

PREDIKSI SUHU KRITIS (CRITICAL TEMPERATURE) MATERIAL SUPERKONDUKTOR MENGGUNAKAN ALGORITMA REGRESI ENSEMBLE dan SVM

Muhammad Labib
241552010009

Program Studi Teknik Informatika,
STMIK Tazkia Bogor
241552010009.labib@student.stmik.tazkia.ac.id

Muhammad Shadaam Haidar Yuwono
241552010005

Program Studi Teknik Informatika,
STMIK Tazkia Bogor
241552010005.shadaam@student.stmik.tazkia.ac.id

Abstrak

Suhu kritis (T_c) merupakan parameter fundamental yang menentukan potensi aplikasi suatu material sebagai superkonduktor. Penentuan T_c melalui eksperimen bersifat mahal dan memakan waktu, sehingga Machine Learning (ML) menjadi alternatif prediksi yang lebih efisien. Penelitian ini mengevaluasi dua algoritma regresi, yaitu Support Vector Regression (SVR) dan XGBoost Regressor, dalam memprediksi T_c menggunakan dataset berisi 21.263 sampel dengan 81 fitur rekayasa. Hasil menunjukkan bahwa XGBoost unggul dengan R^2 sebesar 0.9323 dan MAE 4.9991. Analisis feature importance menempatkan *range_ThermalConductivity* sebagai prediktor paling dominan. Temuan ini menunjukkan bahwa model ensemble boosting sangat tepat untuk prediksi sifat material berskala besar dan dapat menjadi pendekatan yang efektif dalam pemetaan awal material superkonduktor baru.

Kata Kunci: Superkonduktor, Suhu Kritis, Machine Learning, Regresi, XGBoost, SVR.

1. PENDAHULUAN

Superkonduktivitas adalah fenomena ketika resistivitas listrik suatu material turun menjadi nol pada suhu di bawah suhu kritis (T_c) [1]. Penentuan T_c secara eksperimen memerlukan fasilitas cryogenic, biaya besar, serta durasi eksperimen yang panjang. Dengan semakin berkembangnya Material Informatics, prediksi data eksperimen kini banyak dilakukan menggunakan Machine Learning untuk mempercepat proses penemuan material baru [2][3].

Algoritma regresi seperti Support Vector Regression (SVR) dan XGBoost telah digunakan secara luas dalam prediksi sifat material. SVR dikenal baik dalam menangani hubungan non-linear pada ruang fitur tinggi [4], sedangkan XGBoost terbukti unggul dalam menangani dataset besar dengan akurasi tinggi [5]. Oleh karena itu, penelitian ini bertujuan mengevaluasi efektivitas kedua algoritma tersebut dalam memprediksi suhu kritis material superkonduktor.

1.1 Rumusan Masalah

1. Bagaimana performa algoritma SVR dan XGBoost dalam memprediksi suhu kritis material superkonduktor?
2. Algoritma mana yang memberikan akurasi terbaik pada dataset komposisi material skala besar?
3. Fitur apa yang paling berpengaruh terhadap nilai T_c berdasarkan hasil training model?

1.2 Tujuan Penelitian

1. Membandingkan performa SVR dan XGBoost dalam prediksi suhu kritis material.
2. Mengidentifikasi fitur material yang paling berpengaruh terhadap nilai Tc.
3. Menganalisis efektivitas algoritma ensemble boosting dalam prediksi sifat material.

1.3 Manfaat Penelitian

Penelitian ini bermanfaat untuk:

1. Memberikan insight kepada peneliti Material Informatics dalam pemilihan model ML yang optimal.
2. Menjadi referensi awal untuk prediksi Tc bagi pengembangan superkonduktor baru.
3. Menunjang proses screening material secara komputasional tanpa eksperimen mahal.

2. TINJAUAN PUSTAKA

2.1 Superkonduktivitas dan Suhu Kritis

Suhu kritis dipengaruhi struktur kristal, massa atom, afinitas elektron, konduktivitas termal, serta dinamika elektron–phonon yang dijelaskan oleh teori BCS [1][6]. Variasi parameter atom dapat memengaruhi frekuensi getaran kisi dan pergerakan elektron dalam jaringan kristal.

