ABSTRACT

Advances in artificial intelligence (AI) are rapidly transforming our world, with systems now matching or surpassing human capabilities in areas ranging from game-playing to scientific discovery. Much of this progress traces back to machine learning (ML), particularly deep learning, and its ability to uncover meaningful patterns and representations in data. However, true intelligence in AI demands more than raw predictive power; it requires a principled approach to making decisions under uncertainty. This highlights the necessity of probabilistic ML, which offers a systematic framework for reasoning about the unknown in ML models through probability theory and Bayesian inference.

Gaussian processes (GPs) stand out as a quintessential probabilistic model, known for their flexibility, data efficiency, and well-calibrated uncertainty estimates. GPs are integral to many sequential decision-making algorithms, notably Bayesian optimisation (BO), which has emerged as an indispensable tool for optimising expensive and complex black-box objective functions. While considerable efforts have been made to improve GP scalability, performance gaps persist in practice when compared against neural networks (NNs) largely due to their lack of representation learning capabilities. Along with other natural deficiencies of GPs, this limitation has hampered the capacity of BO to address critical real-world optimisation challenges.

This thesis aims to unlock the potential of deep learning within probabilistic methods and reciprocally lend probabilistic perspectives to deep learning. The key contributions are: (1) Extending orthogonally-decoupled sparse GP approximations to incorporate nonlinear NN activation layers as inter-domain features, mitigating the limitations of prior work and bringing predictive performance closer to NNs. (2) Framing cycle-consistent adversarial networks (CYCLEGANs) for unpaired image-to-image translation as variational inference (VI) in an implicit latent variable model, providing a Bayesian perspective on this powerful class of deep generative models. (3) Introducing a model-agnostic reformulation of BO based on binary classification that eliminates restrictions on the underlying representation of the objective function, enabling the seamless integration of flexible modelling paradigms like deep learning to tackle complex optimisation problems. By enriching the interplay between deep learning and probabilistic ML, this thesis advances the foundations of AI, facilitating the development of more capable and dependable automated decision-making systems.