Artificial Intelligence and Sustainability Research Group
Prof. Dr. Ralf Herbrich, Nicolas Alder

MASTER'S PROJECT WINTER TERM 2024

TRUESKILL BEYOND GAUSSIANS

DESCRIPTION

The widespread implementation of artificial intelligence (AI) incurs significant energy use, financial costs, and CO2 emissions. This not only increases the cost of products, but also presents obstacles in addressing climate change. Traditional AI methods like deep learning lack the ability to quantify uncertainties, which is crucial to addressing issues such as hallucinations or ensuring safety in critical tasks. Probabilistic machine learning, while providing a theoretical framework for achieving much-needed uncertainty quantification, also suffers from high energy consumption and is unviable on a truly large scale due to insufficient computational resources. In this project, we will explore how energy-efficient sampling from mixture models can be utilized in sampling-based inference approaches for factor graphs.

Factor graphs are probabilistic graphical models used to represent the joint probability distribution as a factorization of a function. They provide a visual and mathematical way to decompose the probability distribution into simpler, more manageable components.

TrueSkill is a rating system designed for ranking and matchmaking in multiplayer games, leveraging factor graphs for its underlying computations. It estimates players' skill levels to facilitate more accurate and fair matchmaking. For our experiments, we will utilize TrueSkill as our experimental basis.

Inference refers to the process of computing certain probabilities or distributions related to the variables in the factor graph. It involves calculating marginal probabilities or the most likely configuration of variables given some observed data. In the simplest case, the Sum-Product algorithm is employed to obtain marginal probabilities, and the Max-Product algorithm is used to determine the most probable configuration of variables. When a factor graph contains loops, approximation algorithms such as loopy belief propagation are utilized. Typical operations in inference are performed on probability distributions, often Gaussians, due to their closed-form solutions. However, if resulting distributions are non-Gaussian, the closest Gaussian distribution is approximated and further propagated. While this approach keeps operations manageable, it can result in a loss of information for multimodal or complex distributions.

Sampling-based methods provide an alternative for handling complex, high-dimensional distributions for factor graph inference. Approaches, such as Importance Sampling, Gibbs Samling or Markov Chain Monte Carlo (MCMC) techniques generate samples from the posterior distribution to approximate desired quantities. These methods are particularly useful for handling multimodal and non-Gaussian distributions. Despite their flexibility, sampling-based methods face challenges such as ensuring convergence and managing computational complexity.

There are energy-efficient methods for sampling from uniform distributions. Those can be used in a mixture model to represent arbitrary distribution approximations.

Our objective is to explore how energy-efficient mixture models can be utilized in sampling-based approaches for inference in factor graphs. This exploration involves identifying and prototyping sampling-based methods that can leverage such mixture models. It includes thinking about the best parametrization and operations of the mixture model itself. We will also evaluate the quality of corresponding inferences and posterior approximation compared to standard inference approaches.

In this project, we will:

- Explore concepts in probabilistic machine learning algorithms such as factor graphs, sampling-based inference approaches, and mixture models.
- Create an energy-efficient sampling-based inference approach and evaluate its impact on the posterior approximation.
- Code an efficient framework that allows systematic experiments

Through successful completion of this project, you will:

- Pass the entire cycle of a research project: formulate initial research questions, conduct literature research, prototype solutions and conduct experiments, derive novel insights, publish results in a scientific paper
- Improve your programming and teamwork skills
- Deepen your probabilistic inference and sampling knowledge
- Improve your research methodology and academic writing

At the end of the project, we will condense our insights and results into a research submission to be submitted to a top AI/ML conference.

PREREQUISITES

- Prior understanding of the fundamentals of probabilistic learning (e.g., probability theory)
- Familiarity with Python, C++ and/or Julia

CONTACT

You are welcome to contact **Prof. Dr. Ralf Herbrich** (ralf.herbrich@hpi.de) or **Nicolas Alder** (Nicolas.Alder@hpi.de) or to visit us in the Villa, top floor (on Campus II).