King Saud University
College of Medicine

The Role of Artificial Intelligence in Health Informatics: Systemic Literature Review

Dr. Mousa Al-Akhras

BSc, MSc, PhD, Senior MIEEE

m.akhras@seu.edu.sa / mousa.akhras@ju.edu.jo

Mousa.Akhras@ieee.org

Agenda

- Artificial Intelligence
- Know Your Data / Data Preprocessing
- Artificial Intelligence Branches
  - Machine Learning
  - Evolutionary Computation
  - Expert Systems
- Journals Search
- Related Journals
- Sample Articles
- Recommendations

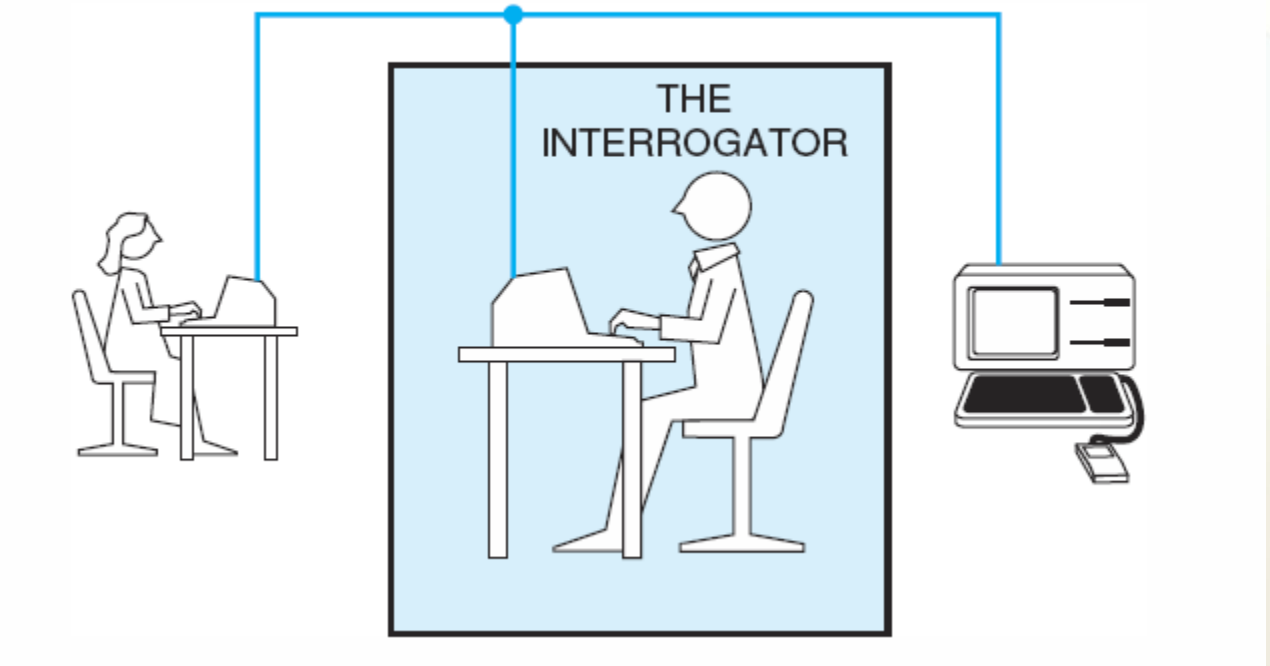
Artificial Intelligence (AI)

- **AI**: The branch of computer science that is concerned with the automation of intelligent behaviour.

- Used to solve problems that can not be solved using algorithmic solution.

King Saud University
College of Medicine

Turing Test: Objective measurement of Intelligence



Turing: An entity deemed intelligent if it demonstrates an ability to achieve human-level performance in all cognitive tasks, sufficient to fool an interrogator

AI Branches

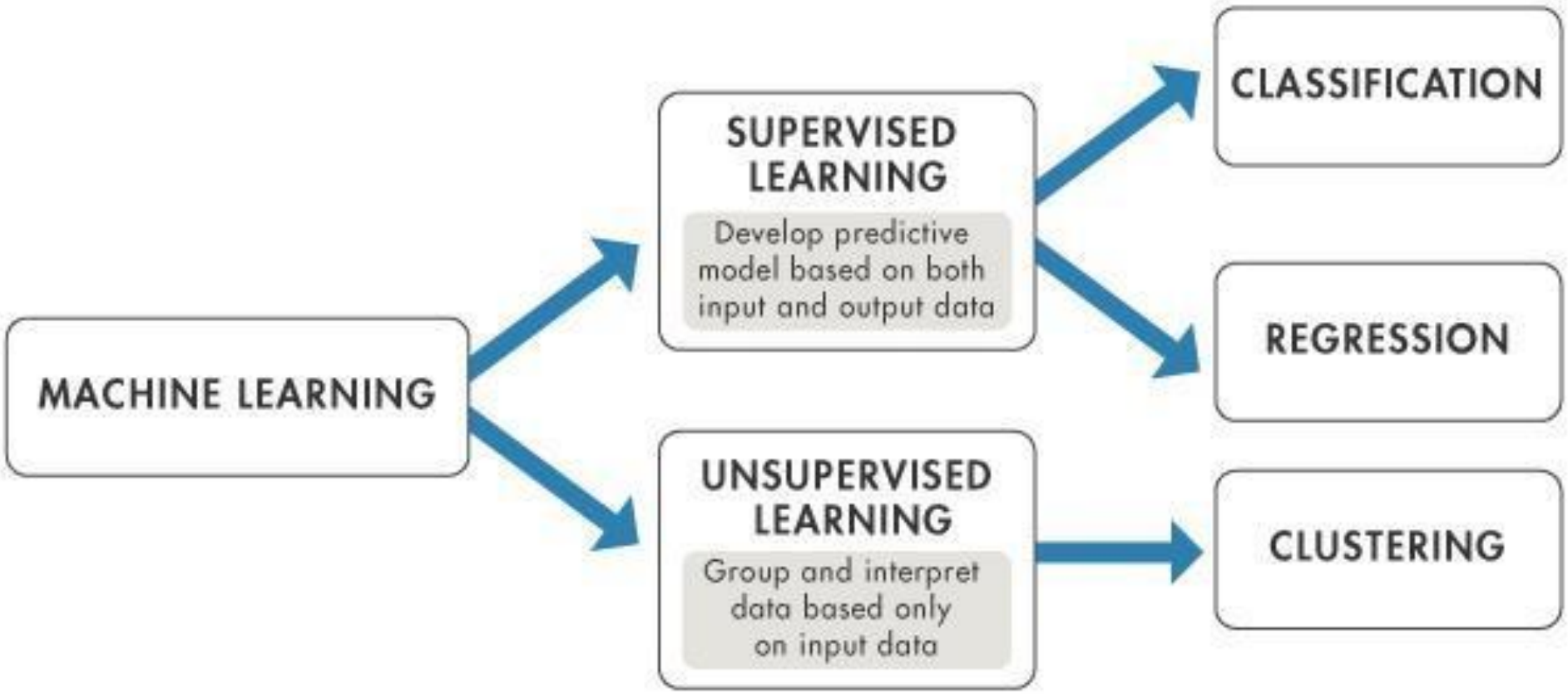
- **Machine learning**, including:
  - Neural networks (NN), Support Vector Machines (SVM)
- **Evolutionary Computation** (EC), including:
  - Genetic algorithms
  - Swarm Intelligence (Ant colony optimization, Particle swarm optimization)
- **Expert Systems**

Machine Learning Tasks

- **Prediction** Methods
  - Use some variables to predict unknown or future values of other variables.
  - E.g. Benign / malignant

- **Description** Methods
  - Find interesting patterns that describe data.

