

Zero-Shot Learning

Preprocessing the Semantic Space in Zero-Shot Learning

Juan José Herrera Aranda, Francisco Herrera, Guillermo Gomez Trenado, Isaac Triguero

Department of Computer Science and Artificial Intelligence, University of Granada, Spain

Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI)

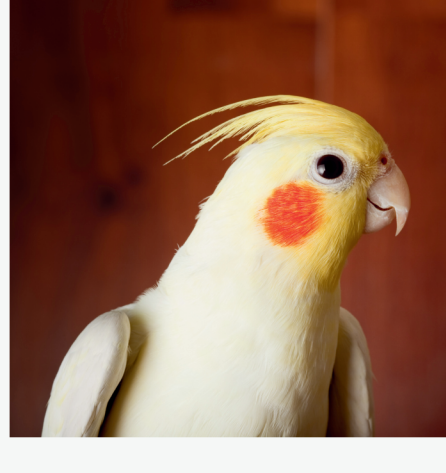


juanjoa@ugr.es



Introduction

Semantic Spaces

	Love Bird Colourful: Yes Amphibian: No Mammal: No Bird: Yes Fish: No Wings: Yes Water: No Diet: Herbivorous		Cockatiel Colourful: No Amphibian: No Mammal: No Bird: No Fish: No Wings: Yes Water: No Diet: Herbivorous		Penguin Colourful: No Amphibian: No Mammal: No Bird: Yes Fish: No Wings: Yes Water: Yes Diet: Carnivorous
---	--	--	--	---	--