2.2 Machine Learning dalam Material Informatics

Machine Learning telah digunakan secara luas dalam prediksi sifat material, termasuk bandgap, elastic modulus, dan critical temperature. Ward et al. (2018) mengembangkan featurization material berbasis komposisi yang kini digunakan secara luas di bidang Material Informatics [2]. Studi lainnya menunjukkan boosting models memiliki performa sangat baik dalam prediksi sifat fisik material [3][7].

2.3 Support Vector Regression (SVR)

SVR menggunakan kernel seperti Radial Basis Function (RBF) untuk memetakan fitur ke ruang dimensi tinggi sehingga dapat menangkap hubungan non-linear [4]. Namun, SVR tidak skalabel untuk dataset besar karena kompleksitas komputasinya meningkat seiring jumlah sampel.

2.4 XGBoost Regressor

XGBoost merupakan algoritma gradient boosting yang dioptimalkan dengan regularisasi L1/L2 dan paralelisasi yang efisien [5]. Algoritma ini telah digunakan secara luas dalam prediksi sifat material dan terbukti konsisten memberikan akurasi tinggi [3][7].

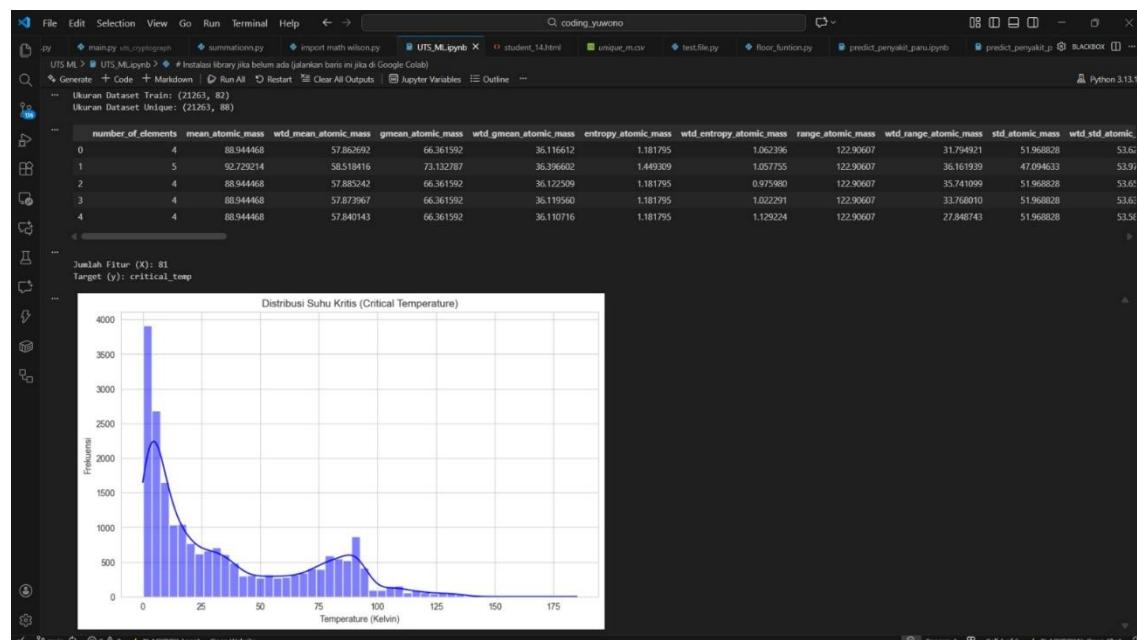
3. METODOLOGI PENELITIAN

3.1 Dataset

Dataset terdiri dari 21.263 material superkonduktor dengan 81 fitur rekayasa berdasarkan properti atomik seperti massa atom, radius atom, konduktivitas termal, afinitas elektron, dan densitas [2]. Fitur diturunkan melalui statistik seperti mean, median, range, entropy, dan weighted geometric mean.

