Laboratoire d'Informatique et d'Automatique pour les Systèmes

Master 2 Internship Proposal

Title: On Understanding and Quantifying Uncertainty in Machine Learning Methods

Laboratory: LIAS/ENSMA

Encadrant(s): Prof. Allel HADJALI (allel.hadjali@ensma.fr)
Prof. Ramdane MAAMRI (ramdane.maamri@univ-constantine2.dz)

Key Words: Machine Learning, Uncertainty, Quantification, Predictive Uncertainty, Predictions Algorithms.

Subject description

Most of the real physical system and everyday scenarios include uncertainty. This is the case for medical diagnosis, weather forecasting, autonomous control of drones and self driving cars and so on. In the literature, two types of uncertainty are distinguished: aleatoric uncertainty (data uncertainty) denotes the one that is inherent to the data, e.g., noise in measurements or natural variability of the inputs, and epistemic uncertainty (model or knowledge uncertainty) related to the model and due to lack of knowledge. Aleatoric uncertainty is irreducible with further data while epistemic uncertainty is reducible given additional information. Measuring the uncertainty is important, so as to support the user in the action to take. For example, when an anomaly is detected, with weak confidence level, another source of information should be added (image, human intervention, etc.) before planning intervention actions. More generally, quantification of the prediction uncertainty allows to trust or not predictions. In fact, incorrect overconfident predictions can be harmful and lead to erroneous decision/prediction.

Quantifying the uncertainty of Machine Learning models plays an important role in making predictions that are more reliable. Providing uncertainty estimates is not only important for safe decision-making in high-risk fields, but it is also crucial in fields where the data sources are highly inhomogeneous and labeled data is rare, such as in remote sensing. Also for fields where uncertainties form a crucial part of the learning techniques, such as for active learning or reinforcement learning, uncertainty estimates are highly important.

The objective of this work is twofold:

- First, provide a comprehensive and complete survey about uncertainty quantification in Machine Learning algorithms.
- Then, lay the first foundations for a robust method to evaluate uncertainty for Machine Learning algorithm predictions. First experiment to validate the method will be carried out as well.

Bibliographie

- E. Hüllermeier, W. Waegeman, Aleatoric and epistemic uncertainty in machine learning: A tutorial introduction, 2019, arXiv preprint arXiv:1910.09457.

- Xinlei Zhou, Han Liu, Farhad Pourpanah, Tieyong Zeng, Xizhao Wang, A survey on epistemic (model) uncertainty in supervised learning: Recent advances and applications, Neurocomputing, Volume 489, 2022, Pages 449-465.
- Moloud Abdar, Farhad Pourpanah, Sadiq Hussain, Dana Rezazadegan, Li Liu, Mohammad Ghavamzadeh, Paul Fieguth, Xiaochun Cao, Abbas Khosravi, U. Rajendra Acharya, Vladimir Makarenkov, and Saeid Nahavandi, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," Information Fusion, vol. 76, pp. 243–297, 2021.
- Y. Gal, Uncertainty in Deep Learning, Vol. 1 (Ph.D. thesis), (3) University of Cambridge, 2016.
- Rußwurm M, Ali M, Zhu XX, Gal Y, Körner M (2020) Model and data uncertainty for satellite time series forecasting with deep recurrent models. In: IEEE international geoscience and remote sensing symposium, IEEE, pp 7025–7028.
- Gawlikowski, J., Tassi, C.R.N., Ali, M. *et al.* A survey of uncertainty in deep neural networks. *Artif Intell Rev* **56** (Suppl 1), 1513–1589 (2023).

LIAS / ENSMA - Téléport 2 - 1 avenue Clément Ader - BP 40109
86961 Futuroscope Chasseneuil Cedex – France
Tél : +33 (0)549498063 - Fax : +33 (0)549498064