Université de technologie de Compiègne – Thesis proposal

Part 1: Scientific sheet	
Thesis proposal title	**Robust ensembling in set-valued classification**
PhD grant	Allocation MESR
Host laboratory	CID team, Heudiasyc lab, UMR CNRS 7253
Thesis supervisors	Thierry DENOEUX (Professor at UTC and senior member of IUF) Vu-Linh NGUYEN (Junior professor (CPJ), holder of the "trustworthy AI chair")
Scientific domain(s)	Computer science
Research work	***Context***: Ensemble learning is a common practice to achieve robustness and improved generalization in machine learning [7]. In classification tasks, explored in this thesis project, ensemble learning typically consists of two phases: *Generation* and *aggregation*. The former trains a set of predictive models that seek a diversity-accurateness trade-off. During the latter, for each instance, the predictions of predictive models are aggregated into the final prediction. Although conventional ensemble learning is advantageous in various application domains, the assumption that predictive models always produce singleton predictions, which should be, in turn, aggregated into a final singleton prediction, can lead to challenging problems in the presence of *low coverage* in ensemble aggregation. By low coverage, we refer to the scenario in which, for the current instance, only a minority or none of the predictive models provide the correct prediction. In fact, low coverage is difficult to handle given the voting nature of the aggregation phase [7], and may lead to an unsatisfactory correctness level of the final predictions. Low coverage is an issue even in the simple *multiclass classification* (MCC) setting, where one has to predict a single class variable. This issue is significantly amplified in the following more challenging classification tasks that require large data sets: *multilabel classification* (MLC), that is, predicting multiple binary variables simultaneously [5] and *multidimensional classification* (MDC), that is, predicting multiple categorical variables simultaneously [6]. ***Research questions***. Predictive models are usually constructed in two steps: for each instance, a scoring vector reflecting the relevance of the possible outcomes is first predicted and used to construct a ranking over the possible outcomes, and the top-ranked outcome is predicted. Therefore, a diversity-accurateness trade-off might be one of the most reasonable targets one can expect from the ensemble generation phase. This is rooted in the fact that diversity and accuracy in ensemble generation are somewhat bidirectional. Even if a diversity-accuracy trade-off is ensured to some extent, it is unclear how it helps to tackle low coverage. This is because low coverage typically appears at the individual level, whereas diversity and accuracy are often assessed at the population level. Therefore, we propose to construct a new ensemble learning framework, especially designed to tackle the low coverage issue, aiming at both coverage and diversity: • *Generation*: Train a set of set-valued classifiers [4, 5, 7] that seek both coverage, i.e., the true outcome is included in the set-valued predictions, and diversity. • *Aggregation*: For each instance, the set-valued predictions of predictive models are aggregated into the final prediction, which can be either a single outcome or a set of outcomes. Can we eventually achieve both coverage and diversity? If so, how could these improvements translate into an enhanced predictive performance of the ensemble?

Research directions (RDs). We plan to first focus on the MCC and MLC settings before tackling the more challenging MDC setting.

Ensemble generation (RD1). This phase aims to gain coverage and diversity.

- *Homogeneous ensembles*: It is not hard to prove that many probabilistic set-valued classifiers (PSVCs) [5, 7] are coverage monotone, i.e., the coverage should not decrease when increasing the reward of being correctly cautious. This makes PSVCs promising candidates for gaining controllable levels of coverage. We shall study how to gain their diversity via different data and model perturbation strategies during the ensemble generation phase: bagging, noise injection, random initialization, and so on.
- *Heterogeneous ensembles*: Allowing predictive models to be heterogeneous, such as mixtures of probabilistic, credal, and evidential classifiers [4, 5, 7] can be a promising way to gain diversity, while still allowing chances to gain high coverage. The high coverage may be given by allowing the predictive models to access the entire and exact training data. The diversity may be naturally given by different backbone uncertainty theories of heterogeneous classifiers. We shall investigate the diversity and coverage of different mixtures of uncertainty theory-based classifiers.

