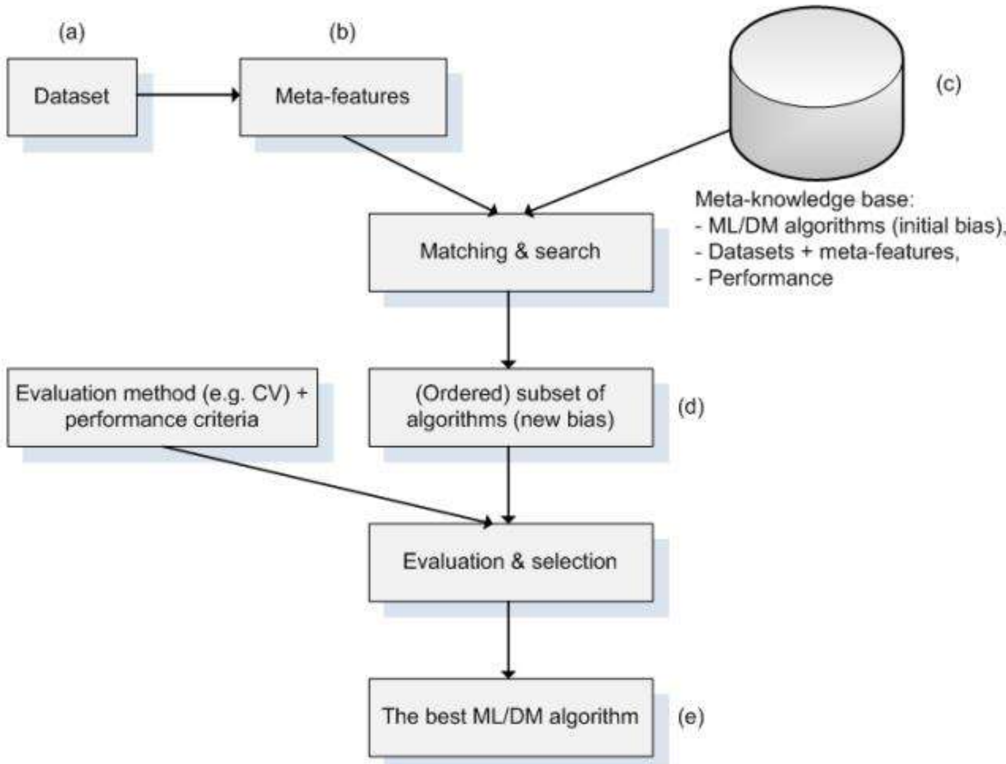


Fig. 1. Selection of machine learning algorithms: Determining the reduced space and selecting the best alternative.

Input/Output of Meta-learning System

- Depends on dataset characteristics or meta features that provide some information to differentiate performance among a given set of learning algorithms
- Measures:
  - number of classes, the number of features, the ratio of examples to features, the degree of correlation between features and target concept, the average class entropy, etc.
- disadvantage:
  - limit on how much information these meta-features can capture, given that all these measures are uni- or bi-lateral measures only

Acquisition of Metaknowledge

- Meta learning in the context of selecting algorithms for machine learning (there are a number of other areas, such as regression, time series forecasting, and optimization)
- rely on expert knowledge
  - representing metaknowledge in the form of rules that match domain (dataset) characteristics with machine learning algorithms
  - Disadvantages:
    - the resulting rule set is likely to be incomplete
    - timely and accurate maintenance of the rule set as new machine learning algorithms become available is problematic
- automatic procedure
  - we need a pool of problems (datasets) and a set of machine learning algorithms that we wish to consider