Machine Learning Problems



What is Data?

Attributes

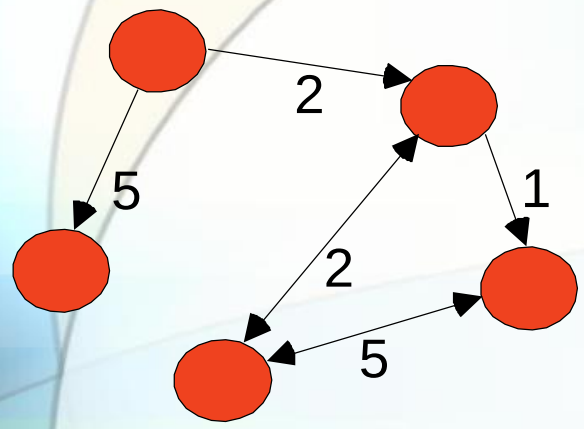
Objects

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

King Saud University
College of Medicine

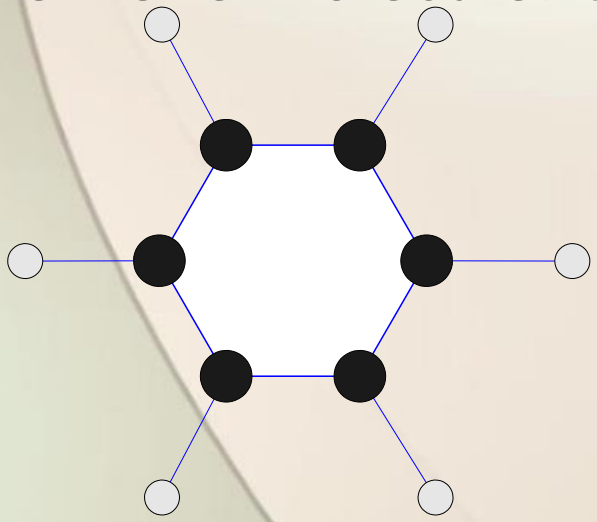
Data Types

Graph Data



Chemical Data

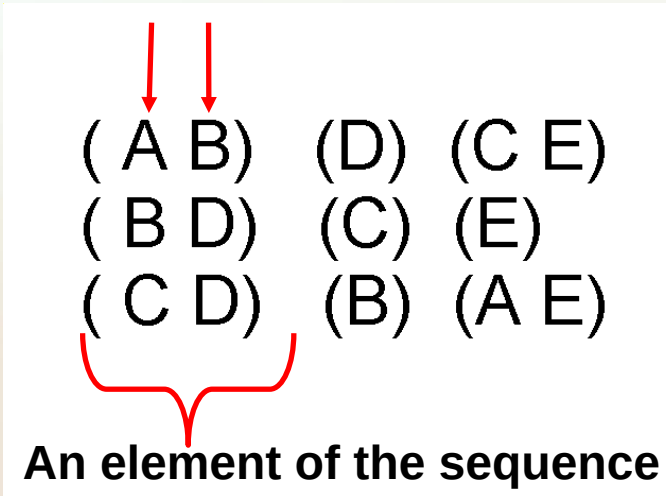
- Benzene Molecule: C_6H_6



Ordered Data

- **Sequences of transactions**

Items/Events



- **Genomic sequence data**

```
GGTTCCGCCTTCAGCCCCGCGCC
CGCAGGGCCCGCCCCGCGCCGTC
GAGAAGGGCCCGCCTGGCGGGCG
GGGGGAGGCGGGGCCGCCCGAGC
CCAACCGAGTCCGACCAGGTGCC
CCCTCTGCTCGGCCTAGACCTGA
GCTCATTAGGCGGCAGCGGACAG
GCCAAGTAGAACACGCGAAGCGC
TGGGCTGCCTGCTGCGACCAGGG
```

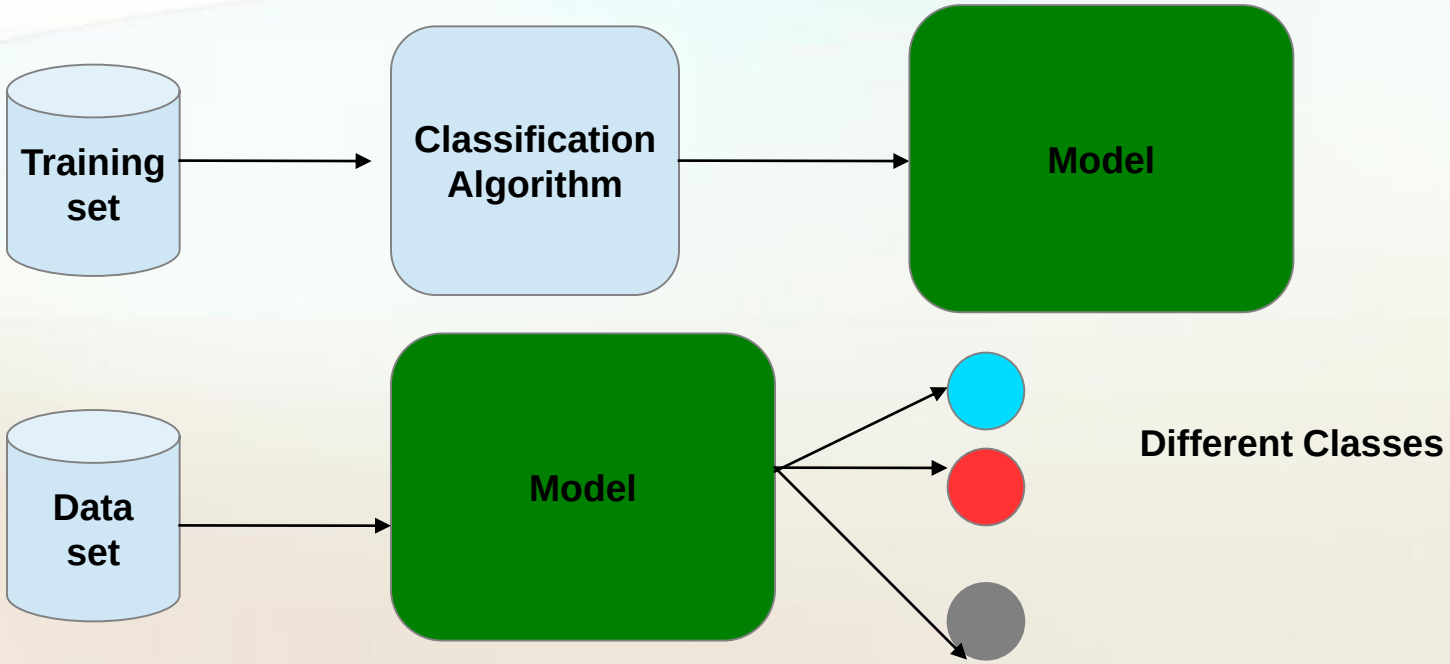
Machine Learning Function :
(1) Association and Correlation Analysis

