Module 2

AI, ML, Deep Learning and Data Mining Methods for 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.	

Arvin Agah - Medical Applications of Artificial Intelligence-CRC Press

2.2 Evolutionary Algorithms, Illustrative Medical Application-Multiagent Infectious Disease Propagation and Outbreak Prediction, Automated Amblyopia Screening System etc.

Infectious disease prediction using machine learning

https://www.nature.com/articles/s41598-021-83926-2

Disease outbreak from Twitter data

Useful link,

Important link

Agentic AI for Epidemiological Forecasting in Ph

Material from book

Automated Amblyopia Screening System

The Hirschberg corneal light reflex test is a screening test that can be used to assess whether a person has strabismus (ocular misalignment)1. It is performed by shining a light in the person’s eyes and observing where the light reflects off the corneas. In a person with normal ocular alignment, the light reflex lies slightly nasal from the center of the cornea (approximately 11 prism diopters—or 0.5mm from the pupillary axis), as a result of the cornea acting as a temporally-turned convex mirror to the observer.

[Material from book]()

2.3 Computational Intelligence Techniques, Deep Learning, Unsupervised learning, dimensionality reduction algorithms

Computational Intelligence Techniques

CI is a set of nature-inspired computing methodologies and approaches to address complex real-world problems where traditional methods fall short. Key pillars include:

1. Artificial Neural Networks (ANNs)

Disease prediction (e.g., cancer, diabetes, heart disease), Medical image analysis (CT, MRI, X-rays), Signal processing (EEG, ECG)

Example: ANN models to detect breast cancer based on mammogram features.

2. Fuzzy Logic (FL)

Handling uncertainty in medical diagnosis, Decision support systems for treatment selection, Modeling symptoms that aren't strictly binary (e.g., "mild fever")

Example: Fuzzy systems for pain assessment or diabetes risk evaluation.

3. Genetic Algorithms (GAs)

Feature selection in diagnostic models, Optimizing medical treatment plans, Scheduling and resource optimization in hospitals

Example: GA to find optimal drug dosages in chemotherapy.

Computational Intelligence Techniques

4. Swarm Intelligence (e.g., PSO, ACO)
Medical image segmentation, Path planning for robotic surgery, Optimizing neural network weights
Example: Particle Swarm Optimization (PSO) for tumor boundary detection in MRI.
5. Deep Learning (DL)
Automatic detection of anomalies in radiology, Natural Language Processing on clinical notes, Genomics and personalized medicine
Example: CNNs for pneumonia detection from chest X-rays.
6. Reinforcement Learning (RL)
Personalized treatment recommendations, Dynamic resource allocation in ICUs
Example: RL-based policies for insulin dosing in type-1 diabetes patients.

https://pmc.ncbi.nlm.nih.gov/articles/PMC4365332/

Computational Intelligence Techniques: Fuzzy logic

Fuzzy C-Means

- An extension of k-means
- Hierarchical, k-means generates partitions
  - each data point can only be assigned in one cluster
- Fuzzy c-means allows data points to be assigned into more than one cluster
  - each data point has a degree of membership (or probability) of belonging to each cluster

Fuzzy C Means Algorithm

Step-1: Randomly initialize the membership matrix using this equation,

$$\sum_{j=1}^{C} \mu_j(x_i) = 1 \qquad i = 1, 2 \ldots k$$

Step-2: Calculate the Centroid using equation,

$$Cj = \frac{\sum_i [\mu_j(x_i)]^m x_i}{\sum_i [\mu_j(x_i)]^m}$$

Step-3: Calculate dissimilarly between the data points and Centroid using the Euclidean distance.

$$D_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Step-4: Update the New membership matrix using the equation,

$$\mu_j(x_i) = \frac{[\frac{1}{d_{ji}}]^{1/m-1}}{\sum_{k=1}^{C} [\frac{1}{d_{ki}}]^{1/m-1}}$$

m = fuzz.
range = (1.25, 2]

Here **m** is a fuzzification parameter.
The range **m** is always [1.25, 2]

Step -5: Go back to Step 2, unless the centroids are not changing.

