**I.Introduction**

**1.1 Abstract**

Optimization problems are ubiquitous in various domains ranging from engineering design to machine learning model tuning. These problems often involve the maximization or minimization of an objective function, which represents a measure of performance or utility. However, in many real-world scenarios, the objective function may be unknown, black-box, or computationally expensive to evaluate. Traditional optimization techniques such as gradient-based methods may struggle in such cases due to their reliance on explicit function derivatives or exhaustive search procedures.

In response to these challenges, Bayesian optimization has emerged as a powerful and versatile approach for solving optimization problems with unknown or expensive objective functions. At its core, Bayesian optimization leverages probabilistic surrogate models, typically Gaussian processes (GPs), to capture the underlying structure of the objective function. By iteratively updating the surrogate model based on observed function evaluations, Bayesian optimization intelligently explores the search space to locate the global optimum while minimizing the number of evaluations required.

Central to the success of Bayesian optimization are the acquisition functions, which guide the selection of the next query point to evaluate the objective function. These acquisition functions trade-off between exploration (sampling from regions with high uncertainty) and exploitation (sampling from regions likely to yield improvement). The choice of acquisition function significantly influences the efficiency and effectiveness of the optimization process.

In recent years, there has been growing interest in leveraging bandit algorithms within the framework of Bayesian optimization. Bandit algorithms, inspired by the multi-armed bandit problem, are well-suited for scenarios where decisions must be made sequentially under uncertainty. In the context of Bayesian optimization, bandit algorithms play a crucial role in efficiently balancing exploration (sampling from regions of high uncertainty) and exploitation (sampling from regions likely to yield improvement) to identify the global optimum with minimal evaluations of the objective function.

The performance of bandit algorithms in Bayesian optimization is influenced by a multitude of factors, including the properties of the optimization problem (e.g., dimensionality, smoothness), the choice of acquisition function, and the algorithm's inherent exploration-exploitation trade-off strategy. Understanding how different bandit algorithms perform under various settings is crucial for practitioners and researchers alike, as it informs the selection of appropriate algorithms and parameter settings for different optimization tasks.

In this thesis, we aim to contribute to the understanding of bandit algorithms in the context of Bayesian optimization by conducting a comprehensive benchmarking study. Specifically, we will compare the performance of various bandit algorithms on a diverse set of benchmark functions, considering different problem settings and acquisition functions. By systematically evaluating and analyzing the performance of these algorithms, we seek to provide insights into their strengths, weaknesses, and applicability in practical optimization scenarios.

* 1. **Motivation and problem statement**

The motivations behind using bandit algorithms in Bayesian optimization stem from the need to address specific challenges inherent in optimization problems with unknown or expensive-to-evaluate objective functions. Traditional optimization methods often struggle in such scenarios due to their reliance on explicit function derivatives or exhaustive search procedures. Bayesian optimization, with its ability to construct probabilistic surrogate models of the objective function, offers a promising alternative. However, the effectiveness of Bayesian optimization hinges on the choice of acquisition functions and their corresponding optimization algorithms.

The motivation to integrate bandit algorithms into Bayesian optimization arises from their inherent ability to efficiently balance exploration and exploitation under uncertainty. Unlike static optimization approaches, bandit algorithms dynamically adapt their decision-making strategy based on past observations, making them well-suited for sequential decision-making in Bayesian optimization. Furthermore, bandit algorithms excel in sample efficiency, minimizing the number of function evaluations required to find the optimum, which is particularly advantageous in settings where function evaluations are costly or time-consuming.

**1.3 Contributions**

The contributions of this thesis lie in bridging the gap in the existing literature on benchmarking bandit algorithms in Bayesian optimization. While previous studies have investigated the performance of Bayesian optimization algorithms, including both bandit-based and non-bandit approaches, there remains a need for a comprehensive benchmarking study specifically focused on bandit algorithms. By systematically evaluating the performance of various bandit algorithms across a diverse set of benchmark functions, this thesis aims to provide insights into their effectiveness, robustness, and scalability in the context of Bayesian optimization.