Ensemble generation (RD2). This phase aims to reliably translate the coverage and diversity into the accurateness of the ensemble.

- *Evidential approach*: Once the coverage and diversity are ensured, it is natural to interpret set-valued predictions as incomplete data (with noise) to construct a Dempster-Shafer mass function from the possible set-valued predictions, and apply evidential decision rules [1, 2, 3] to derive the final prediction, which can be either a singleton or set-valued. We will aim for the scalability of this theoretically sound proposal, especially in MLC and MDC.
- *Probabilistic approach*: It is known that the prediction phase of probabilistic classifiers can be customized to gain accurateness concerning multiple evaluation metrics without requiring retraining the predictive model [5, 6]. We will generalize this idea to ensembles of set-valued classifiers, where the final prediction is a prediction within the feasible domain that optimizes the expected distance/metric to the set-valued predictions. The feasible domain can be customized to produce either a singleton or set-valued prediction. We will gain scalability by formulating and solving optimization problems.

[1] T. Denoeux. A k-nearest neighbor classification rule based on Dempster-Shafer theory. *IEEE transactions on systems, man, and cybernetics*, 25(5):804–813, 1995.

[2] T. Denoeux. A neural network classifier based on dempster-shafer theory. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 30(2): 131–150, 2000.

[3] T. Denœux. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, 172(2-3):234–264, 2008.

[4] L. Ma and T. Denoeux. Making set-valued predictions in evidential classification: A comparison of different approaches. In *ISIPTA*, pages 276–285, 2019.

[5] V.-L. Nguyen and E. Hüllermeier. Multilabel classification with partial abstention: Bayes-optimal prediction under label independence. *Journal of Artificial Intelligence Research*, 72:613–665, 2021.

[6] V.-L. Nguyen, Y. Yang, and C. P. de Campos. Probabilistic multi-dimensional classification. In *UAI*, pages 1522–1533, 2023.

[7] V.-L. Nguyen, H. Zhang, and S. Destercke. Credal ensembling in multi-class classification. *Machine Learning*, 114(1):1–62, 2025.

Keywords	Machine learning, uncertainty quantification, information fusion, belief functions, evidence theory, multilabel classification

Part 2: Job description	
Starting time	01/10/2026
Duration	36 months
Research laboratory	Heudiasyc UMR 7253, Université de Technologie de Compiègne
Requirements	Master 2 or engineer in computer science, good programming skills (Python, PyTorch, TensorFlow, ...) and/or a strong background in mathematics.
Additional missions	Teaching are possible, but not mandatory
Material resources	Shared office, laptop, access to the laboratory's GPU servers and the Jean Zay supercomputer installed at IDRIS, as well as to the laboratory's platforms, ...
Human resources	The laboratory hosts 33 permanent faculty, 6 CNRS researchers, 14 technicians, engineers, and administrative staff in support of research, 46 PhD students, 5 post-doctoral students, and 12 post-docs/engineers on fixed-term contracts.
Working conditions	The supervision team proposes a two-stage pedagogical project. During the first stage, the student will be guided in the choice of results, algorithmic solutions, existing software packages, on which to base their own results, software, and experimental protocols, and to communicate them through scientific articles. Once the necessary knowledge and skills have been acquired, the student will be invited to tackle more difficult problems in collaboration with supervisors and collaborators, to develop the research skills to work both independently and collaboratively. The candidate will be funded by Allocation MESR, and will get financial support for travel (conferences, workshops, summer schools, ...).
Collaborations	To be determined based on the development of the student.
Contact	Applications should include a letter of motivation, a curriculum vitae, and contact information of at least two references. Applications and questions should be sent to: • Thierry Denœux (thierry.denoeux@hds.utc.fr) • Vu-Linh Nguyen (vu-linh.nguyen@hds.utc.fr) Applications should be submitted before **May 10, 2026**.