深度势能生成器(DP-GEN)

Deep Potential GENerator

为什么要建模原子间相互作用

- 计算模拟在材料性质研究和设计中发挥日益重要的作用
- 关键科学问题：材料的微观结构
- 微观结构由原子间相互作用决定

$$\underbrace{E}_{\textbf{势能}} = E(\underbrace{r1, r2, r3, \ldots}_{\textbf{原子坐标}})$$

原子间相互作用建模传统手段

一. 第一性原理计算

KS-DFT：

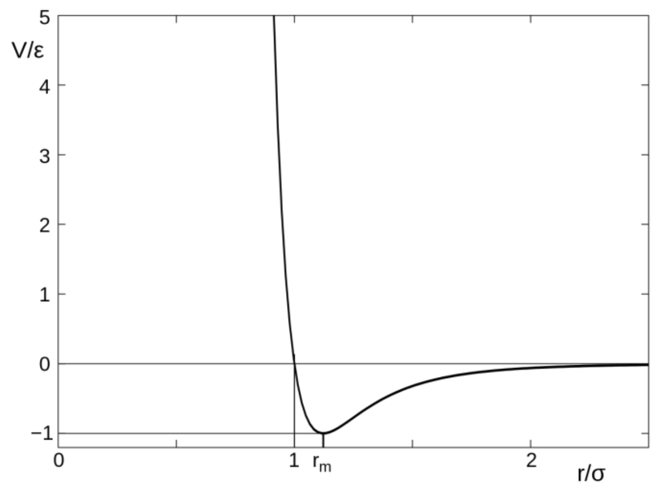
$$E[\rho] = F_{HK}[\rho] + \int \rho(\mathbf{r}) v_{ext}(\mathbf{r}) d\mathbf{r}$$

