Automated Design of Computational Intelligence Techniques for Disease Prediction

Nelishia Pillay and Thambo

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Tutorial Website

https://www.cs.up.ac.za/cs/npillay/SCCI2025Tutorial.htm

Tutorial Outline

- Introduction
  - Disease prediction (TN)
  - Computational intelligence for disease prediction (TN)
  - Multimodal learning for disease prediction (NP)
  - Automated design (NP)
  - Hyper-heuristics for automated design (NP)
  - Evolutionary algorithms for automated (TN)
- Case Studies
  - Evolutionary algorithms: genetic algorithms + genetic programming (TN)
  - Images processing using genetic programming + hyper-heuristics (TN)
  - Neural architecture search using genetic programming + hyper-heuristics (NP)
  - Multimodal learning using hyper-heuristics (NP)
- Discussion + Future research direction(NP+TN)

Introduction:
Disease Prediction

Disease Prediction

- Cancer detection
- Heart disease prediction
- Diabetes prediction
- Alzheimer’s, Parkinson
- Epilepsy
- Glaucoma

Types of Data

- Image data
  - magnetic resonance imaging
  - computed tomography scan
- Electronic medical records
  - health statistics
- Genomic and molecular
  - gene expression sequencing
  - DNA sequencing

Example Heart Disease data

Image

Electronic medical records

age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	Thal	target
52	1	0	125	212	0	1	168	0	1	2	2	3	0
53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
61	1	0	148	203	0	1	161	0	0	2	1	3	0
62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
58	0	0	100	248	0	0	122	0	1	1	0	2	1
58	1	0	114	318	0	2	140	0	4.4	0	3	1	0
55	1	0	160	289	0	0	145	1	0.8	1	1	3	0
46	1	0	120	249	0	0	144	0	0.8	2	0	3	0
54	1	0	122	286	0	0	116	1	3.2	1	2	2	0
71	0	0	112	149	0	1	125	0	1.6	1	0	2	1
43	0	0	132	341	1	0	136	1	3	1	0	3	0
34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
51	1	0	140	298	0	1	122	1	4.2	1	3	3	0
52	1	0	128	204	1	1	156	1	1	1	0	0	0
34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
51	0	2	140	308	0	0	142	0	1.5	2	1	2	1
54	1	0	124	266	0	0	109	1	2.2	1	1	3	0
50	0	1	120	244	0	1	162	0	1.1	2	0	2	1

Diagnosis Complexities

- Overlapping symptoms
- Genetic variability
- Information overload
- Shortage of expertise
- Human cognitive bias
- Data integration
- Disease progression
- Data acquisition tools
- Incomplete data or partial images

Disease Prediction Problems

Data Classification

Image Classification

Image Segmentation

References

1. Khan, A. H., Hussain, M., & Malik, M. K. (2021). ECG Images dataset of Cardiac and COVID-19 Patients. *Data in Brief*, *34*, 106762.