ID	Items
1	{Bread, Milk}
2	{Bread, Diapers, Eggs}
3	{Milk, Diapers, Cola}
4	{Bread, Milk, Diapers }
5	{Bread, Milk, Diapers, Cola}
…	…

market basket transactions

{Diapers} → { Milk} Example of an association rule

Machine Learning Function : (2) Classification



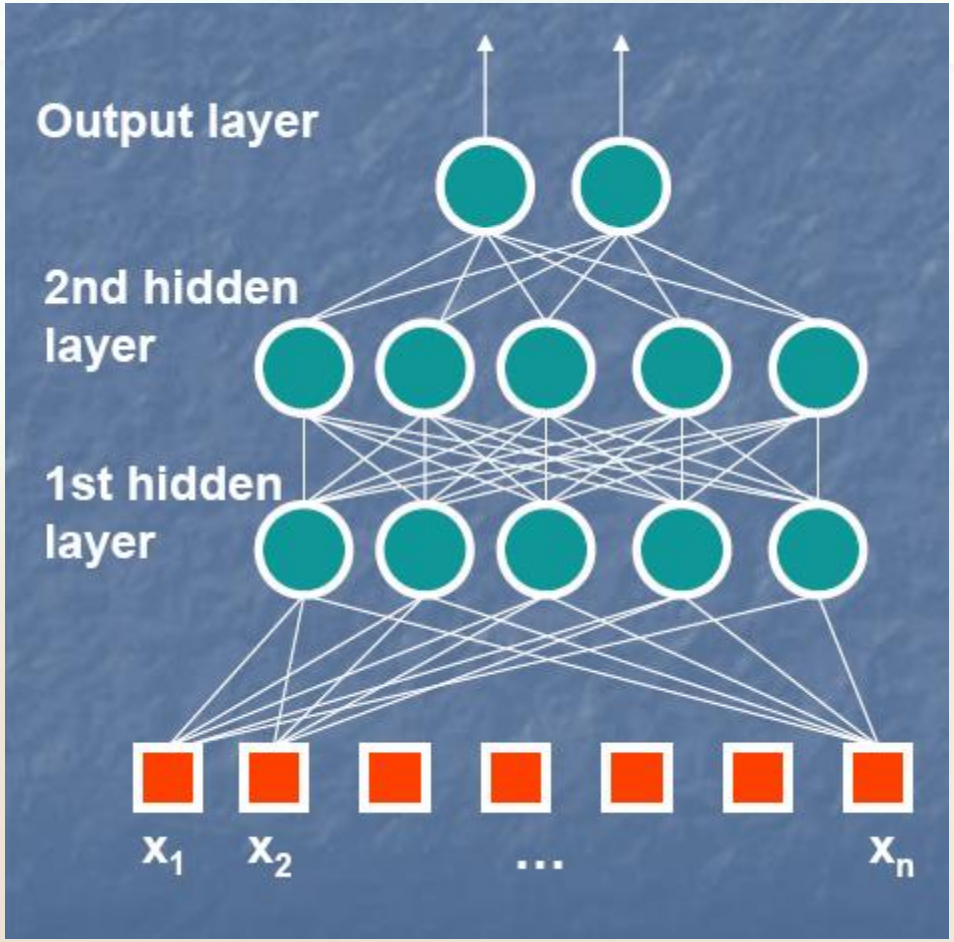
King Saud University
College of Medicine

Machine Learning Function : (2) Classification

Cust ID	age	income	education	Type
1	35	800	udergrad	**risky**
2	26	600	HighSch	**risky**
3	48	1200	grad	**normal**
8	52	2500	udergrad	**good**
44	29	1700	HighSch	**good**

CustID	age	income	Educatin	Type
11	36	850	Udergrd	**?**
27	28	1650	grad	**?**

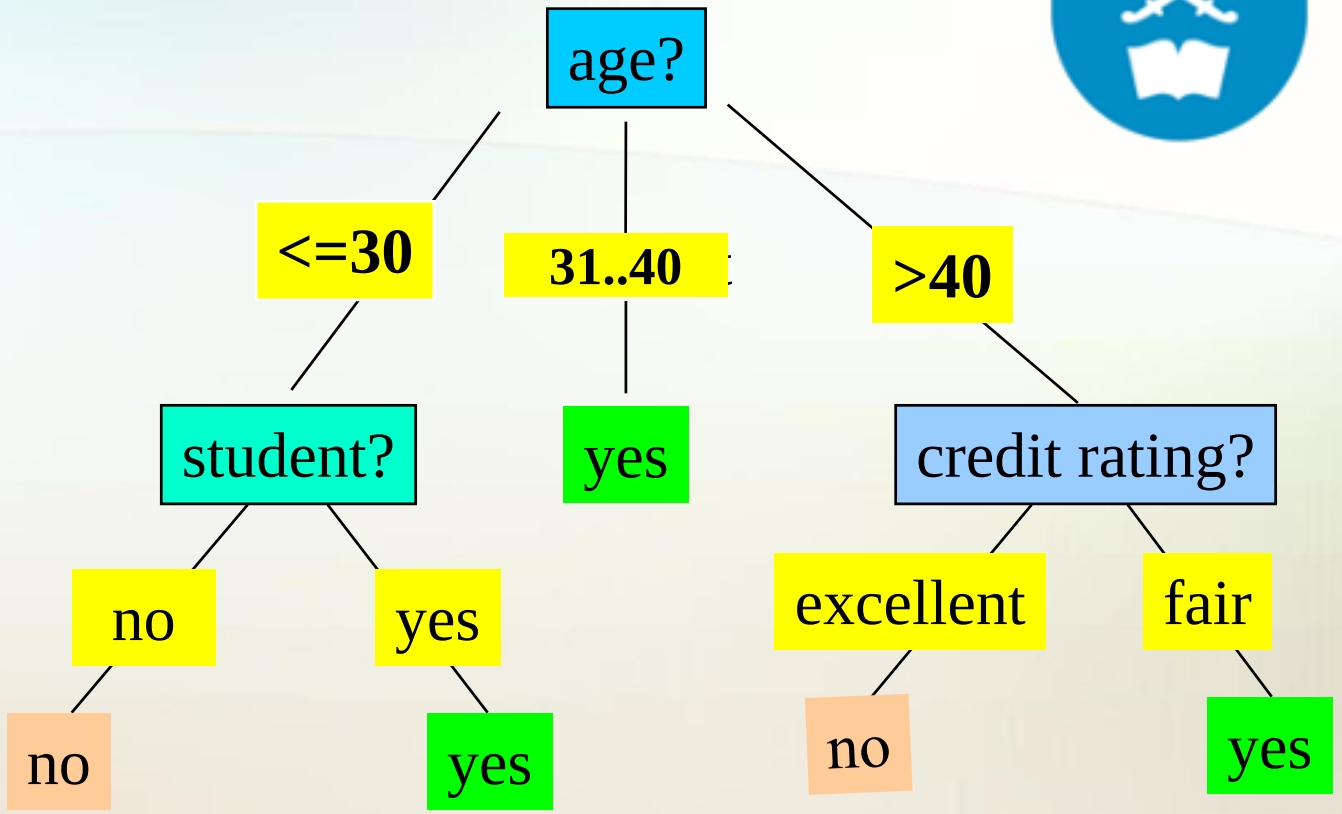
Artificial Neurons – General Structure Multilayers



