Phase 4: (Weeks 8 - 10)

* Week8: finalize baseline model which would be used in active learning pipeline, connect with virtual screening group to get data input and output and solve any conflict on data format
* Week9: construct the running pipeline and produce some initial result
* Week10: Based on week9 results, tune the pipeline and try to increase the accuracy.