Introduction:
Computational Intelligence for Disease Prediction

Evolutionary Algorithms

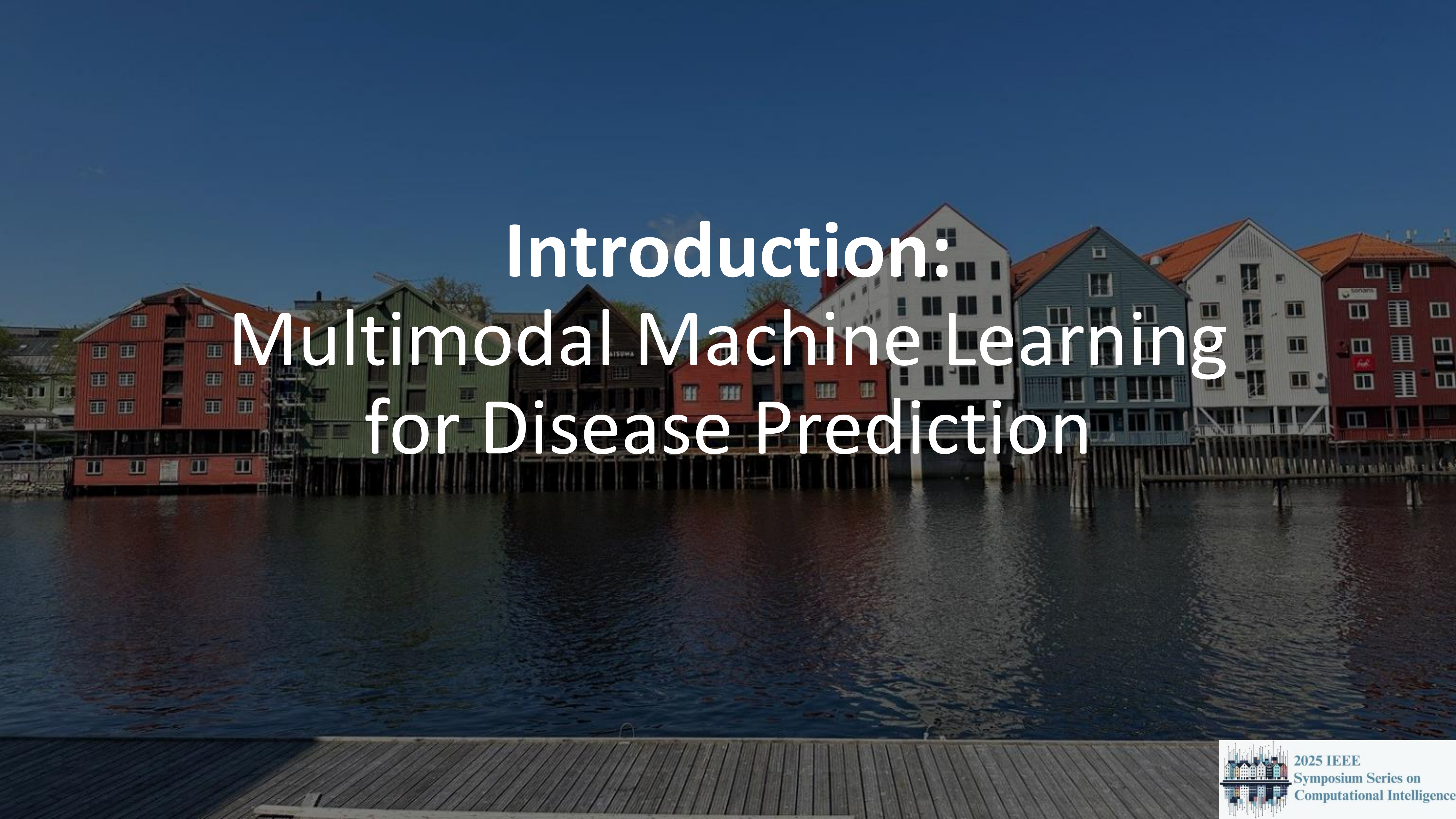
- Genetic programming
- Genetic algorithms
- Grammatical evolution
- Differential evolution
- Multi-objective evolutionary algorithms

Neural Networks

- Autoencoders
- Graph neural networks
- Convolutional neural networks
- Transformers
- Generative adversarial networks
- Ensembles

References

1. Liu, R. et al , Spatial Temporal Co-Attention Learning for Diagnosis of Mental Disorders of Resting State fMRI Data, IEEE Transactions Neural Networks and Appl, 2024.
2. Liang, J. et al, Multiobjective Optimisation Based Network Control Principles for Identifying Personalised Drug Targets with Cancer, IEEE Transactions Evolutionary Transactions, 2024.
3. Ma, H. et al, RS-MAE: Region-State Masked Autoencoder for Neuropsychiatric Disorder Classifications MRI, IEEE Transactions Neural Networks and Appl, 2024.



Introduction:
Multimodal Machine Learning for Disease Prediction

Multimodal Learning

- Learning from more than one type of data
- Types of data
  - Quantitative
  - Image
  - Audio
- Classifier for each combiner approaches
- Automating the design of multimodal classifiers

Introduction:
Automated Design

2025 IEEE
Symposium Series on
Computational Intelligence

Why automated design for computational intelligence for disease prediction?

Automated Design

- Algorithm selection
- Algorithm composition
- Algorithm configuration
- Algorithm generation

Automated Configuration

- Parameters
  - population size
  - selection method
  - application/probability rates
  - lengths and depths
  - …
- Hyper-parameters
  - Learning rate
  - Activation function
  - …
- Operators
- Neural network weights
- Neural network architecture

Automated Generation

- Loss functions
- Activation functions
- Fitness functions
- Genetic operators
- Heuristics
- Metaheuristics

Neural Networks Design

- Determining the type of neural network
- Hyper-parameters
- Determining neural network weights
- Neural network architecture

Evolutionary Algorithm Design

- Parameters
- Genetic operators
- Fitness functions
- Construction heuristics
- Control flow

References

1. Pillay, N., Qu, R. Automated Design of Machine Learning and Search Algorithms, Springer Nature Series, 2021.
2. Qu, R., Kendall, G., Pillay, N. The General Combinatorial Optimisation Problem: Towards Automated Algorithms Design. IEEE Computational Intelligence Magazine, May 2020, Vol. 15, No. 2, pp. 14-23.