Worked out Example

- ***Input:*** Number of Objects = 6 Number of clusters = 2

X	Y	C1	C2
1	6	0.8	0.2
2	5	0.9	0.1
3	8	0.7	0.3
4	4	0.3	0.7
5	7	0.5	0.5
6	9	0.2	0.8

Step-1: Initialize the membership matrix.

Step-2: Find the constraint using the equation

$$Cj = \left[\frac{\sum_i [\mu_j(x_i)]^m x_i}{\sum_i [\mu_j(x_i)]^m}, \frac{\sum_i [\mu_j(y_i)]^m y_i}{\sum_i [\mu_j(y_i)]^m}\right]$$

$$C1 = \left[\frac{1*0.8^2+2*0.9^2+3*0.7^2+4*0.3^2+5*0.5^2+6*0.2^2}{0.8^2+0.9^2+0.7^2+0.3^2+0.5^2+0.2^2}, \frac{6*0.8^2+5*0.9^2+8*0.7^2+4*0.3^2+7*0.5^2+9*0.2^2}{0.8^2+0.9^2+0.7^2+0.3^2+0.5^2+0.2^2}\right]$$

$$C1 = \frac{5.58}{2.32}, \frac{14.28}{2.32}$$

$$C1 = (2.4, 6.1)$$

$$C2 = \left[\frac{1*0.2^2 + 2*0.1^2 + 3*0.3^2 + 4*0.7^2 + 5*0.5^2 + 6*0.8^2}{0.2^2 + 0.1^2 + 0.3^2 + 0.7^2 + 0.5^2 + 0.8^2}, \frac{6*0.2^2 + 5*0.1^2 + 8*0.3^2 + 4*0.7^2 + 7*0.5^2 + 9*0.8^2}{0.2^2 + 0.1^2 + 0.3^2 + 0.7^2 + 0.5^2 + 0.8^2}\right]$$

$$C2 = \frac{7.38}{1.52}, \frac{10.48}{1.52}$$

$$C_2 = (4.8, 6.8)$$

Step- 3 : Find Distance

$$D_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Centroid 1 :

$(1,6)(2.4,6.1) = \sqrt{(1.4)^2 + (0.1)^2} = \sqrt{1.96 + 0.01} = \sqrt{1.97} = 1.40$

$(2,5)(2.4,6.1) = \sqrt{0.16 + 1.21} = \sqrt{1.37} = 1.17$

$(3,8)(2.4,6.1) = \sqrt{0.36 + 3.61} = \sqrt{3.97} = 1.99$

$(4,4)(2.4,6.1) = \sqrt{2.56 + 4.41} = \sqrt{6.97} = 2.64$

$(5,7)(2.4,6.1) = \sqrt{6.76 + 0.81} = \sqrt{7.57} = 2.75$

$(6,9)(2.4,6.1) = \sqrt{12.96 + 8.41} = \sqrt{21.37} = 4.62$

Centroid 2:

$$(1,6)(4.8,6.8) = \sqrt{14.44 + 0.64} = \sqrt{15.08} = 3.88$$

$$(2,5)(4.8,6.8) = \sqrt{7.84 + 3.24} = \sqrt{11.08} = 3.32$$

$$(3,8)(4.8,6.8) = \sqrt{3.24 + 1.44} = \sqrt{4.68} = 2.16$$

$$(4,4)(4.8,6.8) = \sqrt{0.64 + 7.84} = \sqrt{8.48} = 2.91$$

$$(5,7)(4.8,6.8) = \sqrt{0.04 + 0.04} = \sqrt{0.08} = 0.28$$

$$(6,9)(4.8,6.8) = \sqrt{1.44 + 4.84} = \sqrt{6.28} = 2.50$$

Cluster 1		Cluster 2	
Datapoint	**Distance**	**Datapoint**	**Distance**
(1,6)	1.40	(1,6)	3.88
(2,5)	1.17	(2,5)	3.32
(3,8)	1.99	(3,8)	2.16
(4,4)	2.64	(4,4)	2.91
(5,7)	2.75	(5,7)	0.28
(6,9)	4.62	(6,9)	2.50

