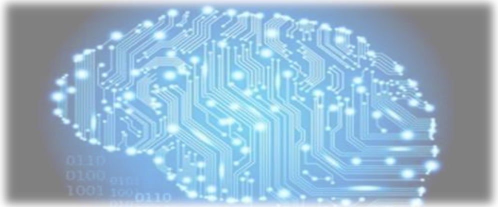
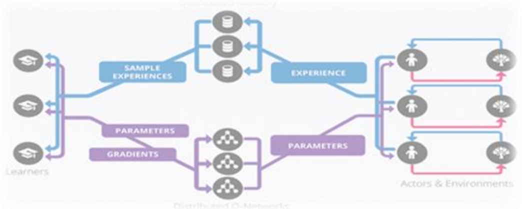




Music Recommendation

AutoML 연구실
학부연구생 조성현
20.12.31

Contents

A many-object feature selection for multi-label classification

- Introduction
- Main Contribution & Background
- Background
- Proposed Method
- Experiment
- Conclusion

Introduction

Introduction

- This study presents many-objective optimization based multi-label feature selection algorithm(MMFS).
- To improve diversity and convergence of NSGA III, they propose an improved NSGA III algorithm with two archives.
- New crossover and mutation operators for feature selection are designed to improve the exploration capability, and the influence of the selection threshold on feature scale and multi-label classification performance in real number coding is studied.
- Commonly used evolutionary algorithms are Genetic Algorithm(GA), Particle Swarm Optimization(PSO), Differential Evolution(DE), Genetic Programming(GP), Ant Colony Optimization(ACO).
- In particular, Evolutionary Algorithms are also effective means of solving multi-objective problems (MOPs).
- Difference between multi-class classification and multi-label classification

Multi-class classification	each sample is related with one label, multiple mutually exclusive classes under the label
Multi-label classification	each sample associates with multiple labels, and are not mutually exclusive

Main Contribution & Background

- ## The main contribution of this study
  - They proposed an improved NSGA III algorithm which uses two archive sets to balance the exploration and development capabilities.
  - To enable the proposed algorithm suitable for multi-label feature selection, the uniform mutation and uniform crossover operator based on the label frequency difference are proposed.
  - In the archive set CA, the selection threshold is adaptively determined with the evolution of the population to balance the feature scale and the classification performance.
- ## Background
  - Many-objective Optimization
  - NSGA III
  - Multi-label learning

Proposed Method

• The proposed method

1. Objective function

- Ranking Loss $$RL(f,\mathfrak{F}) = \frac{1}{n}\sum_{i=1}^{n}\{\frac{1}{|y_i||\overline{y_i}|}|(k,l) \in (y_i * \overline{y_i}), s.t. f(x_i,k) \leq f(x_i,l)\}$$

- Average Precision $$AP(f,\mathfrak{F}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{|y_i|}\sum_{y\in y_i}\frac{|\{y'|rank_f(x,y')rank_f(x_i,y), y' \in y_i\}|}{rank_f(x_i,y)}$$

- Macro-F1 $$MaF(f,\mathfrak{F}) = \frac{1}{q}\sum_{j=1}^{q}\frac{2\sum_{i=1}^{n} y_{ij} f_j(x_i)}{\sum_{i=1}^{n} y_{ij} + \sum_{i=1}^{n} f_j(x_i)}$$

- Coverage $$CV(f,\mathfrak{F}) = \frac{1}{q}\{\frac{1}{n}\sum_{i=1}^{n} max_{y\in y_i} rank_f(x_i,y) - 1\}$$

- Hamming loss $$HL(f,\mathfrak{F}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{q}|f(x_i) \oplus y_i|$$

- Micro-F1 $$MiF(f,\mathfrak{F}) = \frac{2\sum_{i=1}^{n}|f(x_i) \cap y_i|}{\sum_{i=1}^{n}|y_i| + \sum_{i=1}^{n}|f(x_i)|}$$

Proposed Method

- ## The proposed method

2. The proposed many-objective algorithm T-NSGA III

    - CA : mainly focuses on local search to achieve rapid convergence
    - DA : mainly maintains global search to improve the diversity of the population

3. The differences from other algorithms

    (1) Two-Ach2

    - The update methods of CA and DA are different.
    - Different problems lead to different operators.
    - The maintenance of diversity is different in DA.