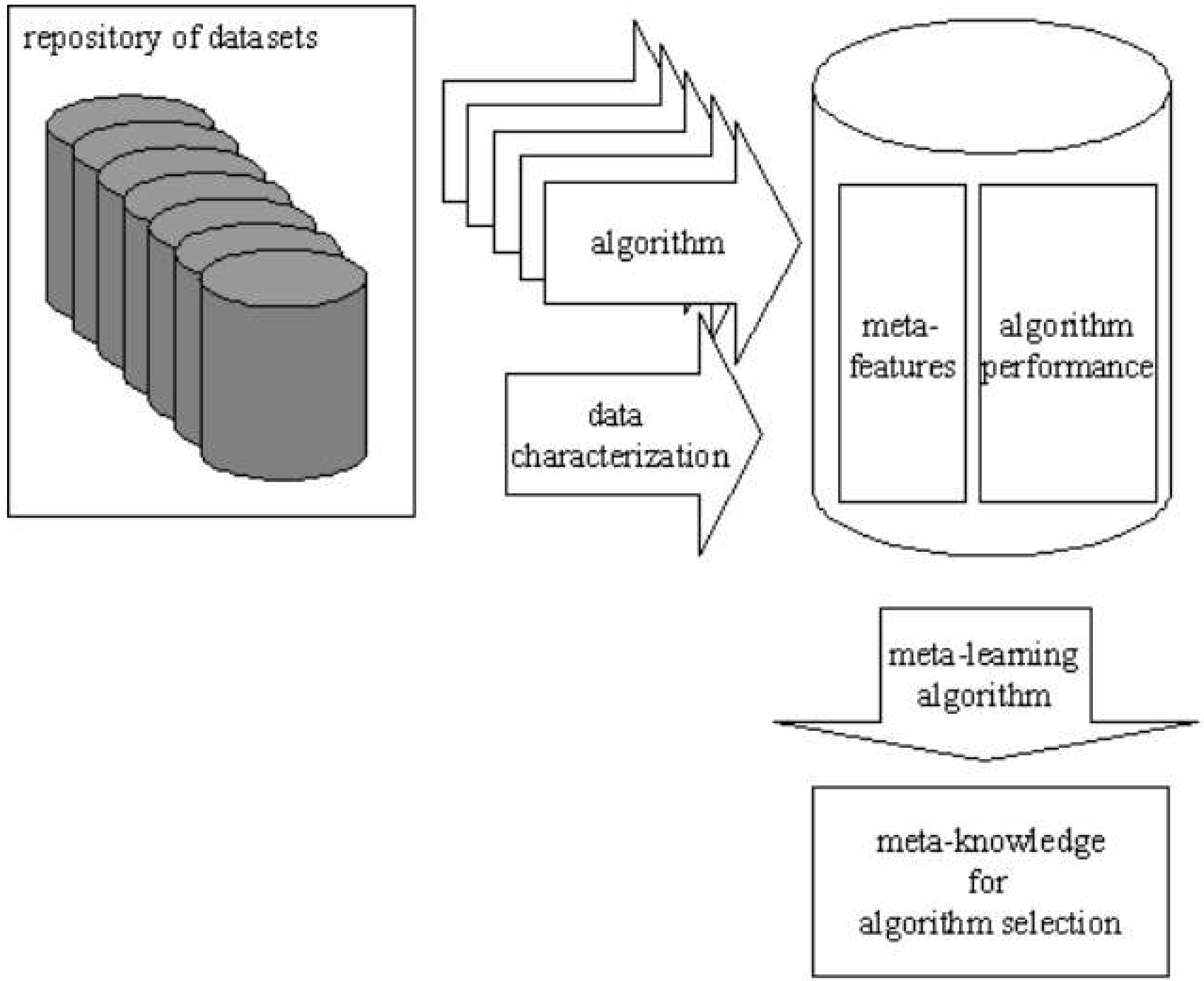


Fig. 2 Acquisition of Metadata for the Metaknowledge base

Meta-learning approaches and its applications

- Metric Learning
  - learning the metric space for predictions
  - provides good results in multiple classification tasks
- Model-Agnostic Meta-Learning (MAML)
  - The neural network is trained using several examples so that the model adapts to new tasks more quickly
  - MAML is a general optimization and task-agnostic algorithm that trains model parameters for fast learning with few gradient updates
- Stacking Generalization
  - subfield of ensemble learning and is used in meta-learning model
  - involves the following steps
    - Learning algorithms are trained using available data
    - A merging algorithm is created to combine all the predictions of these learning algorithms, which are called ensemble members