Step-4 : Update the membership value

$$\mu_j(x_i) = \frac{[\frac{1}{d_{ji}}]^{1/m-1}}{\sum_{k=1}^{c}[\frac{1}{d_{ki}}]^{1/m-1}}$$

here m = 2, i – first data point, j - first cluster

Cluster 1

$\mu_{11} =$ $(1 / d_{11})^{1/2-1} / (1 / d_{11})^{1/2-1} + (1 / d_{21})^{1/2-1}$

$= (1/1.40)^{1} / (1/1.40)^{1} + (1/3.88)^{1} = 0.71 / 0.71+0.25$

$= 0.71 / 0.96 = 0.7$

$\mu_{12} =$ $(1 / d_{12}) / (1 / d_{12}) + (1 / d_{22})$

$=> 1/1.17 / 1/1.17 + 1/3.32 = 0.56 / 0.56+0.30$

$= 0.56 / 0.86 = 0.6$

$\mu_{13} =$ $(1 / d_{13}) / (1 / d_{13}) + (1 / d_{23})$

$=> 1/1.99 / 1/1.99 + 1/2.16 = 0.50 / 0.50+0.46$

$= 0.50 / 0.96 = 0.5$

$\mu_{14} =$ $(1 / d_{14}) / (1 / d_{14}) + (1 / d_{24})$

$=> 1/2.64 / 1/2.64 + 1/2.91 = 0.37 / 0.37+0.34$

$= 0.37 / 0.71 = 0.5$

$\mu_{15} =$ $(1 / d_{15}) / (1 / d_{15}) + (1 / d_{25})$

$=> 1/2.75 / 1/2.75 + 1/0.28 = 0.36 / 0.36+3.57$

$= 0.36 / 3.93 = 0.1$

$\mu_{16} =$ $(1 / d_{16}) / (1 / d_{16}) + (1 / d_{26})$

$=> 1/4.62 / 1/4.62 + 1/2.50 = 0.21 / 0.21+0.4$

$= 0.21 / 0.61 = 0.3$

Cluster 2

$$\mu_{21} = (1 / d_{21}) / (1 / d_{12}) + (1 / d_{21})$$

$$\Rightarrow 1/3.88 \;/\; 1/1.40 + 1/3.88 = 0.25 \;/\; 0.71+0.25$$

$$= 0.25 / 0.96 = 0.3$$

$$\mu_{22} = (1 / d_{22}) / (1 / d_{12}) + (1 / d_{22})$$

$$\Rightarrow 1/3.32 \;/\; 1/1.17 + 1/3.32 = 0.30 \;/\; 0.56+0.30$$

$$= 0.30 / 0.86 = 0.4$$

$$\mu_{23} = (1 / d_{23}) / (1 / d_{13}) + (1 / d_{23})$$

$$\Rightarrow 1/2.16 \;/\; 1/1.99 + 1/2.16 = 0.46 \;/\; 0.50+0.46$$

$$= 0.46 / 0.96 = 0.5$$

$$\mu_{24} = (1 / d_{24}) / (1 / d_{14}) + (1 / d_{24})$$

$$\Rightarrow 1/2.19 \;/\; 1/2.64 + 1/2.19 = 0.34 \;/\; 0.37+0.34$$

$$= 0.34 / 0.71 = 0.5$$

$$\mu_{25} = (1 / d_{25}) / (1 / d_{15}) + (1 / d_{25})$$

$$\Rightarrow 1/0.28 \;/\; 1/2.75 + 1/0.28 = 3.57 \;/\; 0.36+3.57$$

$$= 3.57 / 3.93 = 0.9$$

$$\mu_{26} = (1 / d_{26}) / (1 / d_{16}) + (1 / d_{26})$$

$$= 0.4 / 0.21+0.4 = 0.4/0.61$$

$$= 0.7$$

- Now the New Membership value is

X	Y	C1	C2
1	6	0.7	0.3
2	5	0.6	0.4
3	8	0.5	0.5
4	4	0.5	0.5
5	7	0.1	0.9
6	9	0.3	0.7

Step 5 : Now continue this process until get the same centroids.