$$F_{HK}[\rho] = T[\rho] + V_{ee}[\rho]$$

二. 经验力场

Lennard-Jones 势：

$$V_{\mathrm{LJ}} = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$$



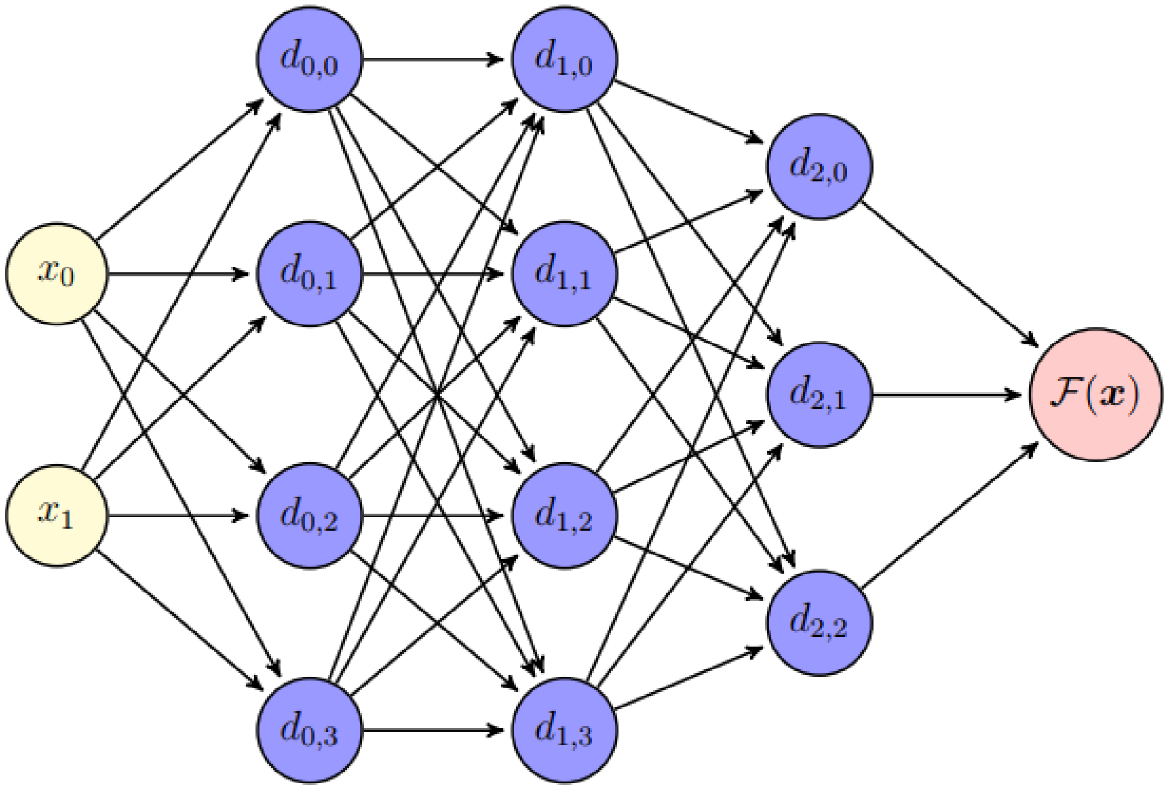
深度学习建模原子间相互作用

$$E = E(r1, r2, r3, ...)$$

3N维函数

- 原子间相互作用建模在数学上是高维函数的表示和逼近问题；
- 传统数学工具对高维函数缺乏有效手段；
- 深度学习为高维函数的逼近提供了有力工具。

深度学习



$$\boldsymbol{x} \xrightarrow{\mathcal{L}^0} \boldsymbol{d}_0 \xrightarrow{\mathcal{L}^1} \boldsymbol{d}_1 \xrightarrow{\mathcal{L}^2} \boldsymbol{d}_2 \xrightarrow{\mathcal{L}^{\text{out}}} \mathcal{F}(\boldsymbol{x})$$

$$\boldsymbol{d}^p = \mathcal{L}^p(\boldsymbol{d}^{p-1}) = \boldsymbol{d}^{p-1} + \Delta t\, \phi\Big(\boldsymbol{W}^p \cdot \boldsymbol{d}^{p-1} + \boldsymbol{b}^p\Big)$$

深度势能平滑模型 (DeepPot-se)

笛卡尔坐标描述原子i截断半径内的局域环境：

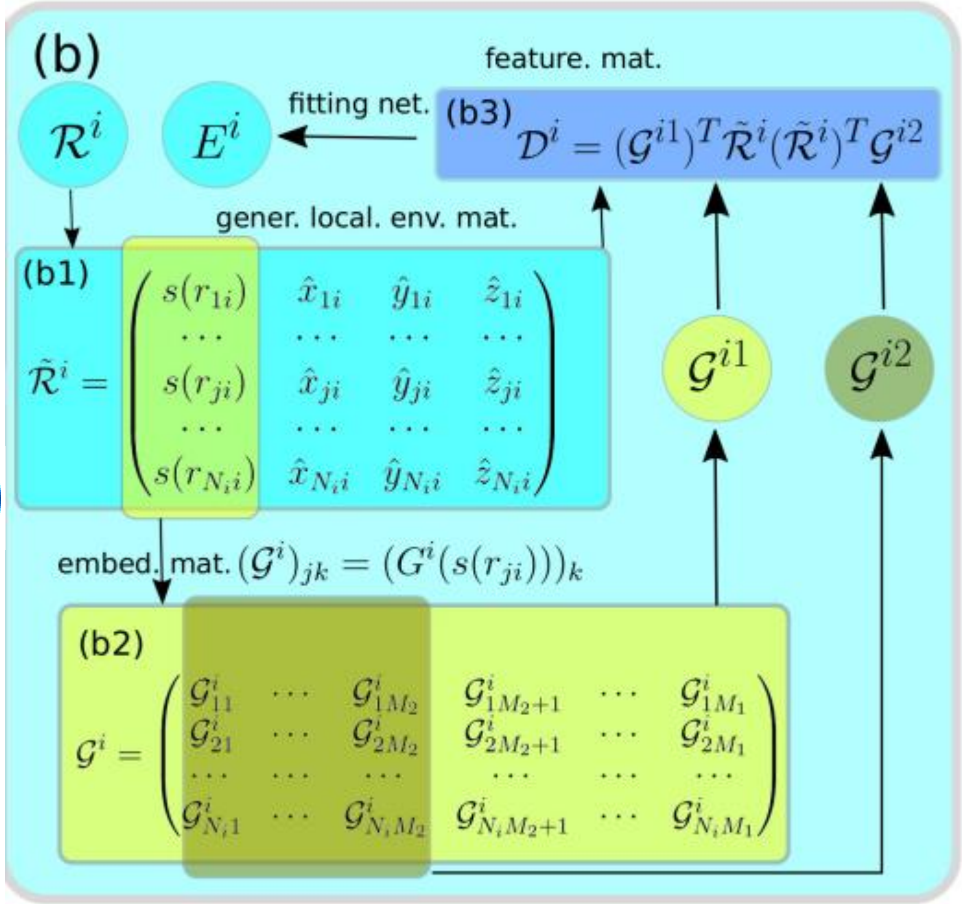
$$\mathcal{R}^i = \{\boldsymbol{r}_{1i}^T, \cdots, \boldsymbol{r}_{ji}^T, \cdots, \boldsymbol{r}_{N_i,i}^T\}^T, \ \boldsymbol{r}_{ji} = (x_{ji}, y_{ji}, z_{ji})$$