Limitations

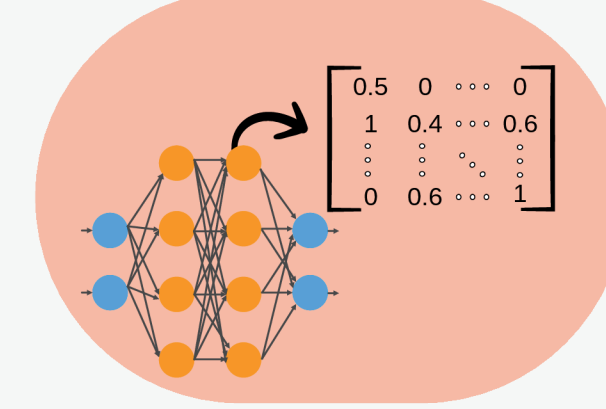
Engineered

Humans Errors
Human Subjectivity
Interdependency



Learned

Internet is biased
Redundancy
Overfitting



Configurations

- ⚠ Inductive vs Transductive
- ⚠ Generalized Zero-Shot Learning

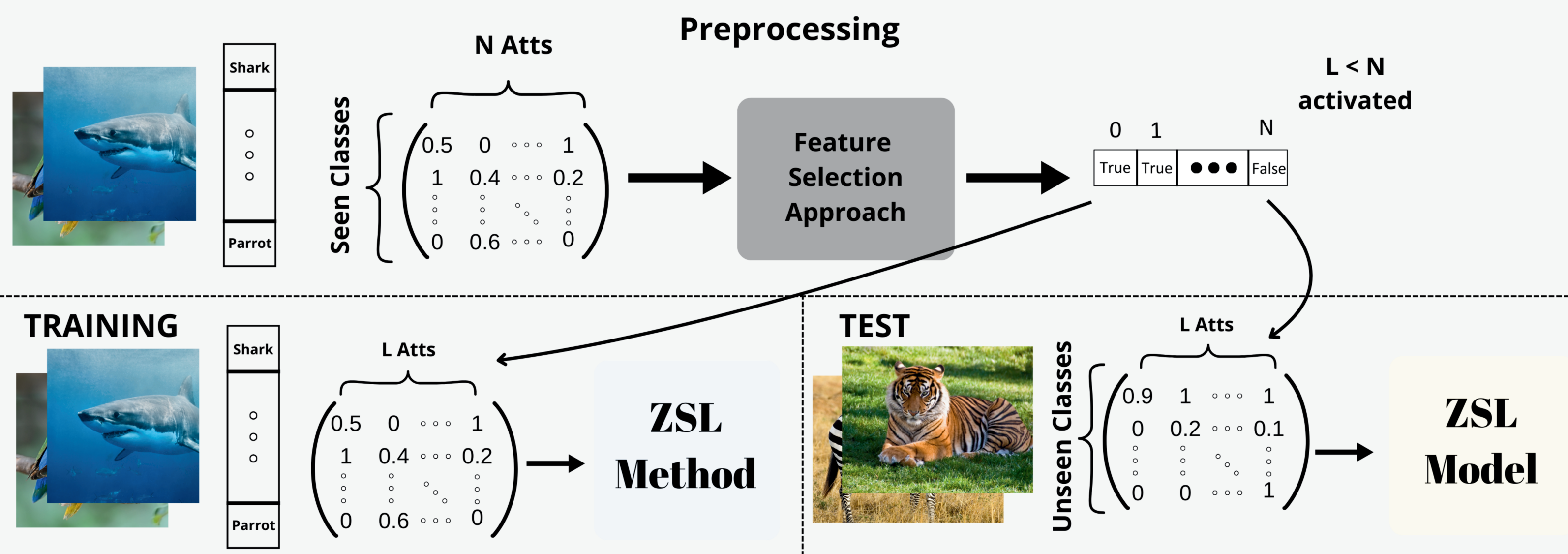
Gaps

- Semantic spaces are not preprocessed
- No robust evaluation protocols exist
- Overlooking the inductive setting

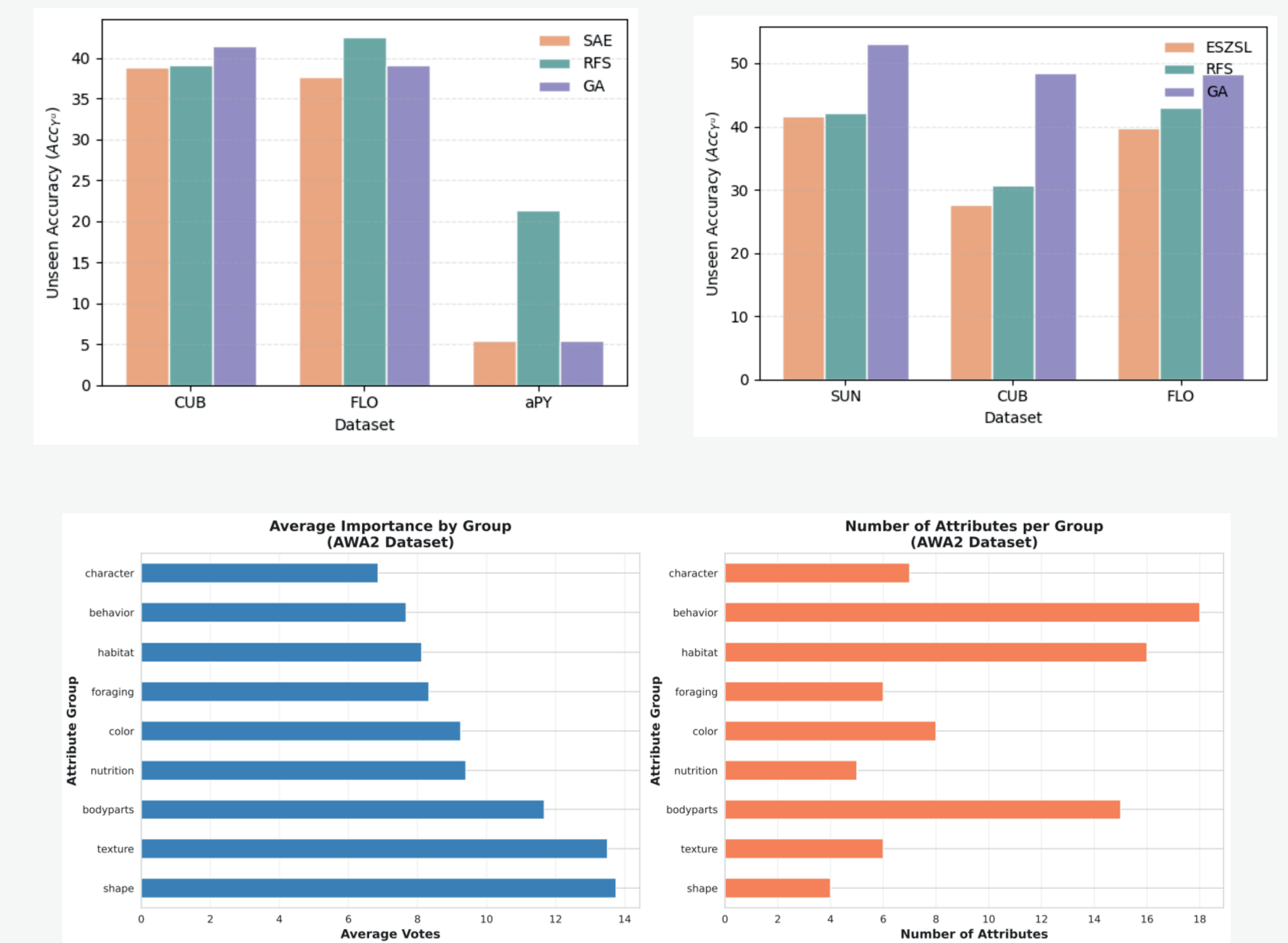
Our Research: Preprocessing the semantic space