Advantages of Meta-Learning

- **Higher model prediction accuracy**
  - Optimization of learning algorithms: For example, optimization of hyperparameters to achieve the best results
  - Helps to learn algorithms better adapt to changing conditions
  - Identifying clues for designing better learning algorithms
- **The faster and cheaper training process**
  - Supporting learning from fewer examples
  - Increase the speed of learning processes by limiting the necessary experiments
- **Building more generalized models**:
  - learning to solve many tasks, not just one task: meta-learning does not focus on training one model on one specific data set

Computational Intelligence

- Theory, design, application and development of biologically and linguistically motivated computational paradigms

- Refers to the ability of a computer to learn a specific task from data or experimental observation

- Main pillars:
  - Neural Networks, Fuzzy Systems and Evolutionary Computation

- Set of nature-inspired computational methodologies and approaches to address complex real-world problems to which mathematical or traditional modelling can be useless for a few reasons
- Processes might be too complex for mathematical reasoning, it might contain some uncertainties during the process, or the process might simply be stochastic in nature

Evolutionary computation

- Involves the study of the foundations and the applications of computational techniques based on the principles of natural evolution
- Can be viewed either as search methods, or as optimization techniques
- Three basic mechanisms drive natural evolution: reproduction, mutation, and selection
- These mechanisms act ultimately on the chromosomes containing the genetic information of the individual (the genotype), rather than on the individual itself (the phenotype)
- The degree of adaptation of each individual (i.e. candidate solution) to its environment is expressed by an adequacy measure known as the fitness function

Evolutionary Algorithms (EA)

- In computational intelligence, an evolutionary algorithm is a subset of evolutionary computation, a generic population-based metaheuristic optimization algorithm
- An EA uses mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, and selection
- 1948, Turing: "genetical or evolutionary search"
- 1962, Bremermann: optimization through evolution
- 1964, Rechenberg: evolution strategies
- 1965, L. Fogel, Owens and Walsh: evolutionary programming
- 1975, Holland: genetic algorithms
- 1992, Koza: genetic programming

Evolutionary Algorithms

- More precisely, evolutionary algorithms maintain a population of structures that evolve according to rules of selection and other operators, such as recombination and mutation

- Each individual in the population receives a measure of its fitness in the environment

- Selection focuses attention on high fitness individuals, thus exploiting the available fitness information