Introduction:
Hyper-Heuristics for Automated Design

2025 IEEE
Symposium Series on
Computational Intelligence

Hyper-Heuristics

- Explores a heuristic space
- Generality
- Selection construction hyper-heuristics
- Selection perturbative hyper-heuristics
- Generation construction hyper-heuristics
- Generation perturbative hyper-heuristics
- Discrete/combinatorial optimization

Hyper-Heuristics - Configuration

Genetic algorithm design:

- Population size
- Number of generations
- Initial depth
- Tournament size
- Offspring depth
- Crossover probability
- Mutation probability

Example:

100,50,4,3,10,0.8,0.6

Hyper-heuristic

- Optimizes the the sequence of low-level heuristics to improve the design
- Low-level heuristics:

Increase a parameter (i_1 i_2 i_3 i_4 i_5 i_6 i_7)

Decrease a parameter (d_1 d_2 d_3 d_4 d_5 d_6 d_7)

Crossover (c)

Hyper-Heuristics - Configuration

$i_1d_3d_2ccci_2$

- Single point search
- Tabu search
- Variable neighborhood search
- Simulated annealing
- Late acceptance hill-climbing
- Genetic algorithms

References

1. Pillay, N., Qu, R. Hyper-Heuristics: Theory and Applications, Springer Nature Series, 2018.

Introduction:

Evolutionary Algorithms for Automated Design

2025 IEEE
Symposium Series on
Computational Intelligence

Genetic Algorithms vs. Genetic Programing

- Genetic algorithms and genetic programming
- Configuration
  - Genetic algorithms
- Automated composition
  - Genetic algorithms
  - Genetic programming
- Automated generation
  - Genetic programming
  - Grammar-based genetic programming
  - Grammatical evolution

Case Studies:
Evolutionary Algorithms

Evolutionary Algorithms

Automated design using grammatical evolution

Automated design using genetic algorithms

Automated Design using Genetic Algorithm

- Chromosome – GP classification algorithms.
- Genes - GP design decisions.
  - GP parameters
  - GP genetic operators
  - GP selection method
  - GP fitness functions
  - GP control flow

Automated Design Using Grammatical Evolution

- Programs – GP classification algorithms.
- Grammar - GP design decisions.
  - GP parameters
  - GP genetic operators
  - GP selection method
  - GP fitness functions
  - GP control flow
- Variable length GE individuals.
- Selection - tournament
- Genetic operators single point.

Automated Design Using a Genetic Algorithm for Image Classification

- Chromosome – image classification pipeline
- Gene
  - noise reduction – gaussian, sobel
  - image descriptors- SIFT, SURF
  - patch/size
  - intensity suppression operators
  - feature extractor - filters
  - thresholding
- Fitness function- SSIM, RMSE
- Genetic operators.

Automated Design of Image Segmentation using Differential Evolution

- DE vectors – GA algorithm design decisions.
  - Mlt thresholds
  - encoding options
  - fitness functions - Otsu , kapur's entropy.
  - quality metrics- iou, dice function
  - selection methods – tournament selection
  - genetic operators – single crossover, uniform mutation
  - GA control flow

References

1. Nyathi , T. Pillay, N Comparison of a genetic algorithm to grammatical evolution for the Automated Design of Genetic Programming Classification Algorithms 2017
2. Nyathi , T. Pillay, N Automated Design of Genetic Programming Classification Algorithms using a Genetic Algorithm. Applications of Evolutionary Computation, 2017
3. Officer, R. Genetic algorithms for the Automated denoising of medical images, Master’s Dissertation, Department of Computer Science, University of Pretoria, South Africa, 2024.
4. Nyathi, T. Automated Design of Multilevel Thresholding using Differential Evolution. International Conference on Artificial Intelligence and Soft Computing, 2024

Case Studies:
Neural Architecture Search

2025 IEEE
Symposium Series on
Computational Intelligence

Neural Architecture Search

Automated design
using genetic programming

Automated design
using hyper-heuristics

Automated Design Using Genetic Programming

- Program – instructions to create neural networks
- Function set
  - AddLayer – combines layers
  - AdjustLayer – parameter change for layer
  - AdjustNet – parameter change for neural network
  - If-Then-Else
  - CheckLayer –layer existence check
  - SwapLayers – swaps layer order
  - ContainsParams – check for parameter existence
  - Comb2 – combines 2 trees
  - Comb3 – combines 3 trees