Global Overview

We aim to design a methodology to perform feature selection in the semantic space to find an optimal subset of attributes to enhance the model performance for the inductive setting. Our purpose is to remove noisy, not to achieve a dimensional reduction.

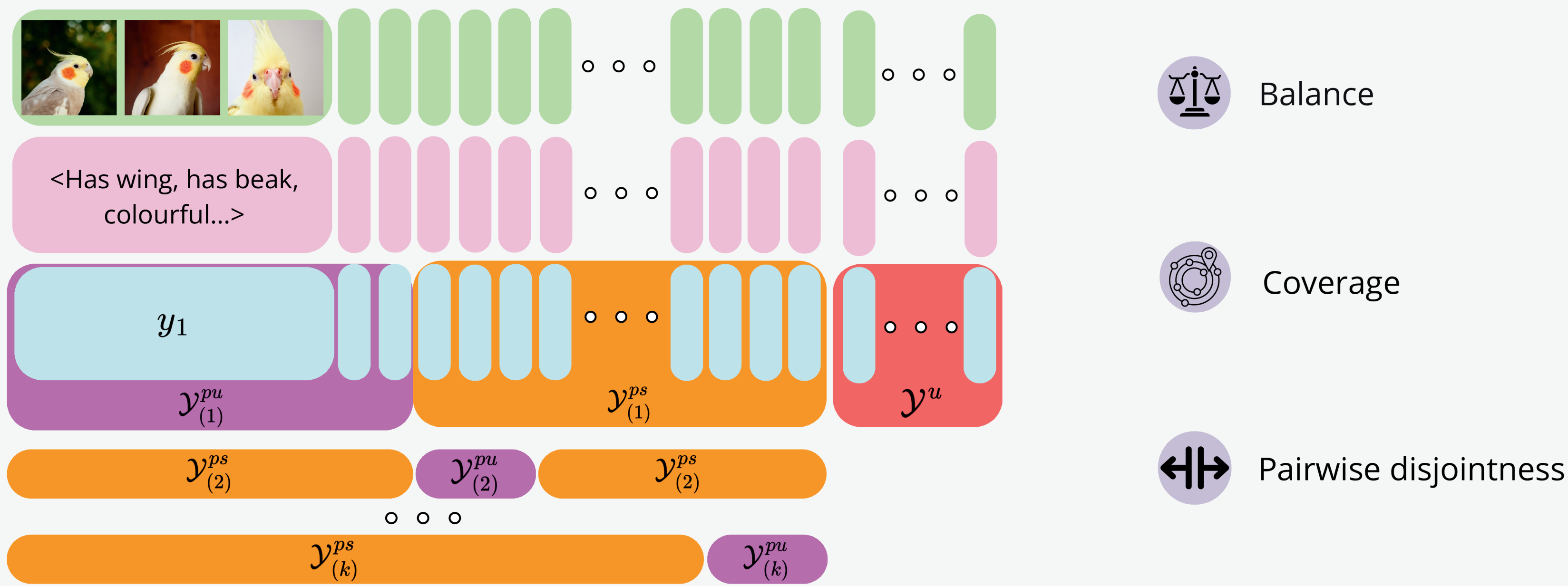


Results



Class Stratified Cross Validation Data Partition

We propose a class-stratified partitioning protocol that simulates unseen classes in an inductive setting. This enables semantic feature selection without accessing unseen semantic information, ensuring a realistic zero-shot evaluation.