广义坐标：

$$\{x_{ji}, y_{ji}, z_{ji}\} \mapsto \{s(r_{ji}), \hat{x}_{ji}, \hat{y}_{ji}, \hat{z}_{ji}\}$$

$$\hat{x}_{ji} = \frac{s(r_{ji})x_{ji}}{r_{ji}}, \ \hat{y}_{ji} = \frac{s(r_{ji})y_{ji}}{r_{ji}}, \ \hat{z}_{ji} = \frac{s(r_{ji})z_{ji}}{r_{ji}}, \text{ and } s(r_{ji})$$

$$s(r_{ji}) = \begin{cases} \frac{1}{r_{ji}}, & r_{ji} < r_{cs}. \\ \frac{1}{r_{ji}}\left\{\frac{1}{2}\cos\left[\pi\frac{(r_{ji} - r_{cs})}{(r_c - r_{cs})}\right] + \frac{1}{2}\right\}, & r_{cs} < r_{ji} < r_c. \\ 0, & r_{ji} > r_c. \end{cases}$$



Deep Potential: 训练

- 能量:

$$E = \sum_i E_i$$

- 力:

$$\mathcal{F} = -\nabla_{\mathcal{R}} E \ \ (\mathcal{F}_{ij} = -\nabla_{\mathcal{R}_{ij}} E)$$

- 维里:

$$\Xi = \mathrm{tr}[\mathcal{R} \otimes \mathcal{F}] \ \left(\Xi_{ij} = \sum_{k=1}^{N} \mathcal{R}_{ki} \mathcal{F}_{kj} \right)$$

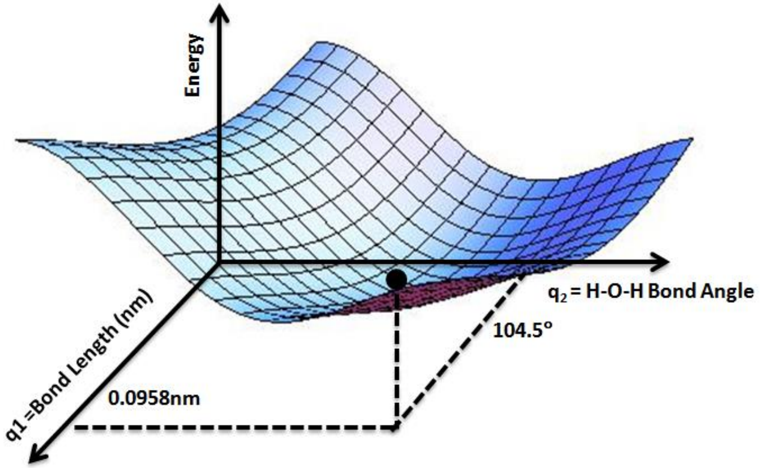
- 损失函数:

$$L(p_\epsilon, p_f, p_\xi) = p_\epsilon \Delta\epsilon^2 + \frac{p_f}{3N} \sum_i |\Delta \boldsymbol{F}_i|^2 + \frac{p_\xi}{9} ||\Delta\xi||^2.$$

$$p(t) = p^{\mathrm{limit}}\left[1 - \frac{r_l(t)}{r_l^0}\right] + p^{\mathrm{start}}\left[\frac{r_l(t)}{r_l^0}\right]$$

