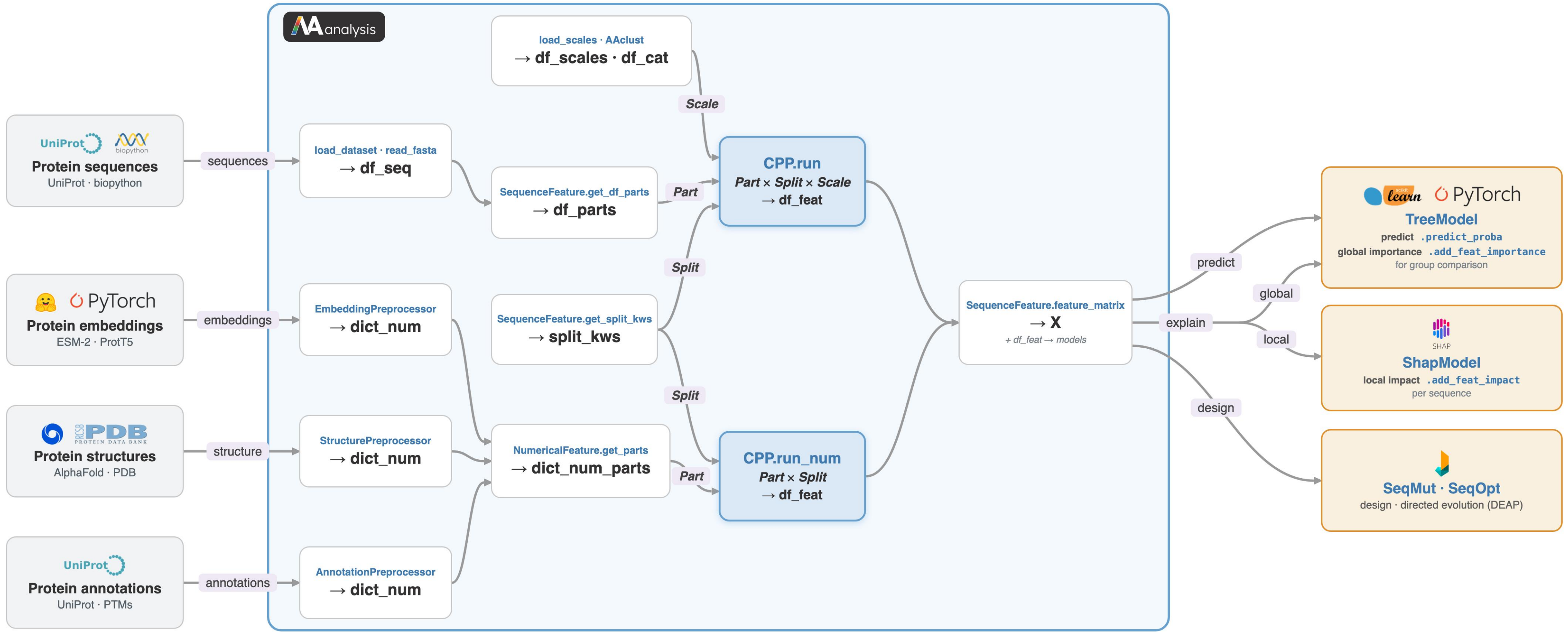


Outside — external data source
 Core — AAanalysis CPP pipeline
 Wrapped — AAanalysis class + external engine (logo)



Core: `df_seq` (`load_dataset / read_fasta`) → `df_parts`; `split_kws` and `df_scales` are built directly. **CPP.run** combines **Part × Split × Scale**; **CPP.run_num** uses numeric values (embeddings ESM-2/ProtT5 or AlphaFold/PDB structures → `dict_num` → `dict_num_parts`). Both produce `df_feat` → `X`. **Wrapped** (amber, outside): **TreeModel** (scikit-learn / PyTorch) predicts and gives **global** feature importance, **ShapModel** (SHAP) gives **local** feature impact per sequence, and **SeqMut · SeqOpt** (DEAP) designs variants — all take `X` and `df_feat`.