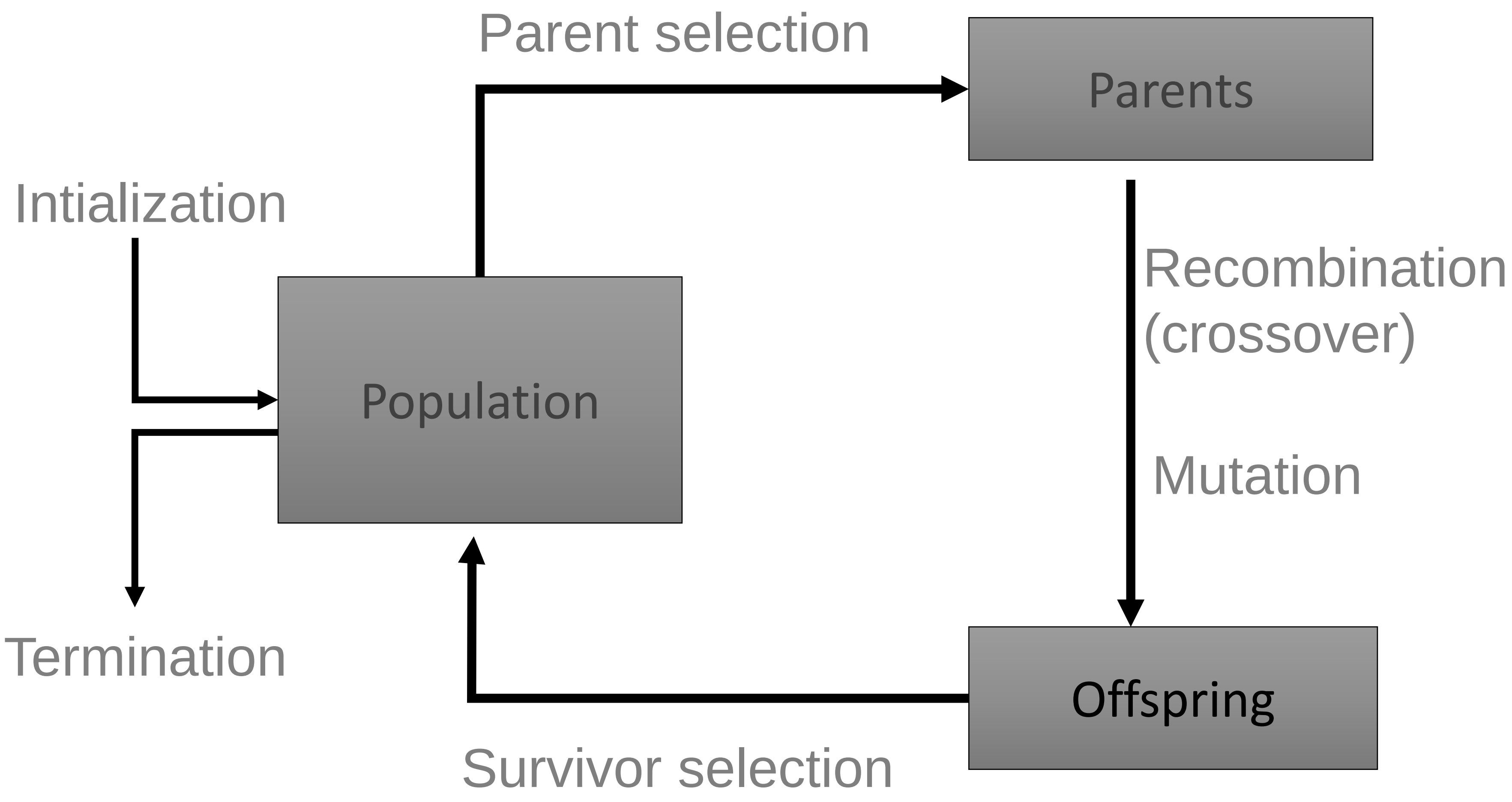
Evolutionary algorithm

- Recombination and mutation perturb those individuals, providing general heuristics for exploration
- Although simplistic from a biologist's viewpoint, these algorithms are sufficiently complex to provide robust and powerful adaptive search mechanisms
- A population of individual structures is initialized and then evolved from generation to generation by repeated applications of evaluation, selection, recombination, and mutation
- The population size N is generally constant in an evolutionary algorithm

Implementation

- Example of a generic single-objective genetic algorithm
- Step 1:
  - Generate the initial population of individuals randomly. (First generation)
- Step 2:
  - Repeat the following regenerational steps until termination:
    1. Evaluate the fitness of each individual in the population (time limit, sufficient fitness achieved, etc.)
    2. Select the fittest individuals for reproduction. (Parents)
    3. Breed new individuals through crossover and mutation operations to give birth to offspring.
    4. Replace the least-fit individuals of the population with new individuals.

General structure of Evolutionary Algorithms (EA)



Sample EA in pseudo-code

```
BEGIN
  INITIALISE population with random candidate solutions;
  EVALUATE each candidate;
  REPEAT UNTIL ( TERMINATION CONDITION is satisfied ) DO
    1 SELECT parents;
    2 RECOMBINE pairs of parents;
    3 MUTATE the resulting offspring;
    4 EVALUATE new candidates;
    5 SELECT individuals for the next generation;
  OD
END
```