Worked out Example

- ***Input:*** Number of Objects = 6 Number of clusters = 2

X	Y	C1	C2
1	6	0.8	0.2
2	5	0.9	0.1
3	8	0.7	0.3
4	4	0.3	0.7
5	7	0.5	0.5
6	9	0.2	0.8

1. Evolutionary Algorithms (EAs)

- Definition: Stochastic search & optimization methods inspired by natural evolution.
- Core Process:

1. Fitness Function – guides evolution by evaluating candidate quality.
2. Selection – fitter candidates chosen as parents.
3. Reproduction Operators:
   - Recombination → combine parents → children.
   - Mutation → alter a single candidate.
4. Iteration (Generations) – cycle continues until termination (quality goal, time/computation limit).

🔑 Mnemonic: “F-SRIT” → Fitness, Selection, Recombination, Iteration, Termination.

---

2. Illustrative Medical Applications

a) Multiagent Infectious Disease Propagation & Outbreak Prediction

- Purpose: Predict and forecast disease spread.
- Data Sources:
- Real-time (e.g., social media).
- Historical epidemiological data.
- Example: Malaria outbreak prediction using rainfall, temperature, case counts + ANNs.

🔑 Mnemonic: “RAIN” → Rainfall, ANN, Infection trends, Number of cases.

---

b) Automated Amblyopia Screening System

- Condition: Amblyopia = “lazy eye,” often linked to strabismus (ocular misalignment).
- Method: Hirschberg corneal light reflex test → shine light in eyes → check corneal reflections.
- Normal Alignment: Reflex slightly nasal to corneal center.
- AI Role: Automated detection → faster, accessible screening.

🔑 Mnemonic: “HAS” → Hirschberg test, Amblyopia, Strabismus.

Master Mnemonic for Module 2.2

👉“Fast Robots Help Against Pandemics & Screen Eyes.”

- Fast → F-SRIT (Fitness, Selection, Recombination, Iteration, Termination)
- Robots → Evolutionary Algorithms (automation, adaptation)
- Help Against Pandemics → Outbreak Prediction (RAIN)
- Screen Eyes → Automated Amblyopia Screening (HAS)

Unsupervised learning

Unsupervised learning in healthcare

Unsupervised learning in healthcare involves using machine learning algorithms to analyze unlabeled data and discover hidden patterns, structures, and relationships within it. This approach is particularly useful for tasks like disease phenotyping, anomaly detection, and data exploration in complex healthcare datasets.

Disease Phenotyping:

Identifying distinct subgroups or subtypes of diseases based on patient data, such as genetic information, electronic health records, or medical imaging.

Anomaly Detection:

Identifying unusual patterns or outliers in patient data that could indicate potential health risks or fraud.

Data Exploration and Pattern Discovery:

Exploring large and complex datasets to uncover hidden relationships and insights that can inform research and clinical decision-making.

Unsupervised learning in healthcare contd…

Personalized Treatment Recommendations:
Identifying patient subgroups that respond differently to treatments, paving the way for more personalized and effective care.
Clustering:
Grouping similar patients or data points based on their characteristics, which can be used for targeted interventions or resource allocation.
Drug Discovery:
Analyzing molecular data to identify potential drug candidates and predict their effectiveness.

Unsupervised learning in healthcare contd…

Algorithms used

K-means Clustering:

Groups data points into k clusters based on their proximity to cluster centers.

Hierarchical Clustering:

Builds a hierarchy of clusters, allowing for the exploration of data at different levels of granularity.

Principal Component Analysis (PCA):

Reduces the dimensionality of data by identifying the principal components that capture the most variance in the data.

