Master 2 Internship Proposal

From Predictive Uncertainty to Domain Adaptation: When Does Uncertainty Reveal a Change of Domain?

Team/project, laboratory, city, country: LITIS, Université de Rouen Normandie / INSA Rouen Normandie; Rouen, France.
Duration/expected start: 6 months/March 2027.
Internship supervisors: Alessandro Leite (aleite@insa-rouen.fr); Fannia Pacheco (fannia.pacheco@univ-rouen.fr); Marianne Clausel (marianne.clausel@univ-lorraine.fr).

General introduction to the topic

Machine-learning models are commonly trained assuming that deployment data follow the training distribution. In practice, changes in populations, sensors, environments, or acquisition conditions produce a mismatch between a *source* domain P_S and a *target* domain P_T. Domain adaptation seeks to transfer predictive knowledge across such shifts, while uncertainty quantification (UQ) estimates the reliability of a model's predictions. In particular, a source-trained model may be more uncertain about target inputs that are poorly represented in its training data. Yet domain shift and uncertainty are not equivalent: a shift need not increase prediction difficulty, and uncertainty can be high even without a shift. This internship will study when predictive uncertainty can meaningfully reveal, measure, or be corrected for a change of domain.

Objective of the internship

The central question is: *under which assumptions can predictive uncertainty be characterized, calibrated, or explained by source–target discrepancy?* The project will investigate a conceptual decomposition $U_T(x) = U_{\text{in}}(x) + U_{\text{shift}}(x)$ and whether the shift-induced term relates to a local discrepancy $D_x(P_S, P_T)$. Specifically, the student will:

Characterize effects of shift: compare covariate, label, conditional/concept and support shifts, including mixtures, using predictive entropy, deep-ensemble variance, epistemic indicators, residual-based uncertainty and conformal set size.

Connect uncertainty with adaptation: identify when uncertainty reflects local source–target mismatch rather than intrinsic data noise or model limitations; analyze relationships among discrepancy, uncertainty, and target prediction error, including counterexamples.

Develop domain-aware uncertainty measures: assess $U_{\text{DA}}(x) = g(U_{\text{base}}(x), D_x(P_S, P_T))$ using density ratios, domain-classifier scores, kernel discrepancies, optimal transport, and local-density or representation-space distances.

Evaluate adaptation strategies: compare adapting the predictor, the uncertainty quantifier, or both; consider importance weighting and weighted conformal prediction for target-domain calibration.

Validate experimentally: construct controlled synthetic benchmarks with increasing shift complexity, then test on real domain-shift datasets; deliver reproducible code, a comparative evaluation, and an internship report.

Expected ability of the student

Master's 2 or final-year engineering level; strong foundations in machine learning and probability/statistics; ability to implement reproducible experiments and critically analyze results. Familiarity with uncertainty quantification, transfer/domain adaptation, or probabilistic learning is appreciated but not required. Interest in both theoretical reasoning and rigorous empirical research is important.

Selected bibliographic references

[1] S. Ben-David et al., "A Theory of Learning from Different Domains," *Machine Learning*, 79, 2010.

[2] Y. Gal and Z. Ghahramani, "Dropout as a Bayesian Approximation," *ICML*, 2016.

[3] B. Lakshminarayanan et al., "Simple and Scalable Predictive Uncertainty Estimation Using Deep Ensembles," *NeurIPS*, 2017.

[4] R. J. Tibshirani et al., "Conformal Prediction Under Covariate Shift," *NeurIPS*, 2019.

[5] A. N. Angelopoulos and S. Bates, "Conformal Prediction: A Gentle Introduction," *Foundations and Trends in Machine Learning*, 16(4), 2023.