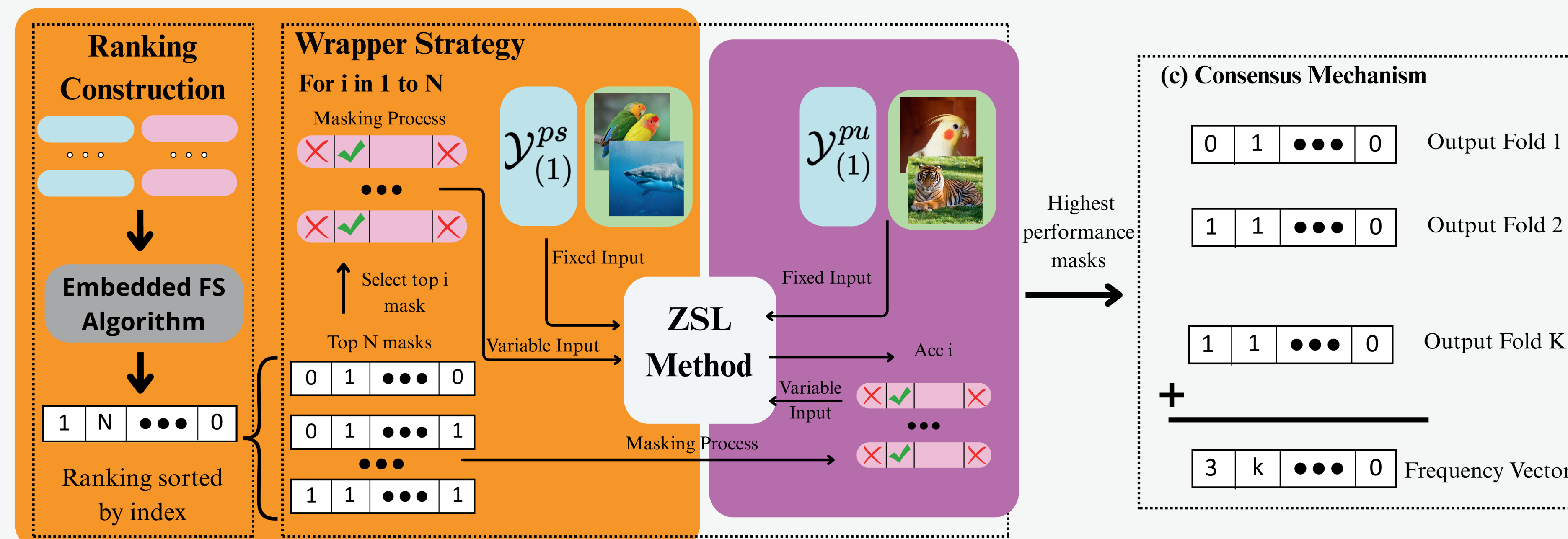
Feature Selection using Genetic Algorithms - GAZeLa

Genetic algorithms that optimizes attributes subsets using unseen-class the previous evaluation protocol. The genetic operators are standard

- 🔍 Global Search over attributes subsets
- ⌚ Might be time-consuming
- 📊 Not individual rankings and simpler than RFS

Ranking-based Feature Embedded-based Selection Method - RFS

Ranking-based Feature Selection (RFS) builds an attribute ranking using embedded feature selection methods on seen-class semantic prototypes. Attribute subsets are evaluated via a class-stratified wrapper on pseudo-unseen classes and consolidated across folds using a stability threshold.



- 🔍 Sensitive to the thresholds
- 📄 Interpretable attributes subsets
- 🔍 Search is limited (Greedy)
- ⌚ Less efficient than GAZeLa

Bibliography

Herrera-Aranda, J. J., Gomez-Trenado, G., Herrera, F., & Triguero, I. (2025). Semantic-Inductive Attribute Selection for Zero-Shot Learning. arXiv preprint arXiv:2510.03260.

Herrera-Aranda, J. J., Herrera, F., & Triguero, I. (2025, June). A First Approach to Refine Semantic Spaces in Zero-Shot Learning with a Genetic Algorithm. In 2025 IEEE Congress on Evolutionary Computation (CEC) (pp. 1-4). IEEE.

Herrera-Aranda, J. J., Herrera, F., & Triguero, I. (2024, October). A preliminary study on preprocessing the semantic space in zero-shot learning. In International Conference on Hybrid Artificial Intelligence Systems (pp. 177-189). Cham: Springer Nature Switzerland.