Autoencoders:

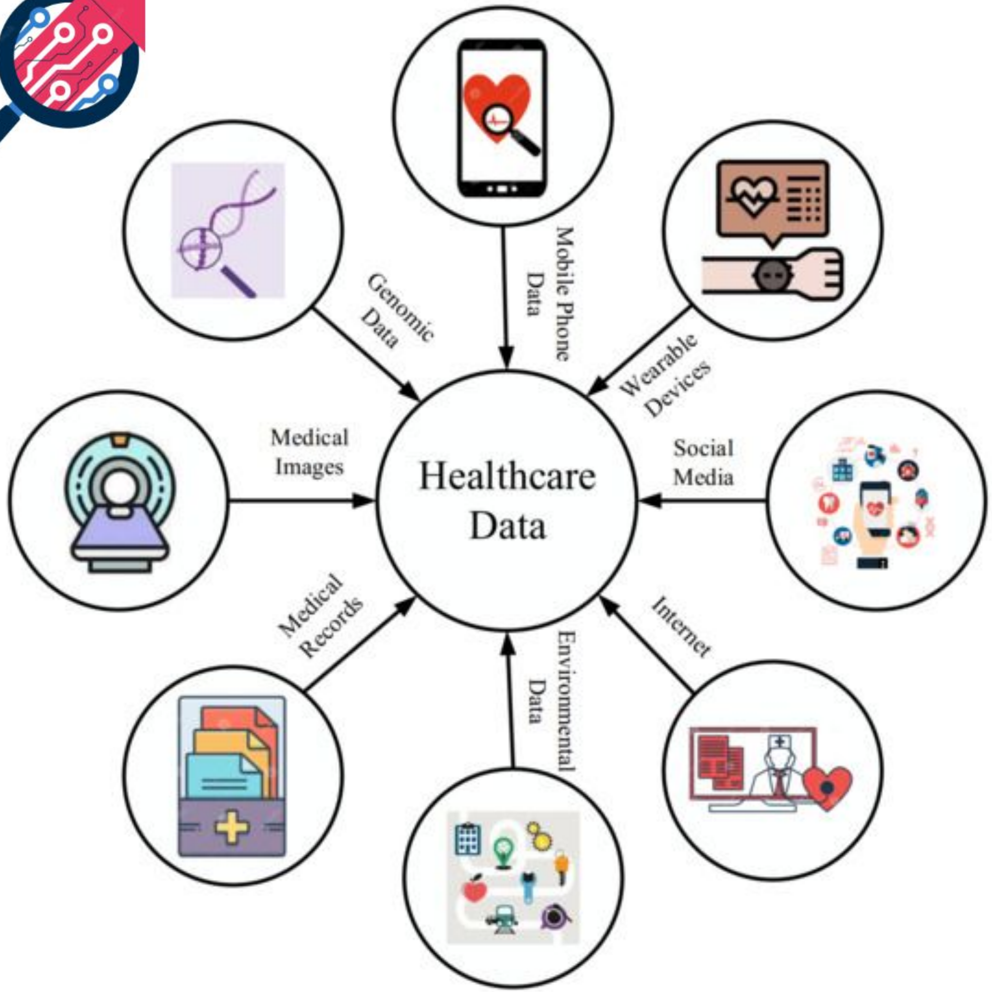
Neural networks that learn to compress and reconstruct data, which can be used for anomaly detection and dimensionality reduction.