Furthermore, this thesis seeks to shed light on the interplay between different factors influencing the performance of bandit algorithms in Bayesian optimization, including problem characteristics, choice of acquisition function, and algorithmic parameters. By conducting a rigorous benchmarking study, this thesis aims to offer practical guidance to practitioners and researchers in selecting appropriate bandit algorithms and parameter settings for different optimization tasks. In doing so, this thesis contributes to advancing the understanding and applicability of Bayesian optimization techniques in real-world optimization problems.

**Chapter 2 Bayesian Optimization Framework**

Bayesian optimization (BO) is a powerful framework for optimizing black-box functions that are expensive to evaluate. It leverages Bayesian inference to construct a probabilistic surrogate model of the objective function, enabling efficient exploration of the search space. This chapter provides a detailed overview of the Bayesian optimization framework, elucidating its key components and the underlying principles governing its operation.

1. Surrogate Model:

The surrogate model in Bayesian optimization serves as a probabilistic representation of the objective function. Gaussian processes (GPs) are commonly employed as surrogate models due to their flexibility and ability to capture uncertainty. GPs provide not only point estimates of the objective function but also uncertainty estimates, crucial for guiding the exploration-exploitation trade-off. The process of updating the surrogate model involves incorporating new observations and updating the posterior distribution over the objective function.

**2. Acquisition Functions:**

Acquisition functions play a pivotal role in Bayesian optimization by quantifying the utility of evaluating the objective function at a particular point in the search space. Popular acquisition functions include Probability of Improvement (PI), Expected Improvement (EI), and Upper Confidence Bound (UCB). These functions balance exploration (sampling from regions of high uncertainty) and exploitation (sampling from regions likely to yield improvement) to guide the search towards the global optimum. The choice of acquisition function is influenced by factors such as the problem domain, computational resources, and the desired trade-off between exploration and exploitation.

**3. Optimization Algorithm:**

The optimization algorithm in Bayesian optimization governs the iterative process of updating the surrogate model and selecting query points to evaluate the objective function. Common optimization algorithms include Sequential Model-based Optimization (SMBO) and its variants, such as the Bayesian Optimization by Gaussian Processes (BOGP) algorithm. These algorithms iteratively update the surrogate model based on observed function evaluations and select the next query point using the acquisition function. The choice of optimization algorithm depends on considerations such as computational efficiency, scalability, and robustness to noise.

**4. Iterative Process:**

Bayesian optimization operates in an iterative manner, where each iteration involves updating the surrogate model based on observed function evaluations and selecting the next query point using the acquisition function. This iterative process continues until a stopping criterion is met, such as reaching a predefined budget of function evaluations or achieving a satisfactory solution. The convergence properties of Bayesian optimization are influenced by factors such as the choice of surrogate model, acquisition function, and optimization algorithm.

**5. Hyperparameter Tuning:**

A key application of Bayesian optimization is hyperparameter tuning, where the goal is to optimize the hyperparameters of machine learning models or algorithms. Bayesian optimization offers an automated and principled approach to hyperparameter tuning, leveraging its ability to efficiently explore the hyperparameter space and identify optimal configurations. Hyperparameter tuning with Bayesian optimization involves specifying a search space for the hyperparameters, defining a suitable surrogate model, and selecting an appropriate acquisition function and optimization algorithm.

**6. Practical Considerations:**

In practice, Bayesian optimization may encounter challenges such as scalability to high-dimensional spaces, noisy observations, and computational constraints. Various techniques have been proposed to address these challenges, including approximate inference methods, parallelization strategies, and surrogate model approximation techniques. Additionally, Bayesian optimization can be combined with other optimization methods, such as evolutionary algorithms and reinforcement learning, to address specific problem domains and constraints.

Conclusion:

This chapter provides a comprehensive and detailed overview of the Bayesian optimization framework, highlighting its key components, principles, and practical considerations. Understanding the Bayesian optimization framework is essential for conducting benchmarking studies and evaluating the performance of bandit algorithms within this framework, as discussed in subsequent chapters.