Text Classification using Naïve Bayes:

1. Data:

The data are attachment with these folders:

Training Data:

    Folder: All-in-many/

    Description: 117 articles (titles and abstract) of Volume 18 of Evolutionary Intelligence Journal.

Training Data Labeling:

    Folder: All-in-many_classification/

    Description: Four classes available :

        1st **class :** Metaheuristics

        2nd **class :** Machine & Deep Learning

        3rd **class :** Combination of Metaheuristics & Machine/Deep Learning

        4th **class :** Others

Testing Data:

    Folder: All-in-many_classification_testing/

    Description: 41 articles (titles and abstract) of Volumes 17 and 14 of Evolutionary Intelligence Journal with labels.

2. Classifying a document using Naive Bayes :

- The **final equation** used to **identify** the **most probable class** $\hat{c}$ for a **test document** d in the **Naive Bayes classifier** is as follows:

$$\hat{c} = \arg\max_{c \in C} P(c) \prod_{i \in \text{POSITIONS}} P(w_i|c)$$

$$\hat{c} = \arg\max_{c \in C} P(c) + \sum_{i \in \text{POSITIONS}} P(w_i|c)$$

POSITIONS: the set of word positions in the test document d.

3. Implementation :

Classifying document no. 2 from the test dataset.

Part 5 \ Text Classification

Naive Bayes classifier

🔘 Content of one volume   ⚪ Content of all articles

Volume n°:

| 18 | − | + |

Lemmatization and Stemming:

Porter Stemmer ⌄

☑ Normalization

Visualization

Training

Testing

Test text

Improved task scheduling in heterogeneous distributed systems using intelligent greedy harris hawk optimization algorithm.
In the context of heterogeneous computer systems, the utilization of task scheduling algorithms is crucial for enhancing system performance. A major challenge in the development of efficient scheduling algorithms for such systems lies in reducing execution time. However, achieving this objective is not a straightforward process. Due to the complexity of these problems, exact algorithms are not feasible for solving them, primarily because of the vast search space involved. Consequently, approximate algorithms become necessary. Among these, meta-heuristic algorithms are notable for their ability to explore large search spaces and generate suitable solutions. Nevertheless, these algorithms are not flawless as they may become trapped in local optima and failing to scan the entire sample space. To address task scheduling problems in heterogeneous computer systems, this research proposes the Intelligent Harris Hawk Optimization algorithm. The algorithm was employed to tackle the task scheduling problem in heterogeneous systems, while also considering common issues associated with such algorithms, including early convergence and limited exploration capability during the exploration phase. To overcome these challenges, the exploration phase was separated from the exploitation phase. This approach aimed to

Test article n°:

| 2 | − | + |

Classification result: class n°1, **correct classification:** class n°1

Accuracy: 56.10%