Automated Design Using Genetic Programming

- Terminal set
  - Shape – layer parameters/options
  - LFunc– activation function
  - AdjustNet – parameter change for neural network
  - NOpt - optimiser
  - NLRate – learning rate
  - NSize– number of layers
  - C – ephemeral constant
- Selection method: tournament
- Fitness functions:
  - Maximise the model accuracy
  - Minimise training time
  - Minimise size of the model

Automated Design Using Iterative Structure-Based Genetic Programming

- Both fitness and structure
- Structure controls exploration and exploitation
- Global level search is used for exploration
- Local level search is used for exploitation
- Similarity indexes
  - global
  - local

Automated Design Using Iterative Structure-Based Genetic Programming

- Image classification
- Results
  - GP performed that GA
  - Iterative structure-based GP performed better than GA
  - Iterative structure-based GP performed better than GP

Automated Design Using Hyper-Heuristics

- NAS operators to create architecture
  - Insert(i) – inserts layer
  - Swap(s) – swaps layers
  - Replace(r) – replace layer
  - Insert connection (c) – connects layers
  - Delete connection (d)– deletes a connection
  - Swap connection (a) – swap two connections
- Example sequence: *iassrdc*
- Single point hyper-heuristic
  - Heuristic selector – choice function
  - Move acceptance – acceptance improvement limited target acceptance (AILTA)

Automated Design Using Hyper-Heuristics

- NAS operators to create architecture
  - Replace – replaces an operator
  - Swap – swaps operators
  - Insert – inserts an operator
  - Remove – connects layers
  - Destroy create – deletes a connection
  - One-point crossover – one point swap
  - Two-point crossover – two point swap
  - Uniform crossover – segment swap

References

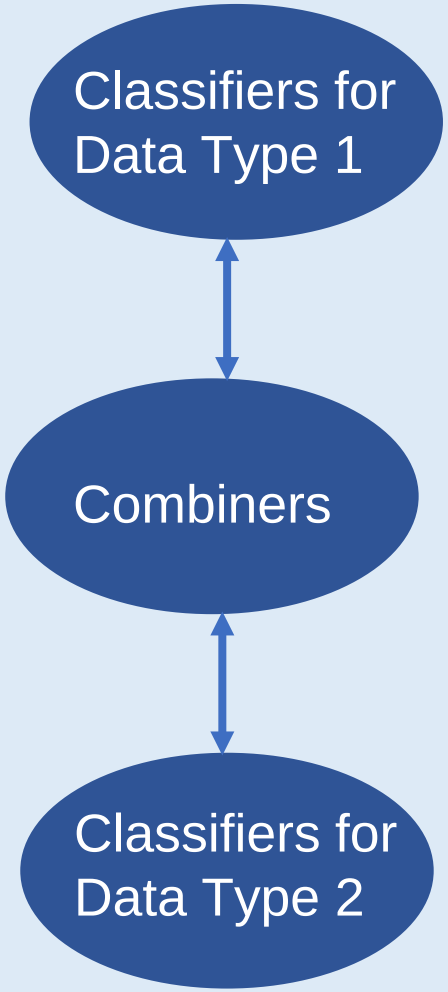
1. Kapoor, R., Pillay, N., A Genetic Programming Approach to the Automated Design of CNN models for Image Classification and Video Shorts Creation. Genetic Programming and Evolvable Machines, 25, 10, 2024, https://doi.org/10.1007/s10710-024-09483-5
2. de Clercq, J. Hyper-Heuristics for Neural Architecture Search, Master's Dissertation, Department of Computer Science, University of Pretoria, 2025.

Case Studies: Multimodal Machine Learning

Multimodal Machine Learning

Automated design using hyper-heuristics

Multimodal Approach – Two Types of Data



Multimodal Approach: Overview

- Example data
  - Chest X-rays (image)
  - Patient information – 10 attributes
- Image classifiers
  - ResNet
  - VGGNet
  - AlexNet
- Data classifiers
  - Multi-layer perceptron
  - Random forest
  - Support vector machines
- Combiners
  - K-Nearest neighbour
  - Multi-layer perceptron
  - Random forest

Multimodal Approach: Hyper-Heuristic

- Genetic algorithms hyper-heuristic
- Chromosome:

  Image classifier;combiner;data classifier
- Fitness – function of accuracy, precision and recall
- Selection method: tournament selection
- Genetic operators: crossover and mutation
- Performance
  - Multimodal performed better than unimodal
  - Automated multimodal performed better than manual multimodal

References

1. Marais, G.N. Automated Multimodal Machine Learning, Master's Dissertation, Department of Computer Science, University of Pretoria, South Africa, 2024.

Thank you

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Postdoctoral Position

2 years
R 370 000 p.a.
Machine learning and optimization
nelishia.pillay@up.ac.za

Discussion and Future Research Directions