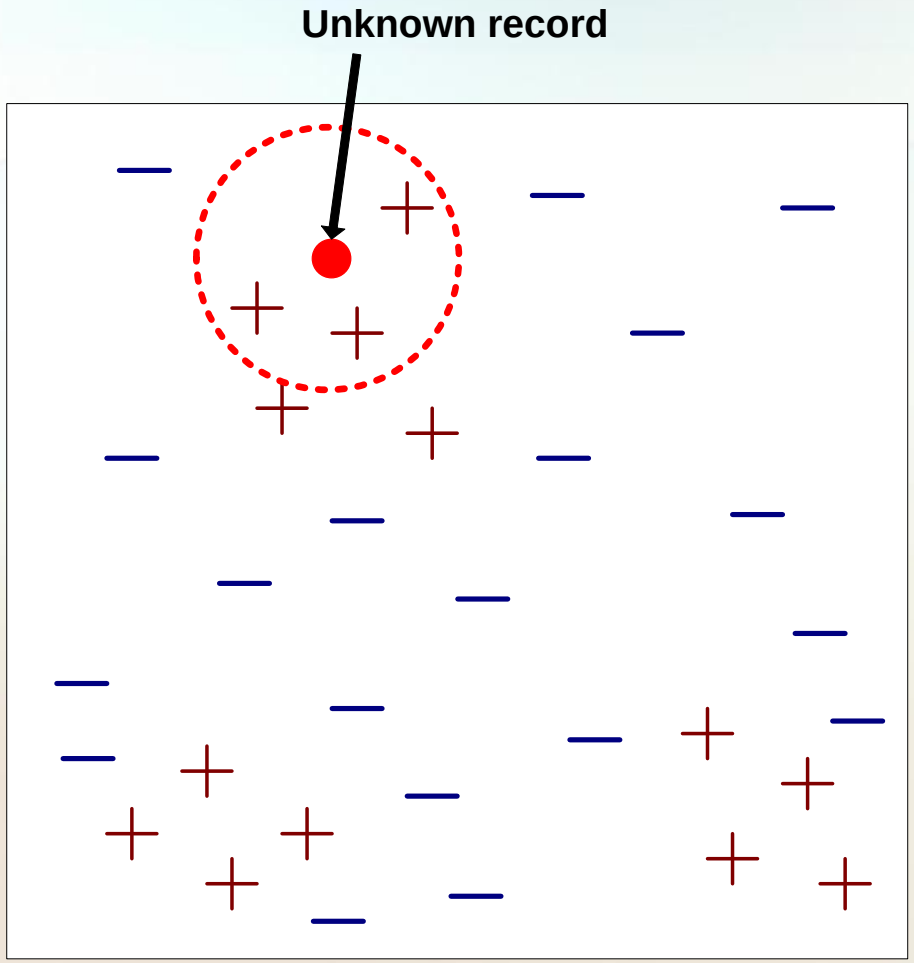
King Saud University
College of Medicine

Machine Learning : (2) Classification – Decision Trees

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31…40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31…40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes
>40	medium	no	excellent	no



Machine Learning : (2) Classification – Nearest Neighbors



Machine Learning : (2) Classification Techniques

- Artificial Neural Networks (ANN)
- Support Vector Machines
- Decision Trees
- Nearest Neighbour
- Rule-Based Classification
- Improve Classification Accuracy: Ensemble Methods
- Naïve Bayes Classifier (Statistical)

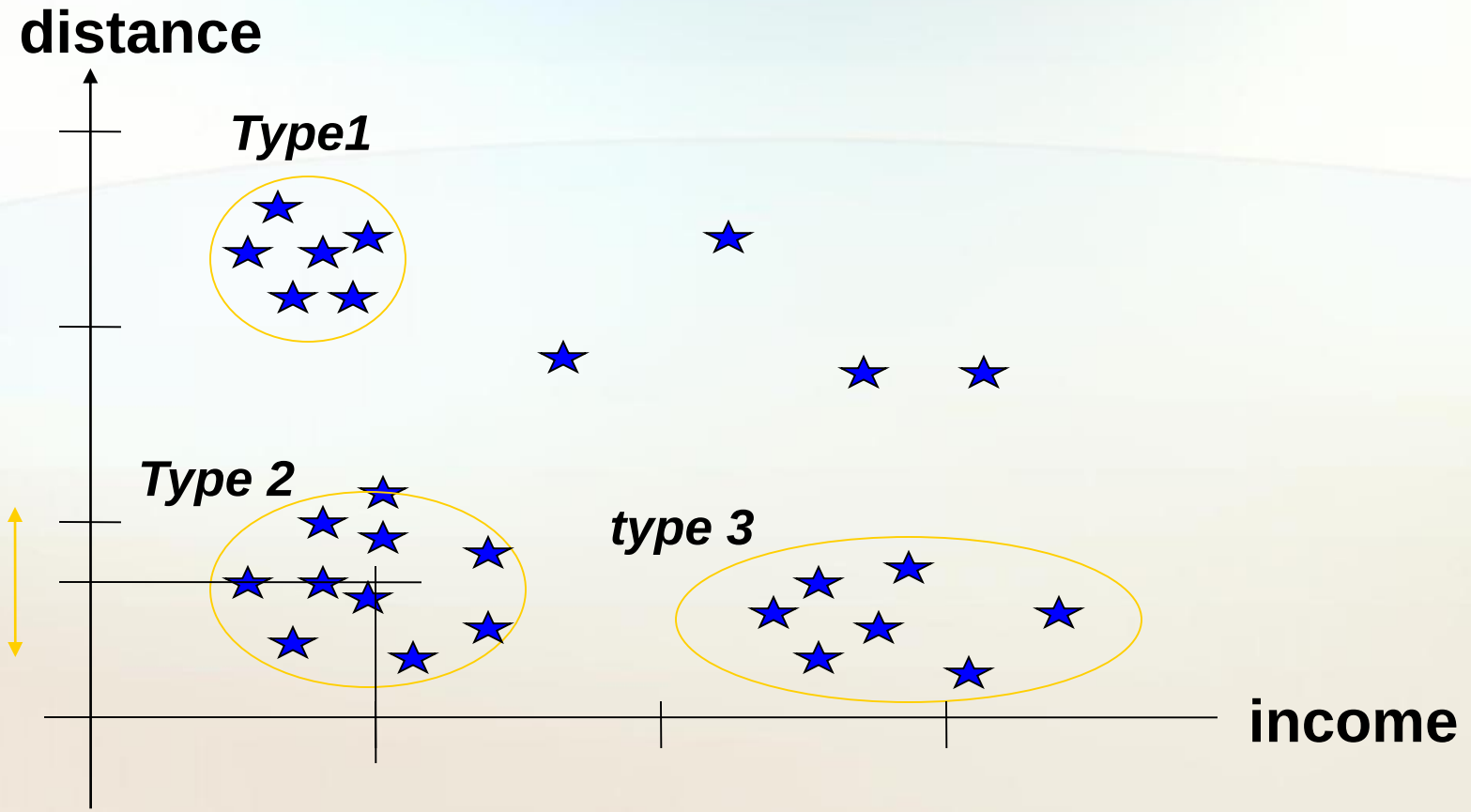
Machine Learning : (3) Cluster Analysis (Grouping)