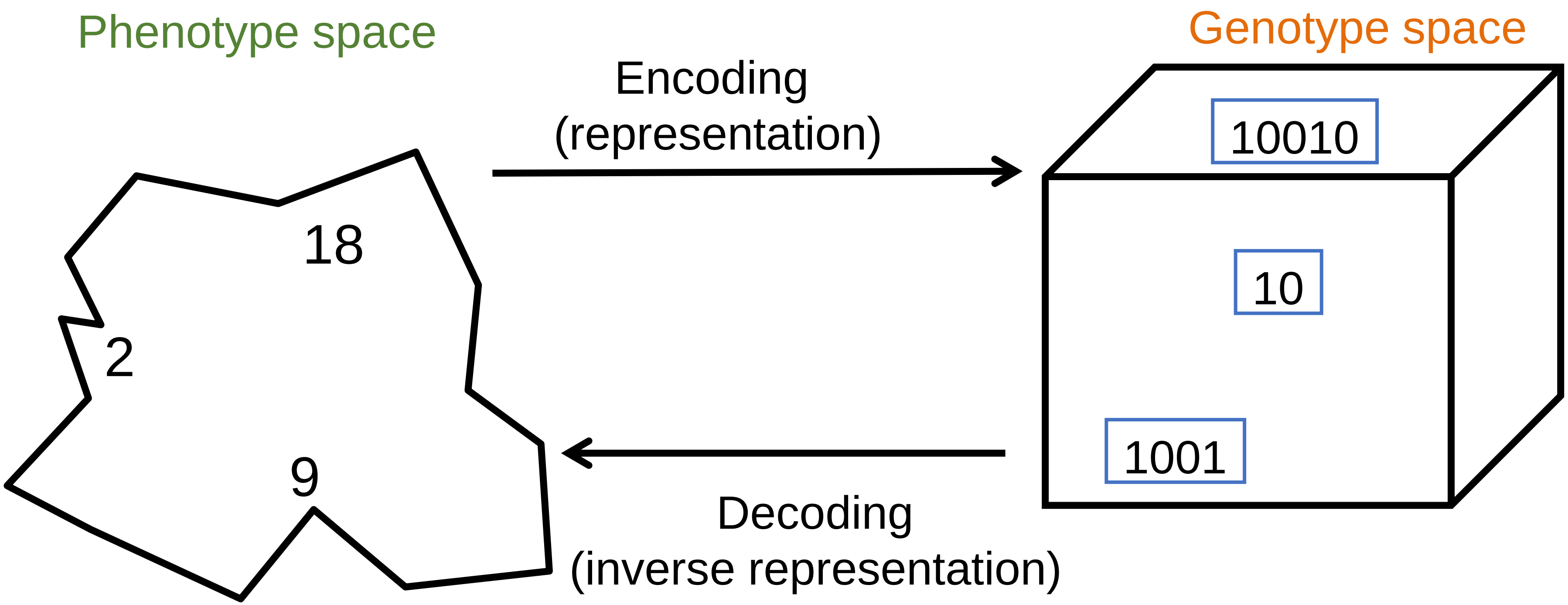
Main EA components: Representation

- Role: provides code for candidate solutions that can be manipulated by variation operators
- Leads to two levels of existence
  - phenotype: object in original problem context, the outside
  - genotype: code to denote that object, the inside (chromosome, “digital DNA”)
- Implies two mappings:
  - Encoding : phenotype=> genotype (not necessarily one to one)
  - Decoding : genotype=> phenotype (must be one to one)
- Chromosomes contain genes, which are in (usually fixed) positions called loci (sing. locus) and have a value (allele)

Each chromosome has two representations.
genotype: The set of genes representing the chromosome.
phenotype: The actual physical representation of the chromosome.

Main EA components: Representation

Example: represent integer values by their binary code



In order to find the global optimum, every feasible solution must be represented in genotype space

Main EA components: Evaluation (fitness) function

- Role:
  - Represents the task to solve, the requirements to adapt to (can be seen as “the environment”)
  - Enables selection (provides basis for comparison)
- A.k.a. *quality* function or *objective* function
- Assigns a single real-valued fitness to each phenotype which forms the basis for selection
  - So the more discrimination (different values) the better
- Typically we talk about fitness being maximised
  - Some problems may be best posed as minimisation problems, but conversion is trivial

Main EA components: Population

- Role: holds the candidate solutions of the problem as individuals (genotypes)
- Formally, a population is a multiset of individuals, i.e. repetitions are possible
- Population is the basic unit of evolution, i.e., the population is evolving, not the individuals
- Selection operators act on population level
- Variation operators act on individual level

Main EA components:Selection mechanism

Role:

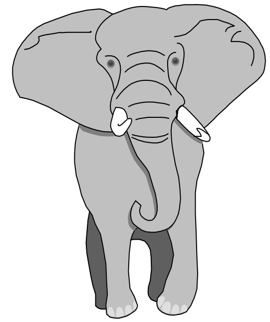
- Identifies individuals
  - to become parents
  - to survive
- Pushes population towards higher fitness
- Usually probabilistic
  - high quality solutions more likely to be selected than low quality but not guaranteed