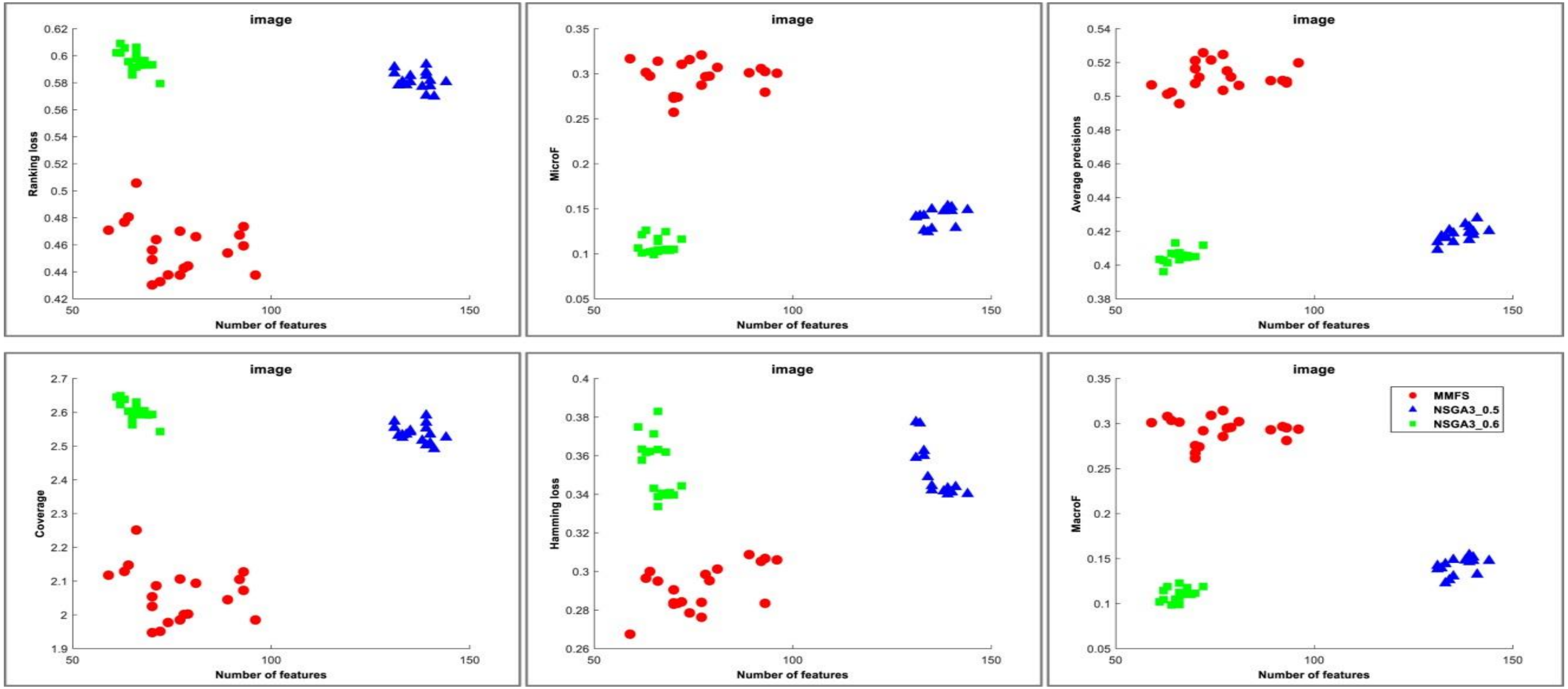
    (2) BNSGA3

    - The optimization objectives are different.
    - Both BNSGA3 and MMFS have improved the NSGA3 algorithm.
    - The maintenance of non-dominated archive sets is different. BNSGA3 uses the traditional NSGA3 method of maintaining archive sets. MMFS combines the methods of NSGA3 and Two-Ach2 and sets two archive sets to balance diversity and convergence.

Experiment

- **Experiment**

Dataset	#Features	#Labels	#Instances	#Training	#Testing	Domain
Image	294	5	2000	1200	800	scene



Experiment

- ## Experiment(Comparison Results)

	MMFS	NSGA3_0.5	NSGA3_0.6	BNSGA3
Ranking Loss	0.4327	0.5705	0.5858	0.4572
Average Precision	0.5258	0.4226	0.4131	0.5044
MicroF	0.3105	0.1477	0.0092	0.2554
Coverage	1.9513	2.5026	2.5627	2.0588
Hamming Loss	0.2843	0.3431	0.3431	0.2858
MacroF	0.2921	0.1463	0.1050	0.2589

Conclusion

- ## Conclusion
  - They propose a many-objective optimization-based multi-label feature selection method. The optimization objectives of the algorithm include six multi-label evaluation criteria and the feature subset size.
  - They adaptively determine the selection threshold in CA to explore the feature space as much as possible. In the two archives, the non-dominated solutions are selected according to the NDS strategy of NSGA III in DA. And they increase the selection pressure on the non-dominated solutions in terms of the convergence information in CA.
  - The positions of excellent individuals can be exchanged through the crossover operation. In the experiment, eleven multi-label datasets are chosen for comparing MMFS with the multi-label feature selection algorithms based on optimized and non-optimized algorithms.
  - In general, MMFS can effectively reduce the data dimensionality and obtain better classification results. However, the calculation cost of the wrapper feature selection algorithms is relatively large.