- Cluster: A collection of data objects
  - similar (or related) to one another within the same group
  - dissimilar (or unrelated) to the objects in other groups

- Unsupervised learning: no predefined classes (i.e., *learning by observations* vs. learning by examples: supervised)

- Good clustering
  - high intra-class similarity: cohesive within clusters
  - low inter-class similarity: distinctive between clusters

Machine Learning : (3) Cluster Analysis (Grouping)



Machine Learning : (3) Clustering –Techniques

- K-Means
- Self-Organizing Feature Map (SOM)
- Fuzzy Clustering

Evolutionary Computation

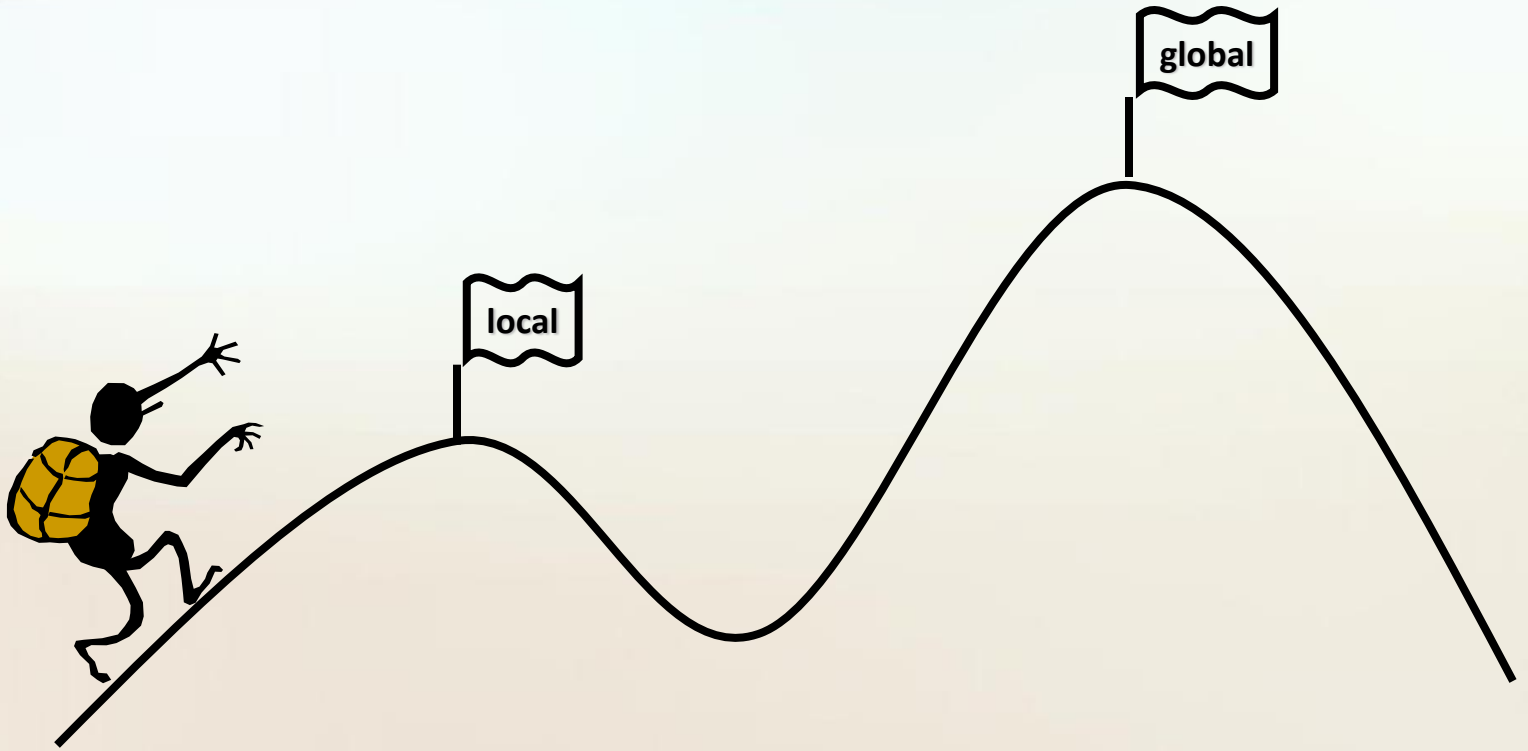
- Genetic Algorithms
- Differential evolution
- Swarm Intelligence (Ant colony optimization, Particle swarm optimization)

- Bio Inspired Optimization Techniques
- Optimization means to find the best (maximum or minimum) solution.

