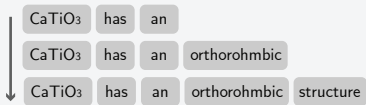
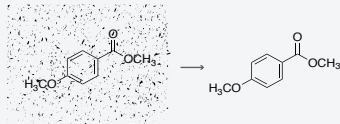


predicting masks

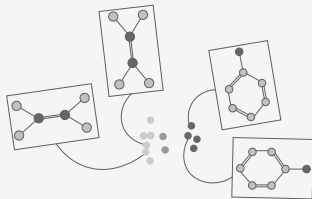
generative



learning from context



learning to denoise

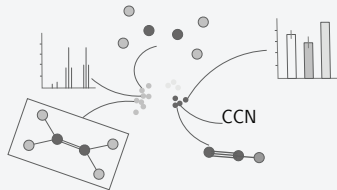


aligning embeddings

contrastive



learning to cluster



cross modal alignment