目标问题

- 目标：精确预测势能面。
- 方法：利用昂贵的第一性原理计算数据，通过神经网络进行训练，获得深度势能模型。



目标问题

势能面上的样本数量：

- 固定原子位置，改变原子种类：

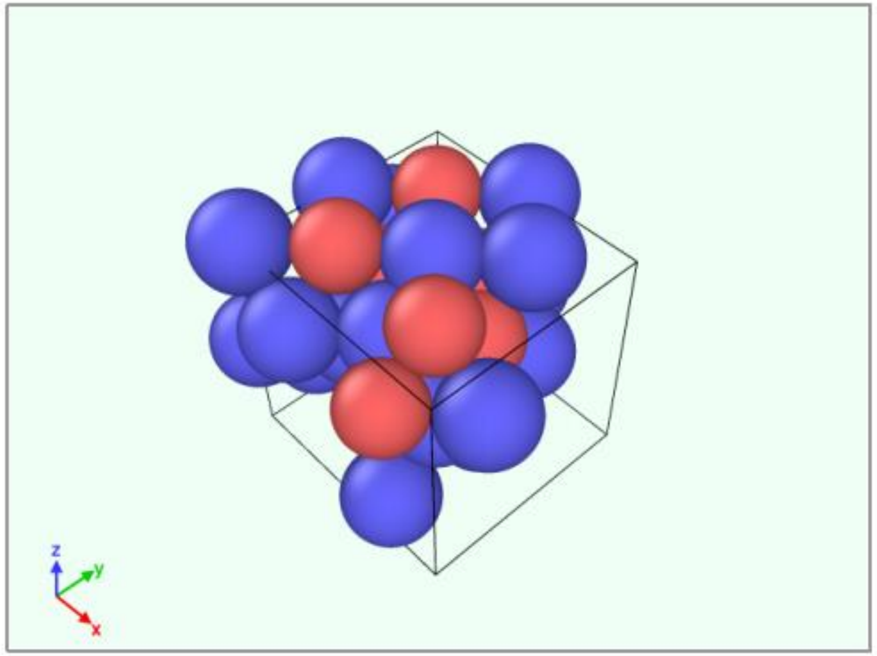
$\Omega = 2^{32}$

- N元素体系：

$\Omega = N^{32}$

- 原子位置不固定：

$\Omega = \infty$



含32个原子的二元合金

目标问题

- 问题一：充足采样
  如何对势能面充足采样，有效覆盖训练可靠模型所需的样本空间?
- 问题二：筛选样本
  如何高效筛选样本，选取训练价值高的构象进行第一性原理计算?

解决策略：同步学习

- 采样器：
  DPMD
- 误差判据：
  模型偏差(model deviation)