Evolutionary Computation

- Finding the best ingredients
- Glass and a half of milk

Hill Climbing

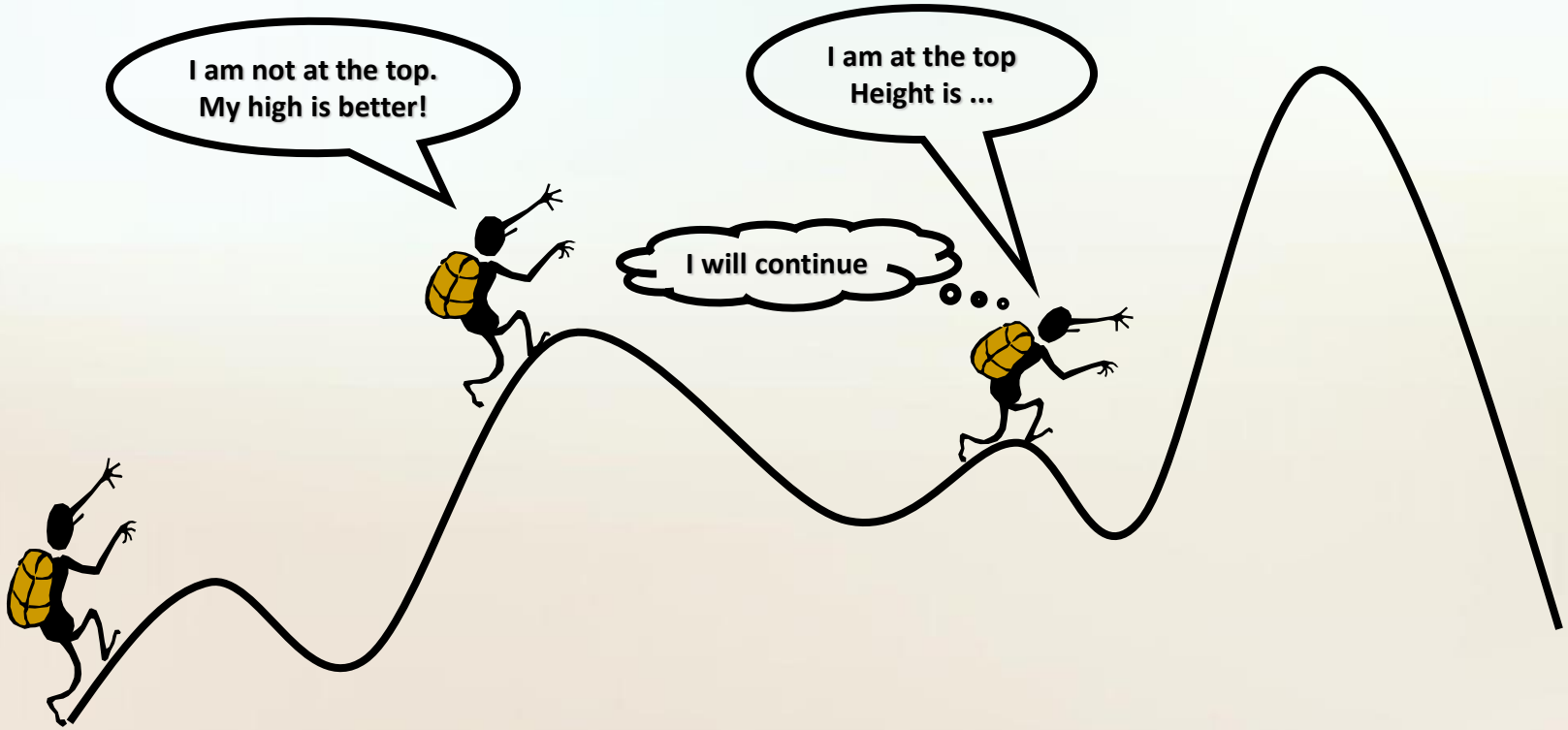


Hill Climbing

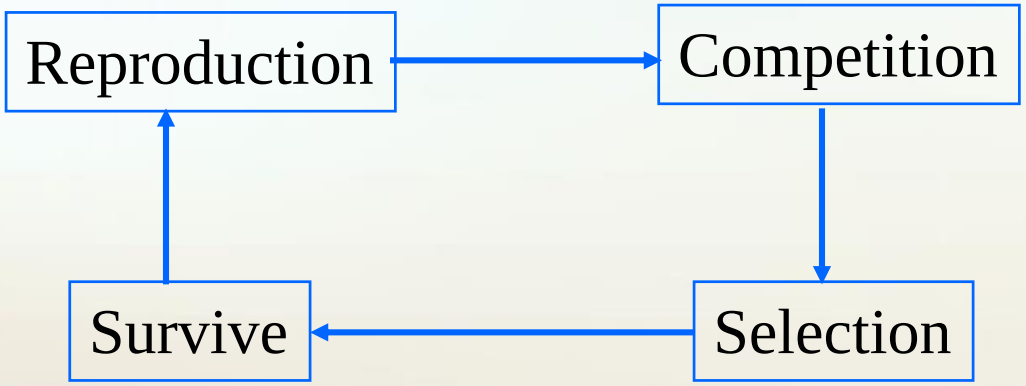
- Multi-climbers

Hill Climbing

- Genetic algorithm



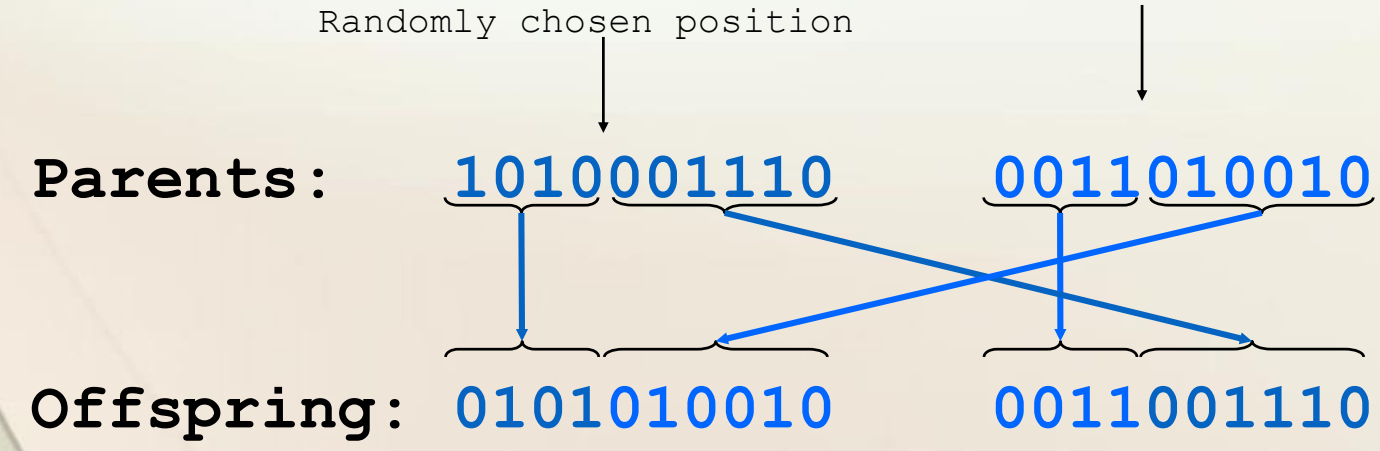
Genetic Algorithm



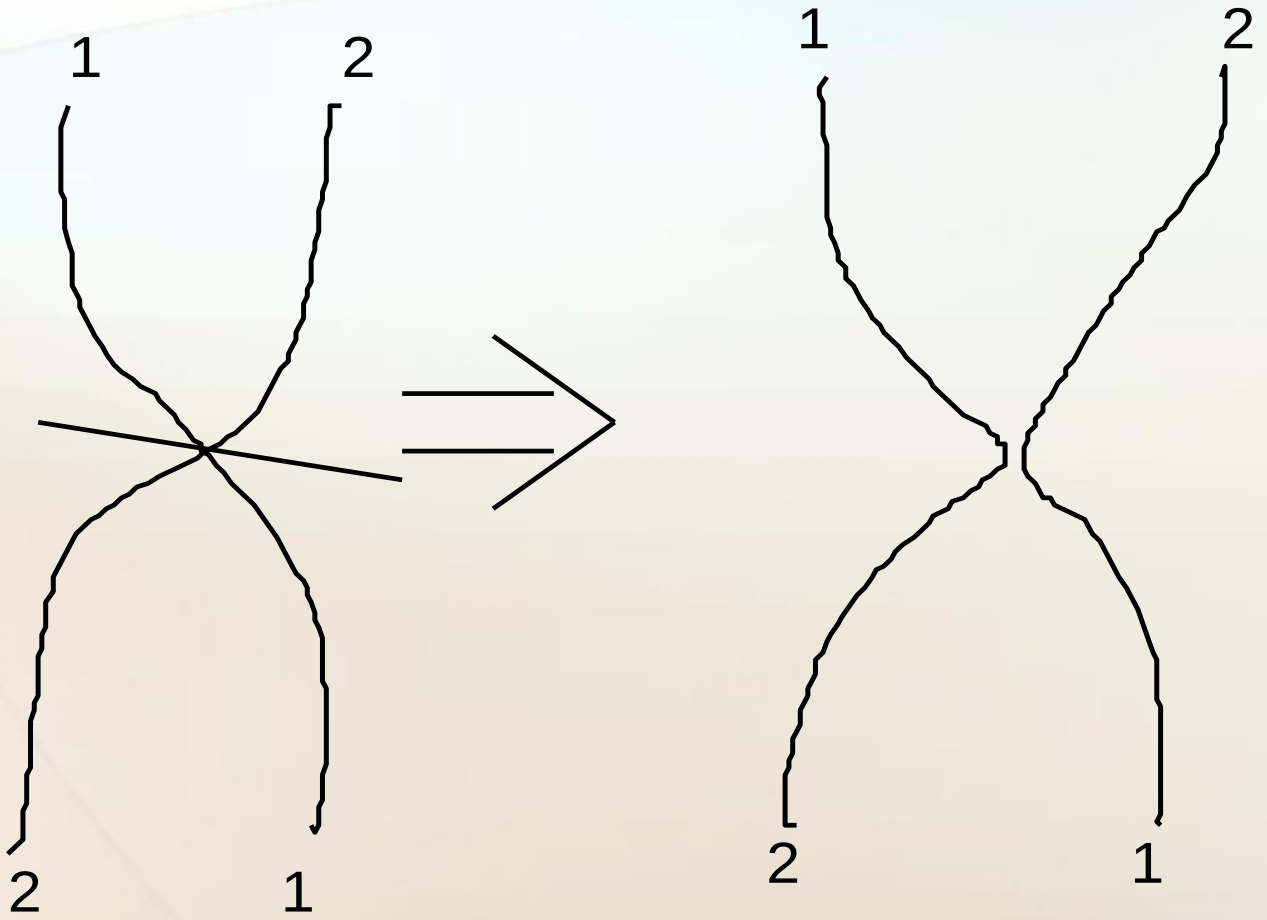
- A robust search and optimization mechanism