Mutation Operators

We might have one or more mutation operators for our representation:

- At least one mutation operator should allow every part of the search space to be reached
- The size of mutation is important and should be controllable
- Mutation should produce valid chromosomes

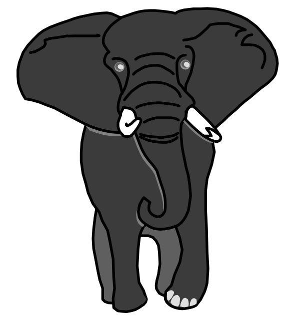
Example: Mutation for Discrete Representation



before 1 1 1 1 1 1 1

after 1 1 1 0 1 1 1

mutated gene



Mutation usually happens with probability p_m for each gene

Example: Mutation for order based representation (Swap)

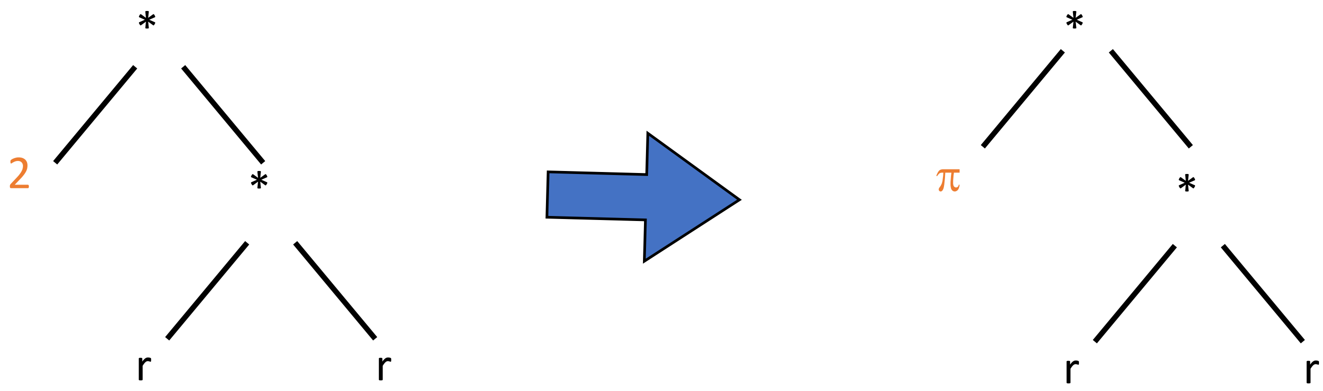
Randomly select two different genes and swap them.

7	3	1	8	2	4	6	5

7	3	6	8	2	4	1	5

Example: Mutation for tree based representation

Single point mutation selects one node and replaces it with a similar one.



Genetic Algorithms

- Widely used in different real-world applications
- **Designing electronic circuits, code-breaking, image processing, and artificial creativity**
- It is a subset of evolutionary algorithms, which is used in computing.
- Uses genetic and natural selection concepts to solve optimization problems
- ***A genetic algorithm is an adaptive heuristic search algorithm inspired by "Darwin's theory of evolution in Nature*** used to solve optimization problems in machine learning.
- It helps solve complex problems that would take a long time to solve

Genetic Algorithms-Terminologies

- **Population:** Population is the subset of all possible or probable solutions, which can solve the given problem.
- **Chromosomes:** A chromosome is one of the solutions in the population for the given problem, and the collection of gene generate a chromosome.
- **Gene:** A chromosome is divided into a different gene, or it is an element of the chromosome.
- **Allele:** Allele is the value provided to the gene within a particular chromosome.
- **Fitness Function:** The fitness function is used to determine the individual's fitness level in the population. It means the ability of an individual to compete with other individuals. In every iteration, individuals are evaluated based on their fitness function.
- **Genetic Operators:** In a genetic algorithm, the best individual mate to regenerate offspring better than parents. Here genetic operators play a role in changing the genetic composition of the next generation.
- **Selection**

Genetic Algorithm-Five phases

- **Initialization**
- **Fitness Assignment**
- **Selection**
- **Reproduction**
- **Termination**

•Refer class notes for sample numerical solved on Genetic algorithm in classroom