采样器

- DPMD势能面采样优势：

  高效

  更新模型改进采样质量

  改变MD条件拓展采样范围

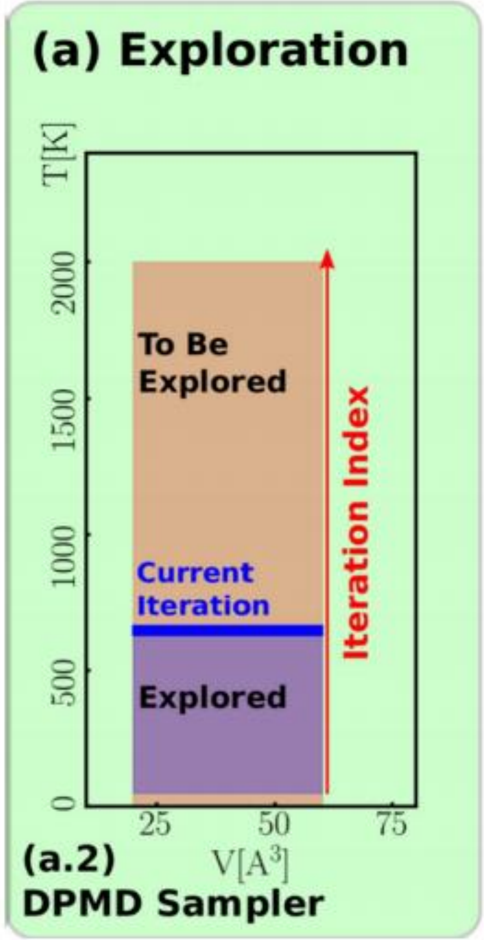
- 可拓展的采样策略：

  增强采样

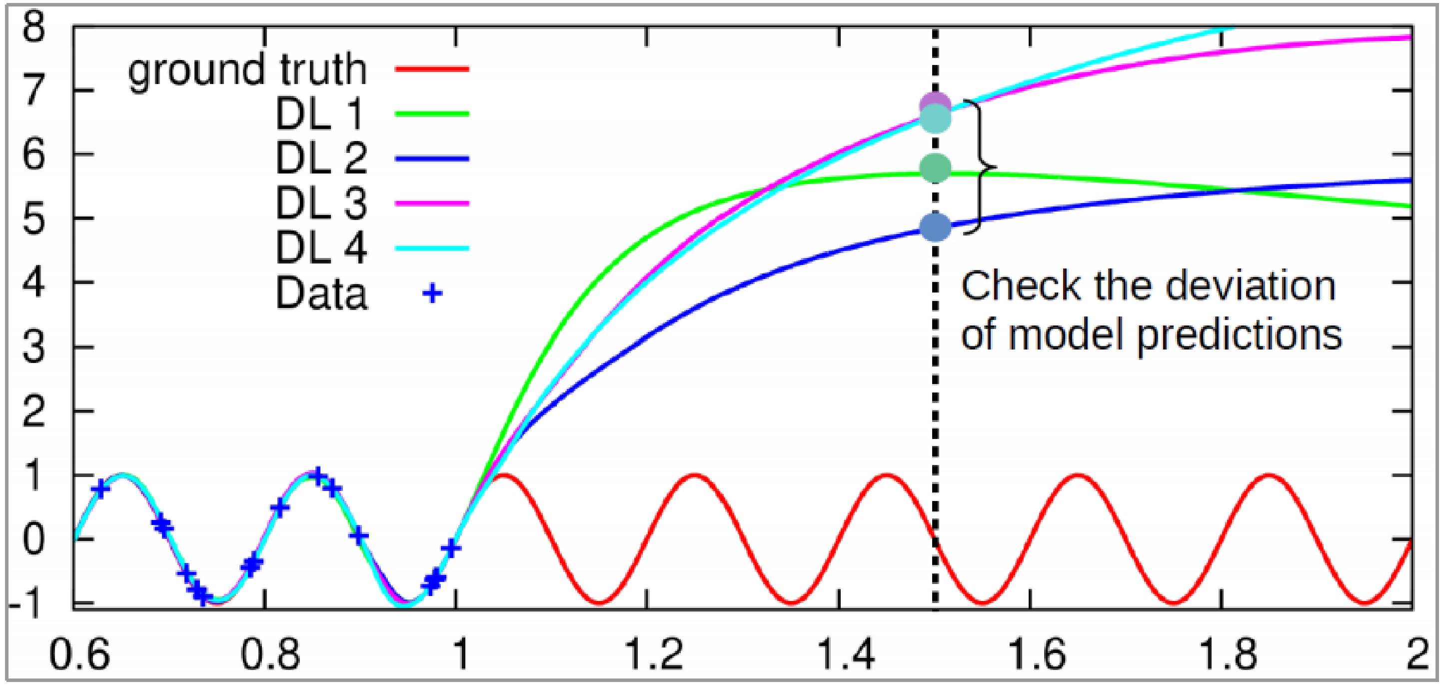
  蒙特卡罗

  结构搜索

  …



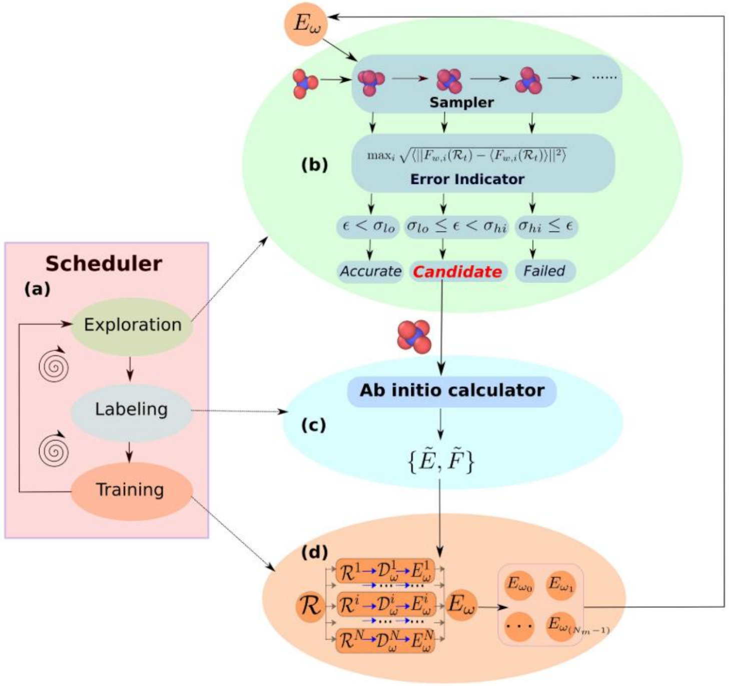