One-point crossover 1

- Randomly one position in the chromosomes is chosen
- Child 1 is head of chromosome of parent 1 with tail of chromosome of parent 2
- Child 2 is head of 2 with tail of 1

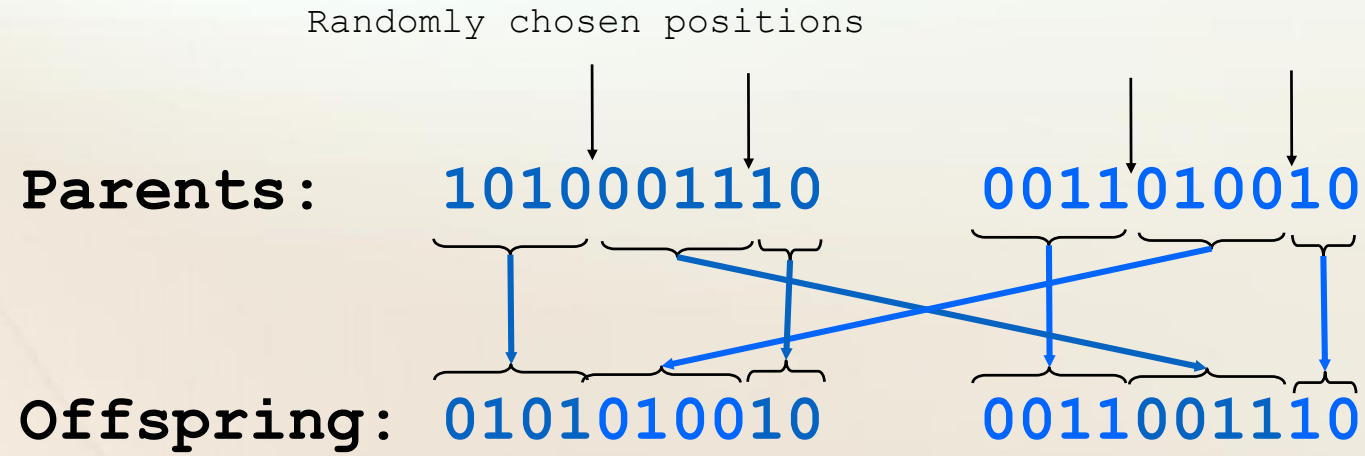


One-point crossover - Nature



Two-point crossover

- Randomly two positions in the chromosomes are chosen
- Avoids that genes at the head and genes at the tail of a chromosome are always split when recombined



King Saud University
College of Medicine

Reproduction Operators - Mutation

Variation: Mutation

- Mutation: changing genes of a parent to create a child
- Swap mutation: Randomly select two genes and swap them

- Inversion mutation: Randomly select a substring and inverse the order of genes inside the substring

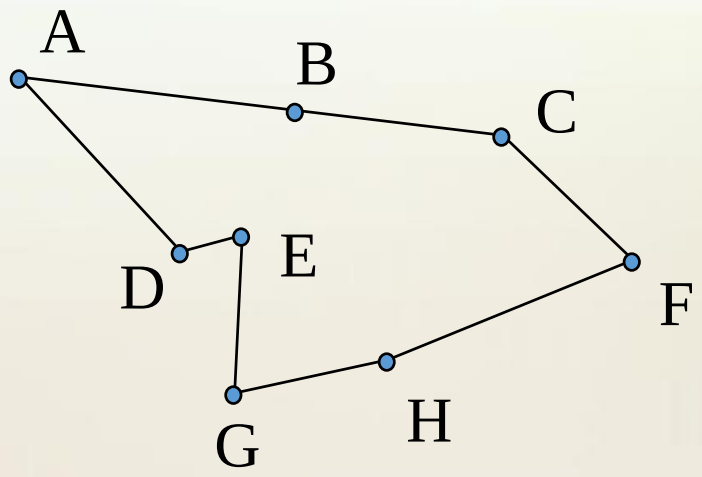
- Scramble mutation: Randomly select a subset of genes and randomly rearrange the alleles (values) of these genes

