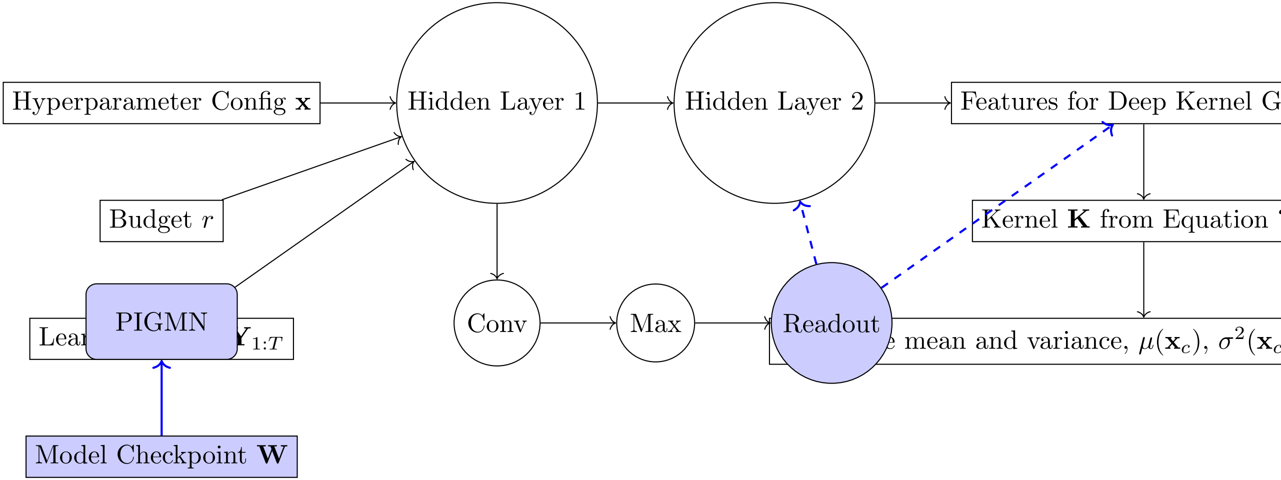

Figure 1: We show an overview of our method, Forecasting Model Search (FMS), which builds on DyHPO's multifidelity method from Algorithm **??**. Novel components of FMS are highlighted in blue and further detailed in Algorithm **??**. We include DyHPO's features from the hyperparameter configuration, budget, and learning curve [**?**]. Notably, we also featurize the model's checkpointed weights $\mathbf{W}$ with a permutation-invariant graph metanetwork (PIGMN) as in Section **??** for input to a deep kernel GP (see Equation **??**/**??**). This provides the HPO with an – often pre-existing – rich source of information, which implicitly includes the architecture, dataset, loss, and optimization process. FMS shows improved predictions about hyperparameter performance across compute budgets (see Table **??**), improved quality of the final selected configuration across compute budgets (see Figure **??**), and a potential to generalize beyond what was seen in training (see Figure **??**). Specific design choices for this surrogate model are detailed in Appendix Section **??**.