R244 Open-Source Project: OtterTune

Session 8 4 December 2024

1

Problem

- DBMSs have hundreds of configuration parameters (knobs)
- Parameters are not independent
- Empirical knowledge required to set correct values
- Large high-dimensional space of configurations
- How do we find the global minima?

OtterTune Architecture



Project Aims

- Explore the metric pruning process
 - Factor Analysis used to reduce dimensionality
 - K-means clustering: retain only the metric closest to the centre and discard the rest
- Explore the workload mapping process
 - Euclidean distance calculated between current observation and all previous observations
- Explore how different components pipeline together
 - Components described independently. No specific detail on the entire pipeline.
- Implement a DNN-based tuning model
 - GPs have limited performance on large datasets and high-dimensional features

Project Work and Plan

- Work so far:
 - Set up OtterTune locally
- Work to do:
 - Understand code
 - Implement modifications to OtterTune
 - Measure performance
 - Write report