误差判据

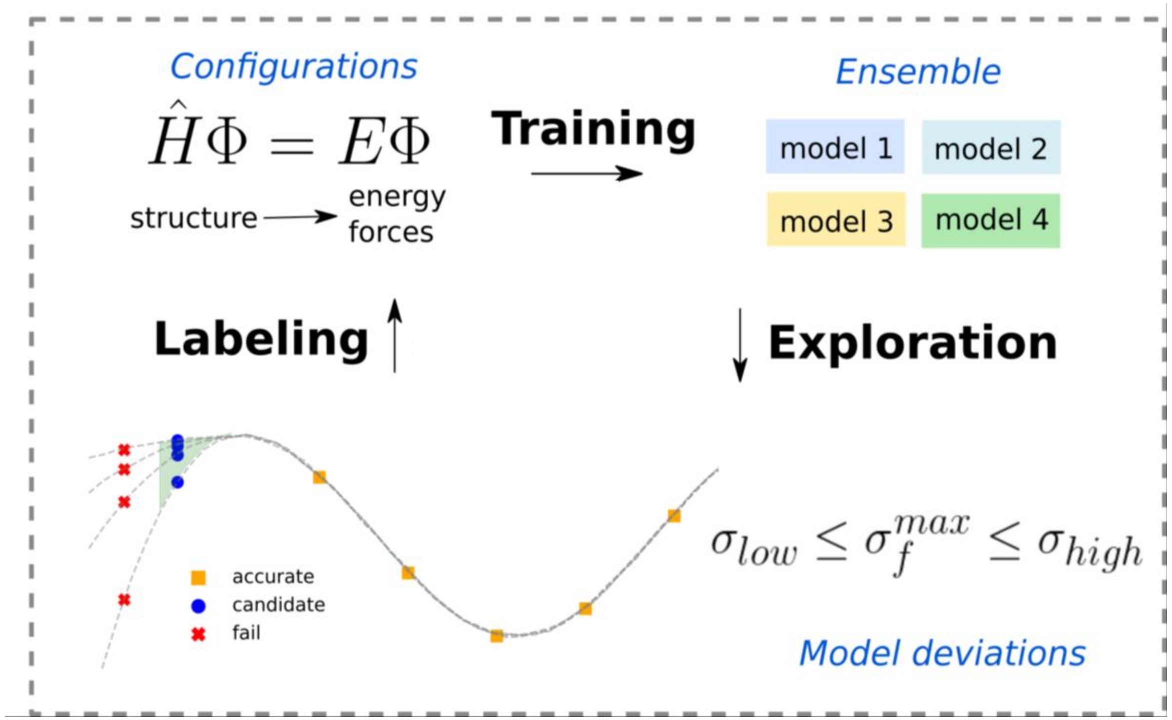


模型偏差(model-deviation): $\epsilon_t = \max_i \sqrt{\langle \| \boldsymbol{F}_{\omega,i}(\mathcal{R}_t) - \langle \boldsymbol{F}_{\omega,i}(\mathcal{R}_t) \rangle \|^2 \rangle}$

