

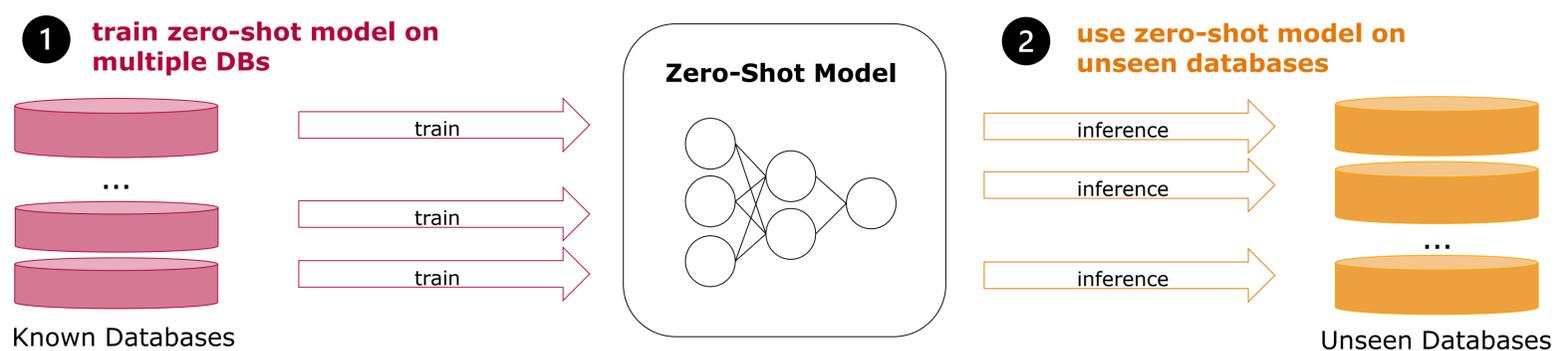
# Research Proposal: Physical Design Tuning with Zero-shot Learning

**Physical Design Tuning** deals with adjusting the physical schema of a working database system to adapt it to changing workload requirements. To minimize the runtime of queries one aims at optimizing the selection of indices and materialized views (MV). Combined, “both indexes and materialized views are physical structures that can significantly accelerate performance.” [1]

**Problem statement:** So far, the task of tuning the physical design of a database was tackled by using classical optimization approaches which come with one major downside: Inaccuracy. Thus, optimizers often ended up choosing non-optimal design options. While newer approaches such as workload-driven or reinforcement learning do improve performance of databases, both options require training queries of different physical layouts for new databases and cannot be used out of the box. [2]

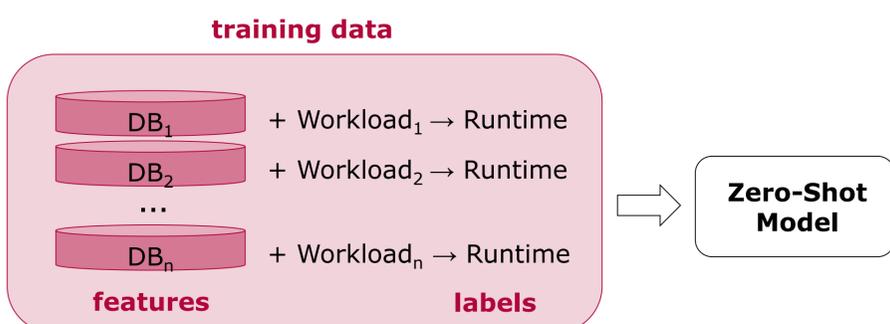
**Why zero-shot learning?** Recent work by Hilprecht et al. has proven the feasibility of zero-shot learning for physical cost estimations. By training a zero-shot model on different databases, Hilprecht et al. were able to outperform the cost estimations of workload-driven models and classical optimization approaches. Most importantly, though: The trained model was used out-of-the box and required no additional training queries for the new database. [2]

This research proposal aims at applying the zero-shot learning approach to the physical design tuning of databases based on indices and MVs.



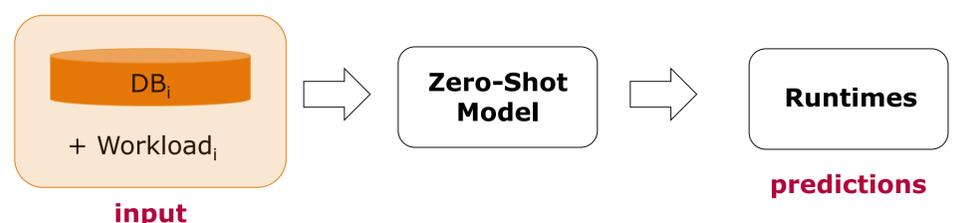
## 1 Train zero-shot model

First results indicate that using the zero-shot learning approach for optimizing the selection of indices and materialized views is feasible [2,3]. Besides elaborating named indications, a model will be trained on zero-shot encodings for different MV-index combinations and the required runtime for different workloads.



## 2 Inference on Unseen Database

After that, the trained model will run in a “what if”-mode and predict runtimes of different MV and index configurations and combinations of unseen databases. Other design tuning approaches will be used as a reference to evaluate the predictions.



[1] S. Agrawal, S. Chaudhuri, and V. Narasayya, “Automated Selection of Materialized Views and Indexes for SQL Databases,” 2000. Accessed: Jan. 25, 2022. [Online]. Available: <http://www.vldb.org/conf/2000/P496.pdf>.  
 [2] B. Hilprecht and C. Binnig, Eds., “One Model to Rule them All: Towards Zero-Shot Learning for Databases,” [www.youtube.com](http://www.youtube.com), Jan. 13, 2022. <https://www.youtube.com/watch?v=s-84ALjPspg> (accessed Jan. 25, 2022).  
 [3] B. Hilprecht and C. Binnig, “One Model to Rule them All: Towards Zero-Shot Learning for Databases,” 2022. [Online]. Available: <https://arxiv.org/pdf/2105.00642.pdf>.