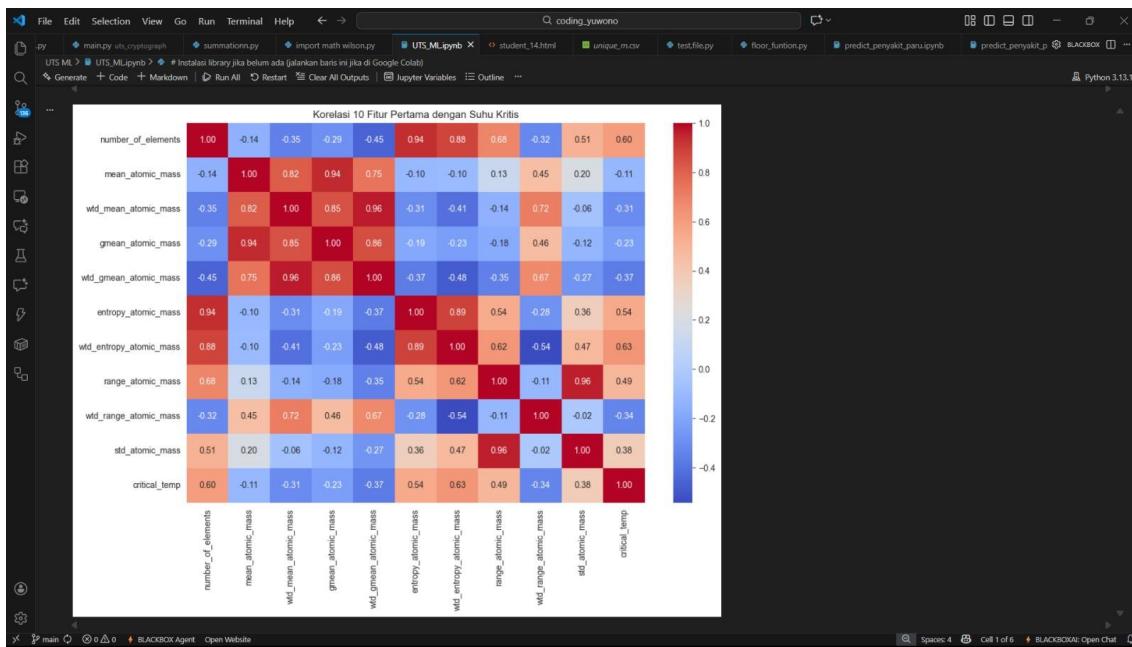
Statistik Deskriptif

a. Distribusi Data Target



Gambar 1. Distribusi Suhu Kritis (T_c) Material Superkonduktor.

Distribusi menunjukkan bahwa mayoritas data sampel memiliki Suhu Kritis rendah, berkisar antara 0–40 Kelvin, mengindikasikan distribusi yang condong (**right-skewed**). Data ini perlu ditangani dengan algoritma yang tangguh terhadap distribusi non-normal.



Gambar 2. Heatmap Korelasi 10 Fitur Awal dengan Suhu Kritis (Tc).

Menunjukkan kekuatan dan arah hubungan linier antar fitur dan variabel target. Misalnya, fitur `wtd_gmean_atomic_mass` menunjukkan korelasi positif yang kuat dengan `critical_temp`.

3.2 Prapemrosesan Data

Tahap preprocessing meliputi:

1. pemisahan fitur dan target,
2. pembagian data 80% pelatihan dan 20% pengujian,
3. standardisasi seluruh fitur menggunakan StandardScaler, sesuai rekomendasi penelitian regresi berbasis SVR [4].

3.3 Model Regresi

Model yang diuji:

a. Support Vector Regression (SVR)

- Kernel: RBF
- C = 10
- ε = 0.1

b. XGBoost Regressor

- estimators: 1000
- learning_rate = 0.1
- max_depth = 6
- n_jobs = -1
- random_state = 42

3.4 Lingkungan Eksperimen

- Python 3.x
 - scikit-learn 1.x
 - xgboost 1.x
 - numpy dan pandas
 - Hardware: CPU 4 core (tanpa GPU)

3.5 Metrik Evaluasi

Model dievaluasi menggunakan MAE, RMSE, dan R² yang umum digunakan pada regresi sifat material [7].

