# Tutorial for using MLIP code

MLIP code uses MTP form for making machine learned interatomic potentials. Data required for making the interatomic potentials is fed in the form of OUTCAR files which are generated by running DFT simulations. In order for the MLIP code to make potential file, first OUTCAR files need to be converted into cfg format. The cfg format is the format read by the MLIP code. To convert OUTCARS into cfg format, run the following script shown below

### Script for converting OUTCAR files to cfg format (making\_cfg-reading\_OUTCAR\_last-cong.sh)

#!/bin/bash/

n=14

for ((ivar=0; ivar<=n; ivar+=1))

do

./mlp convert-cfg OUTCAR-$ivar temp.txt --input-format=vasp-outcar --last

cat temp.txt >> train.cfg

done

where, n is the number of OUTCAR files in a folder. The “mlp” is the binary file obtained after compiling the MLIP code. The “convert-cfg” is the command for converting OUTCAR to cfg format. The “OUTCAR-$ivar” is the name of the OUTCAR file, “temp.txt” is the name of the temp file to store the data, “—input-format=vasp-outcar” represents the command for telling the code the format of the input files, and “last” means to make cfg from the last values from the OUTCAR files. Then “temp.txt” file data is concatenated to “train.cfg” file.

Obtained “train.cfg” file is used for building the machine learning potentials. For building the ML potentials following files are required.

* init.mtp : This file contains the general parameters used for training the model. This is a general file obtained from the website. Some of the parameters can be changed according to user’s requirement and the order of the potentials can also be chosen, also found on web.
* train.cfg : This file contains the configurations used for training the model.

### Use the following command to build the machine learning potentials (making\_potential.sh):

Make sure to input minimum distance and number of species in init.mtp.

./mlp train init.mtp train.cfg --trained-pot-name=pot.mtp --valid-cfgs=test.cfg

Where “—trained-pot-name=pot.mtp” is the filename to store the built potentials. We can also validate the potentials by providing the “test.cfg” file containing the test configurations but it’s optional. After running the above command, a file named “pot.mtp” will be generated which can be used for making further predictions. A code for extracting energies and volumes is also provided by the name,

### extracting\_energy\_vol.py.

This code requires only the filename in the .cfg format to run and outputs a text file and a figure for E-V curve.

Potential file named pot.mtp can now be used for calculating energy, force, and stress for the new configurations. For making prediction on new structures, input file should be in format of cfg. Thus, a code named

### “read\_POSCAR\_make\_cfg.py”

is written which converts POSCAR files to cfg format with no/little user input. All the POSCARs should be in cartesian coordinate before processing it through this code. All the POSCARs can be provided at once for making a single cfg file which then can be used for making prediction. Following command is used for making prediction

### ./mlp calc-efs pot.mtp test.cfg test\_predictions.cfg

Above command will generate test\_prediction.cfg file containing predicted forces and energies. Obtained cfg file can then be also be used for making E-V curve using the python code mentioned above.