

# Ringkasan Jurnal MLatticeABC

*Generic Lattice Constant Prediction of Crystal Materials Using Machine Learning*

**Elviana**

NIM. H021231006

Mata Kuliah: Machine Learning  
Laboratorium Fisika Teori dan Komputasi

## 1. Identitas Jurnal

- **Judul:** *MLATTICEABC: Generic Lattice Constant Prediction of Crystal Materials Using Machine Learning*
- **Penulis:** Yuxin Li, Wenhui Yang, Rongzhi Dong, Jianjun Hu
- **Tahun:** 2020

## 2. Latar Belakang

Konstanta kisi ( $a, b, c$ ) dan sudut sel satuan ( $\alpha, \beta, \gamma$ ) adalah parameter fundamental yang mendeskripsikan struktur periodik kristal. Sifat elektronik, stabilitas fasa, hingga performa material sangat dipengaruhi oleh nilai konstanta kisi.

Metode tradisional seperti XRD dan DFT akurat tetapi mahal serta memakan waktu. Sebagian besar model *machine learning* sebelumnya hanya bekerja untuk keluarga material tertentu sehingga tidak mampu digeneralisasi. Oleh karena itu, dibutuhkan model prediksi yang lebih cepat, akurat, dan dapat digunakan untuk berbagai jenis material.

## 3. Tujuan Penelitian

Penelitian ini bertujuan:

1. Mengembangkan model ML yang generik, bukan hanya spesifik keluarga material tertentu.
2. Menggunakan dataset besar lebih dari 125.000 material dari Materials Project.
3. Mencapai akurasi tertinggi terutama pada sistem kubik, yang hanya memiliki parameter  $a$ .

## 4. Dataset

Dataset diperoleh dari Materials Project berjumlah 125.278 material dengan tujuh sistem kristal: triclinic, monoclinic, orthorhombic, tetragonal, trigonal, hexagonal, dan cubic.

Nilai  $a$ ,  $b$ , dan  $c$  diambil dari berkas CIF menggunakan format *conventional standard*. Distribusi nilai konstanta kisi menunjukkan rentang lebar, mencerminkan keragaman komposisi material.

## 5. Deskriptor dan Fitur

Penulis menggunakan dua kelompok deskriptor:

1. **Enhanced Magpie**, memuat informasi elemental seperti nomor atom, elektronegativitas, jari-jari kovalen, elektron valensi, band gap, dan fraksi unsur.
2. **Element atom no stat**, mencakup statistik jumlah atom dalam sel satuan (min, max, mean, varians, total atom). Fitur total atom terbukti paling berpengaruh pada akurasi prediksi.

## 6. Model Machine Learning

Tiga model dibandingkan:

- **Random Forest (RF)** – performa terbaik.
- **Deep Neural Network (DNN)** – akurat tetapi kompleks.
- **Gaussian Process Regression (GPR)** – akurasi rendah pada dataset besar.

Evaluasi menggunakan 90% data latih dan 10% data uji dengan metrik RMSE, MAE, dan  $R^2$ .

## 7. Hasil Utama

### 7.1 Sistem Kubik

$$R^2 = 0.979, \quad MAE = 0.138, \quad RMSE = 0.416$$

Model RF dengan kombinasi deskriptor penuh berhasil menangkap hubungan nonlinier antara komposisi dan konstanta kisi dengan akurasi sangat tinggi.

## 7.2 Semua Sistem Kristal

Prediksi  $a$  tetap akurat pada sistem heksagonal, trigonal, dan tetragonal ( $R^2 \approx 0.84\text{--}0.89$ ). Akurasi menurun pada sistem monoclinic karena simetri rendah membuat struktur lebih kompleks.

## 7.3 Perbandingan Model

MLatticeABC mengungguli CryspNet dan Roost pada hampir semua sistem kristal, terutama pada kristal kubik.

## 8. Diskusi

Tiga temuan penting:

1. Ukuran dataset besar meningkatkan generalisasi model.
2. Penambahan data dari sistem kristal lain menurunkan akurasi; pelatihan per-sistem lebih efektif.
3. Fitur statistik jumlah atom adalah penentu performa.

## 9. Kesimpulan

MLatticeABC berhasil menghasilkan model generik prediksi konstanta kisi dengan akurasi tinggi. Kombinasi deskriptor komposisi dan statistik atom membuat model ini unggul dari pendekatan sebelumnya dan relevan untuk tujuan *materials discovery*, screening cepat, dan prediksi sifat material.

## Referensi

Dataset dan kode sumber tersedia di repositori resmi: <https://colab.research.google.com/drive/1h86jXWH1xiAHciaHJHM5NFx-L9SVliA9?usp=sharing>