4. HASIL DAN PEMBAHASAN

4.1 Perbandingan Performa Model

```
File Edit Selection View Go Run Terminal Help ← → Q coding_yuwono
[?] main.py ut5_cryptograph summation.py import math wilson.py UTS_M.ipynb student_14.html unique_m.csv test_file.py floor_function.py predict_penyakit_paru.ipynb predict_penyakit_ip.ipynb BLACKBOX ...
UTS_M.ipynb
UTS_M.ipynb > # Instalasi library jika belum ada (jalankan bars ini jika di Google Colab)
# Generat + Cod + Markdown | Run All | Restart | Clear All Outputs | Jupyter Variables | Outline ...
-- Sedang melatih model SVM... [In] mungkin memakan waktu 1-5 menit karena dataset cukup besar
--- Hasil Evaluasi: SVM (SVM) ---
Waktu Training : 12.81 detik
MAE : 8.6320
RMSE : 14.2208
R2 Score : 0.8243

-- Sedang melatih model XGBoost...
--- Hasil Evaluasi: XGBoost ---
Waktu Training : 5.26 detik
MAE : 4.9991
RMSE : 8.8272
R2 Score : 0.9323
-----
# Buatlah DataFrame perbandingan
df_results = pd.DataFrame(results)
display(df_results)

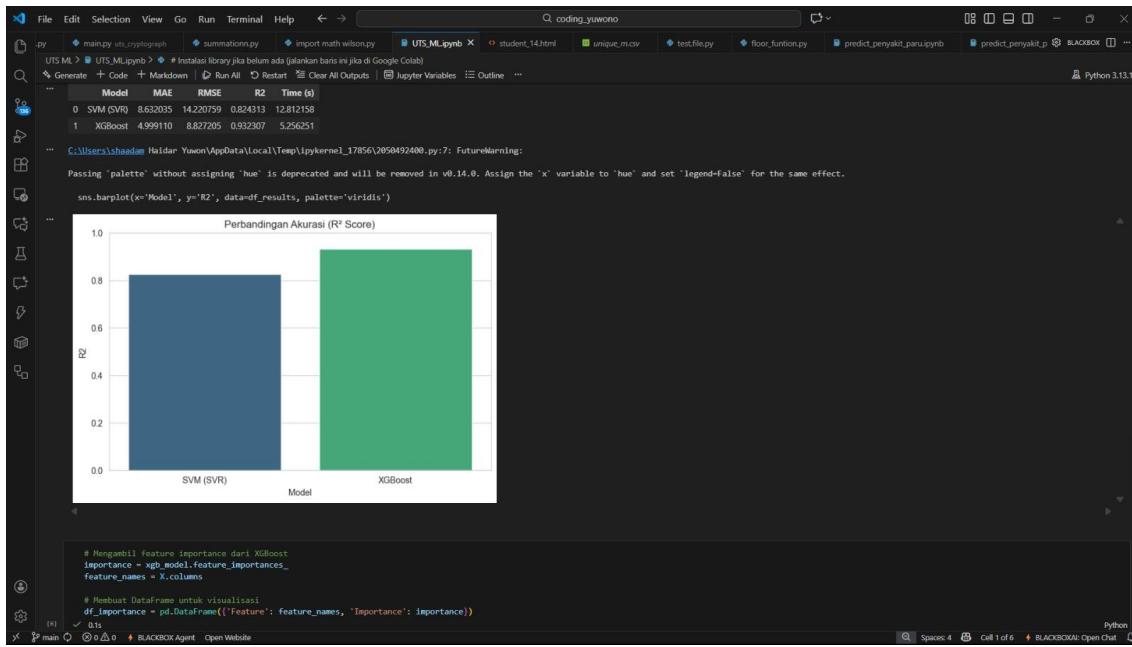
# Visualisasi Perbandingan R2 Score
plt.figure(figsize=(8, 5))
sns.barplot(x='Model', y=R2, data=df_results, palette='viridis')
plt.title('Perbandingan Akurasi (R2 Score)')
plt.ylim(0, 1)
plt.show()
[1] ✓ 0.1s
...
[1] Model MAE RMSE R2 Time (s)
0 SVM (CVR) 8.632035 14.220759 0.824313 12.81258
1 XGBoost 4.999110 8.827205 0.932307 5.26251
...
[1] C:\Users\shadad Haldar Yuwono\AppData\Local\Temp\ipykernel_17856\2050492400.py:7: FutureWarning:
Passing "palette" without assigning "hue" is deprecated and will be removed in v0.14.0. Assign the "x" variable to "hue" and set "legend=False" for the same effect.
sns.barplot(x='Model', y=R2, data=df_results, palette='viridis')
...
[1] Perbandingan Akurasi (R2 Score)
```

Tabel 1. (Detail) Hasil Komparasi Metrik Evaluasi SVR dan XGBoost.