King Saud University
College of Medicine

Traveling Salesperson Problem

Population

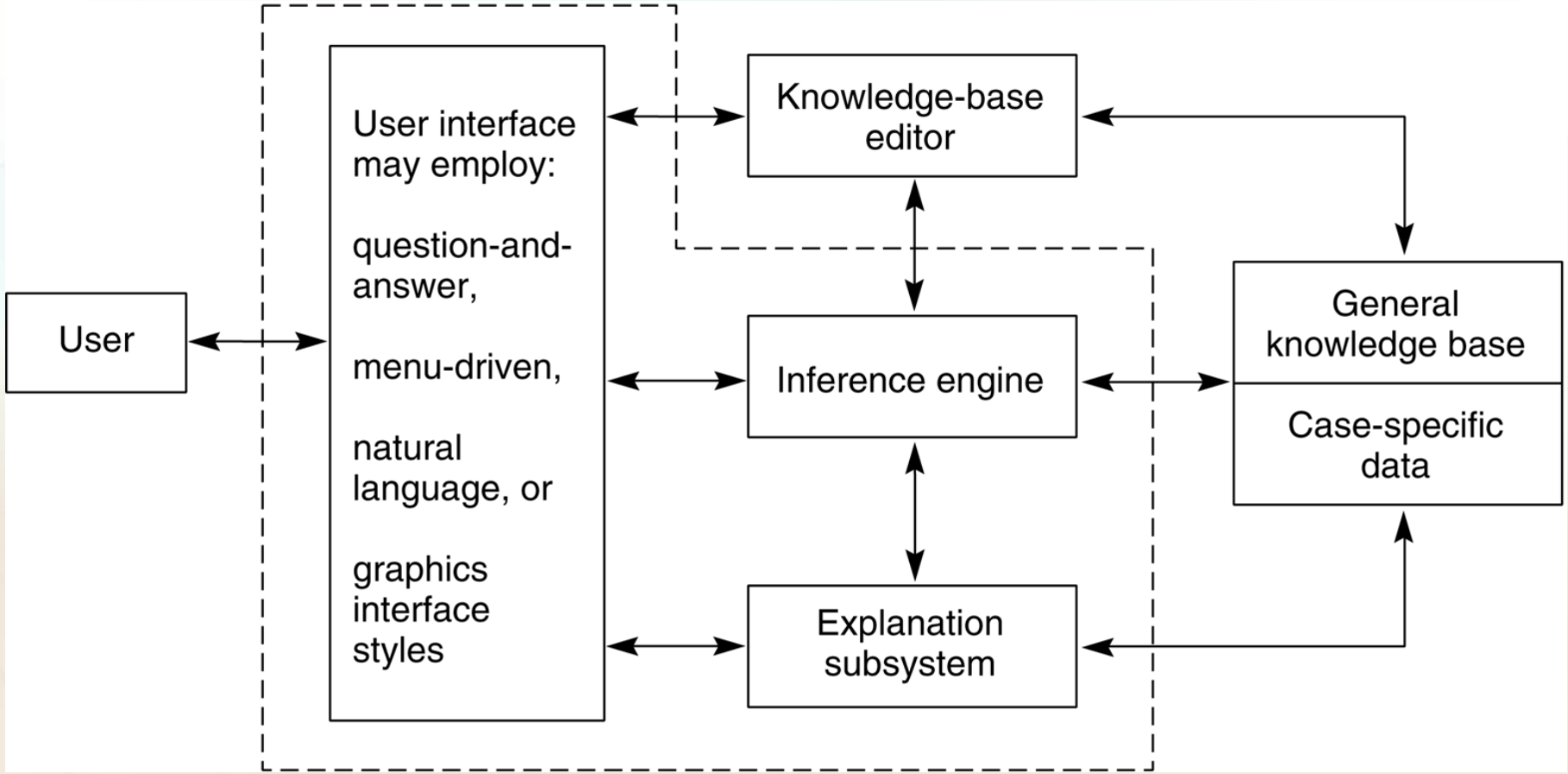
A	B	C	F	H	G	E	D
G	D	A	H	E	C	F	B
C	H	B	F	A	G	D	E
D	C	H	E	G	B	F	A



Expert Systems

- Programs to do reasoning and to solve problems, **diagnosis**
- Modelling an expert: Doctor (diagnose illness), Geologist (discover minerals)
- A **domain-specific knowledge** from a **domain expert** obtained from an **AI specialist** (knowledge engineer).
- e.g. **MYCIN** : Medical system developed in mid 1970s by the medical school at **Stanford** university. Discover bacterial infections with uncertain or incomplete information.
- +ve: Save time, save money, replace the expert in rural areas or when not available, acquire experience from experts

Expert Systems



Practical Data Mining

http://www.cs.waikato.ac.nz/ml/weka/index.html

Journals Search Criteria

- We searched the literature for the use of AI techniques in health related problems
- **Publishers**: Elsevier, IEEE, Springer, and Sage.
- **Journal Title**: “Health” keyword
- **Article Metadata**: “Neural Networks” or “Genetic Algorithms” or “Expert Systems”

Related Journals

- **Elsevier** – 70 Journals
  - Clinical e**Health**
  - Disability and **Health** Journal
  - **Health** Outcomes Research in Medicine
  - **Health** Professions Education
  - International Journal of Hygiene and Environmental **Health**
  - Journal of Pediatric **Health** Care
  - Operations Research for **Health** Care
  - Public **Health**

- **IEEE** – 2 Journals
  - IEEE Journal of Translational Engineering in **Health** and Medicine
  - IEEE Journal of Biomedical and **Health** Informatics

Related Journals

- **Springer** – 8 Journals
  - Administration and Policy in Mental **Health** and Mental **Health** Services Research
  - Advances in **Health** Sciences Education
  - Air Quality, Atmosphere & **Health**
  - Applied Health Economics and **Health** Policy
  - Archives of Public **Health**
  - Archives of Women's Mental **Health**
  - Asian Journal of Gambling Issues and Public **Health**
  - Australia and New Zealand **Health** Policy

Related Journals

- **Sage** – 14 Journals
  - American Journal of **Health** Promotion
  - American Journal of Men's **Health**
  - Asia Pacific Journal of Public **Health**
  - Canadian Journal of Kidney **Health** and Disease
  - Clinical Medicine Insights: Reproductive **Health**
  - Clinical Medicine Insights: Women's **Health**
  - DIGITAL **HEALTH**
  - Environmental **Health** Insights
  - Evaluation & the **Health** Professions
  - Global Advances in **Health** and Medicine
  - Global **Health** Promotion
  - Global Pediatric **Health**
  - **Health**: An Interdisciplinary Journal for the Social Study of **Health**, Illness and Medicine
  - **Health** Education & Behavior

AI Articles in 4 main publishers*

Publisher (Journals)	Neural Networks	Genetic Algorithms	Expert Systems	Total
Elsevier (70)	241	359	5527	6127
IEEE (2)	123	24	41	188
Springer (8)	57	42	1263	1362
Sage (14)	56	70	2926	3052
Total	477	495	9757	**10729**

*: Valid to 16/11/2018

Sample Articles - 1

- J. Shell and W. D. Gregory, "Efficient Cancer Detection Using Multiple **Neural Networks**," in ***IEEE*** *Journal of Translational Engineering in Health and Medicine*, vol. 5, pp. 1-7, **2017**.

- introduce a portable desktop prototype device that provides highly accurate neural network classification of malignant and benign tissue based on

Sample Articles - 2

- F. Miao *et al*., "A Novel Continuous Blood Pressure Estimation Approach Based on **Data Mining** Techniques," in **IEEE** *Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1730-1740, Nov. **2017**.

- Proposes a continuous blood pressure (BP) estimation approach that combines data mining techniques with a traditional mechanism-driven model.
- A **genetic algorithm**-based feature selection method was used to select BP indicators for each subject.
- Multivariate linear regression and support vector regression were employed to develop the BP model.

Sample Articles - 3

- Y. Wang, P. Li, Y. Tian, J. Ren and J. Li, "A Shared Decision-Making System for Diabetes Medication Choice Utilizing Electronic Health Record Data," in *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 5, pp. 1280-1287, Sept. 2017.

- Use shared decision-making for type-2 diabetes mellitus (T2DM) patients utilizing electronic health record (EHR) data in addition to clinical situations to provide a recommended list of available antihyperglycemic medications.

Recommendations

- Study your problem carefully.
- If the problem can not be solved using manual solution, AI will not help.
- If the problem can be used using conventional computation methods, no need to use AI.

Recommendations

- AI Solution:
  - For **classification**, **regression** problems, use **machine learning**.
  - For **grouping** problems use **clustering**.
  - When you need an **expert** program, use **expert systems**.

- Multidisciplinary research needs two+ partners.

King Saud University
College of Medicine

Thank You for Watching!

Any Questions?