真实'力'误差: $\hat{\epsilon}_t = \max_i \sqrt{\langle \| \boldsymbol{F}_{\omega,i}(\mathcal{R}_t) - \widetilde{\boldsymbol{F}}_{\omega,i}(\mathcal{R}_t) \|^2 \rangle}$

DP-GEN主流程





Exploration： 探索样本空间，调用MD模拟软件；
Labeling： 增加标记样本，调用DFT计算软件；
Training： 训练新的模型，调用DeePMD-kit。

DP-GEN的基本命令

- DP-GEN基本命令：

  dpgen sub-command PARAM MACHINE

- **sub-command** 为DP-GEN支持的任务类型，基本流程中包括 init_bulk/init_surf, run, test。

  init_bulk/init_surf: 为体相/表面体系准备初始数据。

  run: 深度势能生成器主流程。

  test: 计算参考体系性质测试模型表现。

- **PARAM**: 指代输入参数文件，各类型任务需要独立的参数文件。
- **MACHINE**: 机器配置文件。

Run 主流程

- 命令：dpgen run param machine
- 多次迭代
- 每一迭代包含三个步骤：
  - ✓ 00.train (Training)
    训练新的模型
  - ✓ 01.model_devi (Exploration)
    探索样本空间
  - ✓ 02.fp (Labeling)
    增加标记样本

迭代序号	各迭代的阶段序号	进程
0	0	make_train
0	1	run_train
0	2	post_train
0	3	make_model_devi
0	4	run_model_devi
0	5	post_model_devi
0	6	make_fp
0	7	run_fp
0	8	post_fp
1	0	make_train

机器配置文件

- MACHINE 为 Task Dispatcher 指定三类计算任务的机器环境与资源需求。
  train (DeepMD-kit)
  model_devi (MD-LAMMPS)
  fp (DFT-VASP)
- Slurm, PBS, ALI, AWS, shell...
- 配置环境，加载软件模块等。
- 指定CPU, GPU, 内存, 时限。

  ✓ 自动化 可通用
  ✓ 智能选择最合适的计算机器提交任务
  ✓ 任务收发与断点恢复

```
"train": [
  {
    "machine": {
      "machine_type": "lsf",
      "hostname": "210.34.15.205",
      "port": 22,
      "username": "fqgong",
      "work_path": "/home/fqgong/temp/train"
    },
    "resources": {
      "node_cpu": 4,
      "numb_node": 1,
      "task_per_node": 4,
      "partition": "large",
      "exclude_list": [],
      "source_list": [
        "/share/base/scripts/export_visible_devices -t 800"
      ],
      "module_list": [
          "cuda/9.2",
          "deepmd/1.0"
            ],
      "time_limit": "23:0:0"
    },
    "python_path": "/share/deepmd-1.0/bin/python3.6"
  }
```

机器配置文件

```
"model_devi": [
  {
    "machine": {
      "machine_type": "lsf",
      "hostname": "210.34.15.205",
      "port": 22,
      "username": "fqgong",
      "work_path": "/home/fqgong/temp/md"
    },
    "resources": {
      "node_cpu": 2,
      "numb_node": 1,
      "task_per_node": 2,
      "partition": "large",
      "exclude_list": [],
      "source_list": [
        "/share/base/scripts/export_visible_devices -t 800"
      ],
      "module_list": [
          "cuda/9.2",
          "deepmd/1.0",
          "gcc/4.9.4"
              ],
      "time_limit": "23:0:0"
    },
    "command": "lmp_mpi",
    "group_size": 5
  }
```

```
"fp": [
  {
    "machine": {
      "machine_type": "lsf",
      "hostname": "121.192.191.51",
      "port": 6666,
      "username": "fqgong",
      "work_path": "/old.data/fqgong/test1/fp"
    },
    "resources": {
      "cvasp": false,
      "task_per_node": 24,
      "numb_node": 1,
      "node_cpu": 24,
      "exclude_list": [],
      "with_mpi": true,
      "source_list": [
      ],
      "module_list": [
          "intel/17u5",
          "mpi/intel/17u5"
          ],
      "time_limit": "12:00:00",
      "partition": "medium",
      "_comment": "that's Bel"
    },
    "command": "/share/apps/vasp/5.4.4/bin/vasp_std",
    "group_size": 30
  }
]
```