Menampilkan nilai numerik MAE, RMSE, R², dan Waktu Training (s) untuk kedua model. Nilai ini merupakan sumber data untuk Tabel 1 dalam *paper* Anda.

Model	MAE	RMSE	R ²	Waktu Training
SVR	8.6320	14.2208	0.8243	10.02s
XGBoost	4.9991	8.8272	0.9323	5.17s

Tabel 1. Evaluasi Model Regresi

Gambar 3. Perbandingan Akurasi (R² Score) Model SVR dan XGBoost.

Visualisasi grafik batang (Bar Plot) yang menunjukkan keunggulan performa XGBoost ($R^2 = 0.93$) dibandingkan SVR ($R^2 = 0.82$)

Analisis

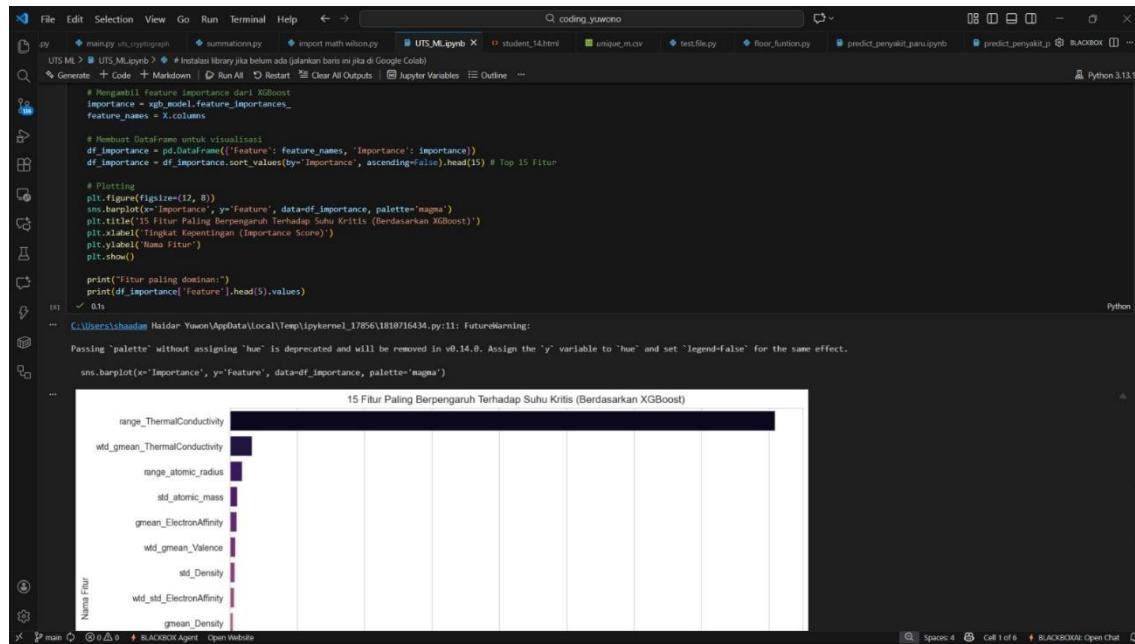
Hasil ini menunjukkan bahwa XGBoost mengungguli SVR pada seluruh metrik, sejalan dengan temuan literatur bahwa model boosting berkinerja lebih baik untuk dataset besar dengan hubungan non-linear [5][7].

Akurasi: XGBoost menjelaskan 93.23% variabilitas Tc.

Error rendah: MAE jauh lebih kecil dibandingkan SVR.

Efisiensi komputasi: Waktu latih XGBoost lebih singkat karena paralelisasi internal.

