| **Working** | **Goals** | **How-To-do** | **Progress** |
| --- | --- | --- | --- |
| Running models   * Confgf * CGCF | Running models   * Confgf * CGCF * ConfVAE * GRAPHGD * DL4 | 1. Go to repo and follow instructions. **30 minutes** should be enough to get the repo running in cluster. 2. Run the commands to see if it is generating conf**.5 min** | 1. Tried to train ConfVAE but not working.   **CONFVAE**: No available chkpoints of the repo.  Tried to train the model on cluster and Zain’s laptop but taking too long.  I have trained the model and put those checkpoints in models . Successfully.  I have found chkpoints and Generated the conf. Chkpoints were in logs  Need to Test them.   1. Trying with Graphgd    1. Graphgd had issues of tensorflow. Able to solve them and run on the cli terminal but then get error of no gpu    2. In bashfile I get error that no Tensorflow.   Graphgd is generating conf. Output in Rnd\_junk desktop local. |
| A Visualization   * CGCF | B Visualization   * CGCF * Confgf * ConfVAE | C  1. **CONFGF**: Try to extract the RDKIT mol obj for visualization.    1. Open notebook **Visulize\_Conformers\_C** and instead of taking smiles take that rdkit obj in array.    2. Try to visualize it as done for CGCF, taking just one mol. 2. **CONFVAE**:Visualize it the same way as other two, depending on the output. Try to extract the RDKIT mol obj for visualization. 3. RDKIT mol obj holds info about 3d structure. |  |
|  | D **Evaluation**  Comparison of Models   1. RDKIT 2. PYMOL - Extrafit prop 3. Align does 3D 4. Compute RMSD val | E  1. From dataset QM9, take 10 molecules.    1. Open any of your notebook    2. Load pickle file and take first 10 molecules or take their RDkit mol.    3. Go to any of the Kaggel project and try to take molecules from there if not able to find    4. Ask divin or umer bhai to explain how to do it    5. ~~Ask them to do it~~ 2. Look for unique conformations    1. Use Extrafit method in Pymol to align the molecules.  **rdkit.Chem.rdMolAlign module** can be used for the alignment of molecules. It has a reference mol and prbmol (mol to be aligned). This method outputs the RMSD value. I can assign a threshold value for performance measure e.g if the RMSD\_value <0.5 the conf of both algorithms are similar if RMSD\_value >0.5 then the conf generate are not simialr.(if needed). This way I can check the uniqueness of these confromations generated by the algorithms.  * 1. **How to Check accuracy of mol?**  1. Come up with a common comparison tool for these confromations. 2. Save the cofnormations in a separate file. 3. In ConfVAE in eval.py some methods to compare by RMSD | F  ConfVAE has Mean and Cov values in output file.  Check if PYMOL and Rdkit method can evaluate 3d mol **rdkit.Chem.rdMolAlign module is for 3D**  * 1. **PYMOL: On wiki page it says** graphical representation of aligned atom pairs as lines in the 3D viewer.   **BOTH OF THEM CAN HANDEL 3D.**   * 1. RMSD value is being generated for all these 3 repos.   2. <https://www.labxchange.org/library/items/lb:LabXchange:6d307d41:html:1#:~:text=To%20align%20two%203D%20structures,structure%20of%20X%20onto%20Y.> TutoriaL TO ALIGN 3D MOLECULES.   3. <http://pldserver1.biochem.queensu.ca/~rlc/work/teaching/pymol/alignment/> Tutorial to get RMSD value from PYMOL   4. <http://people.reed.edu/~glasfeld/pretty/align.html> aNOTHER SIMPLER TUTORIAL TO ALIGN 2 MOL.   5. <https://www.researchgate.net/post/how_can_I_calculate_RMSD_using_Pymol> ANother way to execute RMSD. |
| Report   * Introduction * State of the art- tids bits | Report   * Abstract * Intro fixes * State of the art * Methodologies * Solution * Evaluation * Results * Experiments * Images * Tables * Printing | 1. Complete Introduction 2. Complete State of the art 3. . 4. . 5. . 6. . |  |
|  |  | Details of each model |  |

| Problem | Solution |
| --- | --- |
| E1: The loaded QM9 file has all the same smiles which means:   * I am either not appending all of the mol or * The file has wrong data. * I am loading wrong file | * The input file has mol+conf . I have confirmed it from paper and paper of dataset. * I will put the output in a dataframe and generate a table of those confirmations to put it in report. |
|  | <https://pubs.rsc.org/en/content/articlelanding/2019/me/c9me00039a> read it for ML |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

python scripts/run.py --train\_path=/scratch/zkarim2s/graphdg/iso17\_split-0\_train.pkl --test\_path=/scratch/zkarim2s/graphdg/iso17\_split-0\_test.pkl