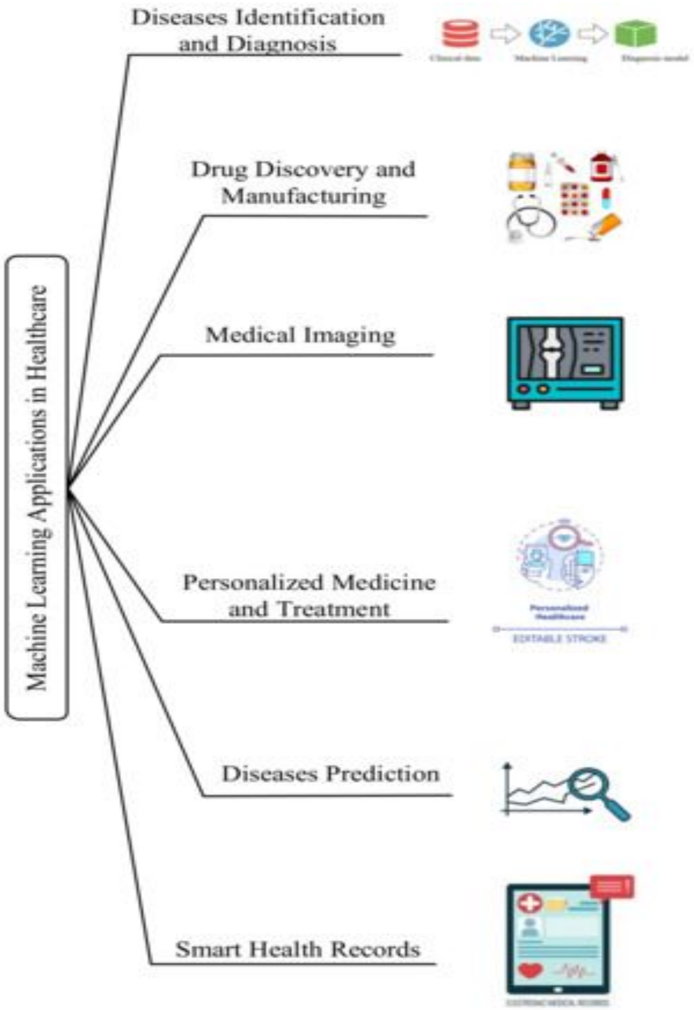
Unsupervised learning in healthcare contd…

https://link.springer.com/article/10.1007/s12553-023-00805-8

https://pubmed.ncbi.nlm.nih.gov/36378293/

Dimensionality Reduction


Healthcare Data
Mobile Phone Data
Wearable Devices
Social Media
Internet
Environmental Data
Medical Records
Medical Images
Genomic Data


Machine Learning Applications in Healthcare
Diseases Identification and Diagnosis
Drug Discovery and Manufacturing
Medical Imaging
Personalized Medicine and Treatment
Diseases Prediction
Smart Health Records

Dimensionality Reduction

Although the message is broadly the more data, the better, datasets can often contain many variables and as a result make capturing the signal of the data a more laborious task.
Whether it's **sparse values, missing values, identifying relevant features, resource efficiency, or more straightforward interpretation**, dimensionality reduction algorithms are very useful to data scientists.

What data is relevant to understanding future risk of disease or adverse event?
Dimensionality reduction refers to converting a dataset of many dimensions into fewer dimensions while concisely representing similar data

Dimensionality Reduction Algorithms

The need for dimensionality reduction becomes apparent when you consider the wealth of data available to make narrow decisions:

• Mobile phones collect hundreds of data points including calls, texts, steps, calories burned, floors climbed, Internet usage, and so forth. What data is best for understanding phone usage?
• Brands on social media are collecting data on engagement and interactions such as comments, likes, followers, sentiment, and mood. What data is best for understanding attitudes toward health?
• Medical health records contain a wealth of information, but only some information is relevant in predicting disease risk or illness progression.

https://www.mdpi.com/2227-7390/9/22/2970

Dimensionality Reduction Algorithms

Dimension reduction is extremely useful in machine learning tasks, with a plethora of benefits:

Fewer dimensions result in quicker computations when compared to the original dataset.

By default, dimension reduction algorithms reduce the space required for storage.

Reducing data into less than three dimensions enables visualization and easier understanding.

Redundant data is removed, which improves the performance of the machine learning model.

Noise is removed, which improves model performance

https://www.sciencedirect.com/science/article/abs/pii/S0026265X21006949#:~:text=Spectroscopic%20and%20chromatographic,popular%20multivariate%20tools.

https://www.datacamp.com/tutorial/principal-component-analysis-in-python