General ideas

* Optimize the set Sherpa (in progress)
  + Optimize regular learning set
    - Done (see results in
  + Optimize over transfer set (real goal)
    - Transfer set to pytorch for better readability/ build out (done)
    - Go to batching (framework done)
      * Optimize batch sizing for the set
* Observe results of optimizing hyper parameters etc
* Take optimized set and do neural regression on multiple
  + Optimize regression weighting on testing or simply use even
* Cluster analysis
  + Group cluster by PCA
  + Look into transfer learning between clusters

Cluster of materials or cluster of MOF?

* + Force bad performance
  + Just simply use one for the other
  + Transfer learn one to another

Draft---

Abstract:

* Replication results
* Cluster results
  + Clustering understanding
  + General Clustering and transfer
  + Topology Cluster and transfer
  + Sub cluster transfer

Introduction:

* Need to find papers---
* MOF machine learning
* MOF history
* MOF significance and use cases
* MOF clustering
  + No examples?
* Transfer learning clustering
  + Not much

Methods/Results:

* Two main goals- replicate and improve, cluster transfer learning
* Replication
  + Get similar values for direct and transfer learning
* Improvement
  + Pre model
    - Look at clustering
    - Look at PCA
  + Model
    - Improve epoch approach
    - Batching
  + Hyperparameters
    - Optimized hyper parameters
* Clustering
  + Clustering based on PCA general understanding
    - Performance PCA
    - Clustering on PCA
    - Cluster on topology median PCA
  + Clustering transfer
    - Clustering results
      * Error per epoch / time to learn
      * Cluster final performance comparison
    - Cluster types
      * Full cluster of all points
      * Cluster on topology
      * Subcluster on topology

Draft:

Abstract:

Introduction:

Methods:

The goal of the study was twofold. The first was to make sure the previous paper [1] was replicable and to conduct a more thorough attempt for transfer learning on the tasks outlined in the original paper. To this end, this entailed not just the transfer approach outlined in the previous paper, but a more complete one using all facets of machine learning for the best results. This includes a more comprehensive understanding of the data prior to creating the model, improving the model itself, optimizing hyperparameters in the transfer learning case. These should all lead to a more accurate understanding of transfer learning in cases and proved the results of the previous paper [1].

To being, there was a study of the data itself before it was processed. This was conducted by first passing the features of significance, in this case 'void fraction', 'Vol. S.A.', 'Grav. S.A.', 'Pore diameter Limiting', 'Pore diameter Largest']

The second goal of the study came about after the analysis of the data prior to transfer learning. It occurred that there was significant [FIGURE???] differences in the set after applying a PCA analysis.

Results:

References:

Data Set:

(30) Colón, Y. J.; Gómez-Gualdrón, D. A.; Snurr, R. Q. Topologically Guided, Automated Construction of Metal−Organic Frameworks and Their Evaluation for Energy-Related Applications. Cryst. Growth Des. 2017, 17, 5801−5810.