参数设置文件

```
"sys_configs_prefix": "/old.data/fqgong/example/",
"sys_configs":    [
  [
    "init/01.scale_pert/sys-0004-0001/scale-1.000/000000/POSCAR",
    "init/01.scale_pert/sys-0004-0001/scale-1.000/000001/POSCAR",
    "init/01.scale_pert/sys-0004-0001/scale-1.000/00000[2-9]/POSCAR"
  ],
  [
    "init/01.scale_pert/sys-0004-0001/scale-1.000/00001*/POSCAR"
  ]
],
"sys_batch_size": [
  8,
  8
],

"_comment": " 01.model_devi ",
"model_devi_dt": 0.002,
"model_devi_skip": 0,
"model_devi_f_trust_lo":  0.05,
"model_devi_f_trust_hi":  0.15,
"model_devi_clean_traj":  false,
"model_devi_jobs": [
  {
    "sys_idx": [0],"temps": [50],"press": [1],"trj_freq": 10,"nsteps": 1000,"ensemble": "nvt","_idx": "00"
  },
  {
    "sys_idx": [1],"temps": [50],"press": [1],"trj_freq": 10,"nsteps": 3000,"ensemble": "nvt","_idx": "01"
  }
],
```

- 指定exploration中分子动力学模拟的初始结构
- 选取合适的模型力偏差上下限
- 设置md的步长以及相关条件

参数设置文件

```
"_comment": " 02.fp ",
"fp_style":              "vasp",
"shuffle_poscar":        false,
"fp_task_max":           30,
"fp_task_min":           8,
"fp_pp_path":            "/old.data/fqgong/example/fp",
"fp_pp_files":           [ "POTCAR_H","POTCAR_C"],
"fp_incar":              "INCAR_methane"
```

VASP单点能计算

CP2K单点能计算

```
"_comment":          " 02.fp ",
"fp_style":          "cp2k",
"shuffle_poscar":    false,
"fp_task_max":       400,
"_comment":          "the maximum number of stcs to calc.",
"fp_task_min":       5,
"fp_pp_path":        ".",
"fp_pp_files":       [],
"fp_params": {
    "FORCE_EVAL":{
        "DFT":{
            "BASIS_SET_FILE_NAME": "/data/fqgong/common-files/cp2k/BASIS_MOLOPT",
            "POTENTIAL_FILE_NAME": "/data/fqgong/common-files/cp2k/GTH_POTENTIALS",
            "UKS": "T",
            "MULTIPLICITY": 2,
            "MGRID":{
                    "CUTOFF": 400
            },
            "SCF":{
                    "EPS_SCF": 3.0E-7,
                    "MAX_SCF": 50,
                    "OUTER_SCF": {
                            "EPS_SCF": 3.0E-7,
                            "MAX_SCF": 20
                    },
                    "OT": {
                            "MINIMIZER": "CG",
                            "PRECONDITIONER": "FULL_SINGLE_INVERSE",
                            "ENERGY_GAP": 0.1
                    }
            },
            "XC":{
                    "XC_FUNCTIONAL":{"_": "PBE"},
                    "XC_GRID":{
                            "XC_SMOOTH_RHO": "NN50",
                            "XC_DERIV": "NN50_SMOOTH"
                    },
                    "vdW_POTENTIAL":{
                            "DISPERSION_FUNCTIONAL": "PAIR_POTENTIAL",
                            "PAIR_POTENTIAL":{
                                    "TYPE": "DFTD3",
                                    "PARAMETER_FILE_NAME": "/data/fqgong/common-files/cp2k/dftd3.dat",
                                    "REFERENCE_FUNCTIONAL": "PBE"
                            }
                    }
            }
        },
        "SUBSYS":{
                    "KIND":{
                            "_": ["Ag", "O"],
                            "POTENTIAL": ["GTH-PBE-q11", "GTH-PBE-q6"],
                            "BASIS_SET": ["TZV2P-MOLOPT-SR-GTH", "TZV2P-MOLOPT-GTH"]
                    }
        }
    }
}
```

实例演示：甲烷力场构建