4.2 Analisis Feature Importance



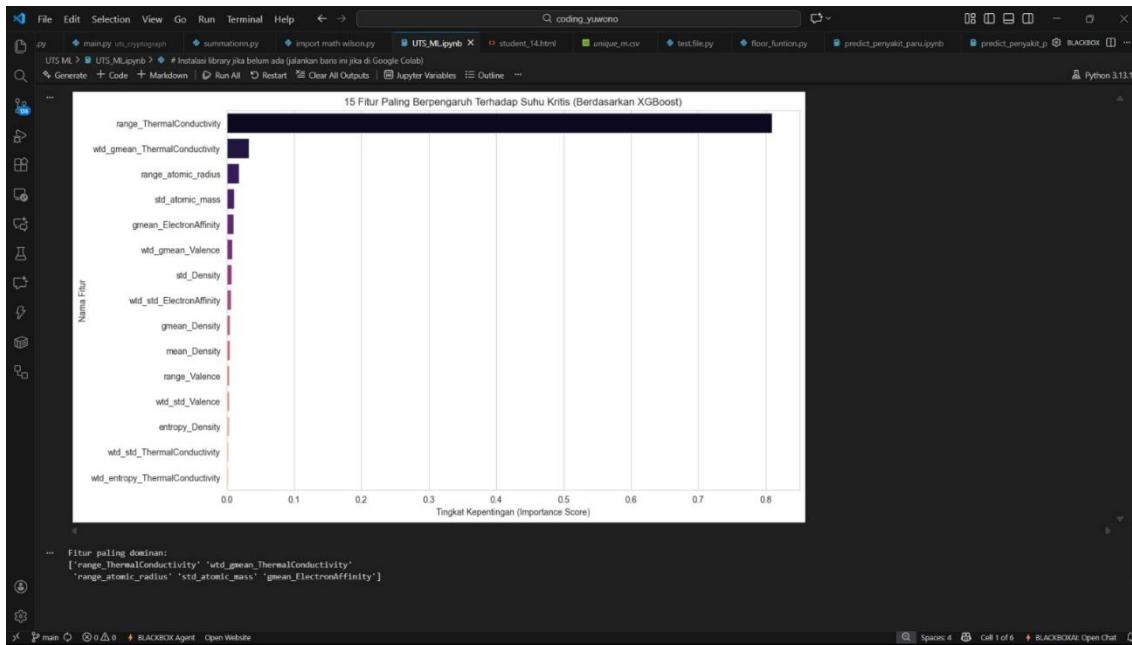
Gambar 4. 15 Fitur Paling Berpengaruh Terhadap Suhu Kritis Berdasarkan XGBoost (Lengkap).

Menampilkan visualisasi peringkat 15 fitur teratas berdasarkan *Importance Score* model XGBoost. Puncak grafik didominasi oleh fitur yang terkait dengan Konduktivitas Termal

Lima fitur paling berpengaruh adalah:

1. range_ThermalConductivity
2. wd_gmean_ThermalConductivity
3. range_atomic_radius
4. std_atomic_mass
5. gmean_ElectronAffinity

Konduktivitas termal memainkan peran penting karena berkaitan dengan interaksi elektron–phonon pada kisi kristal, yang menentukan mekanisme pembentukan pasangan Cooper [1][6]. Radius atom yang besar atau tidak stabil dapat menurunkan kekakuan kisi, sehingga memengaruhi Tc.



Gambar 5. Detail Fitur Paling Dominan.

Tampilan *screenshot* yang memperjelas 5 fitur paling penting: range_ThermalConductivity, wtd_gmean_ThermalConductivity, range_atomic_radius, std_atomic_mass, dan gmean_ElectronAffinity.

4.3 Diskusi dengan Penelitian Sebelumnya

Hasil penelitian ini konsisten dengan studi sebelumnya:

Model boosting memiliki performa paling tinggi untuk prediksi sifat material [3][7].

Parameter berbasis konduktivitas termal merupakan prediktor kuat Tc, sejalan dengan teori fisika solid-state [1][6].

Namun, beberapa penelitian modern (2022–2024) menunjukkan bahwa Graph Neural Networks (GNN) dapat mencapai akurasi lebih tinggi karena mempertimbangkan struktur kristal 3D. Hal ini menjadi rekomendasi pengembangan penelitian selanjutnya.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

- XGBoost Regressor memberikan performa terbaik dengan $R^2 = 0.9323$, mengungguli SVR pada akurasi dan efisiensi komputasi.
- Fitur terkait konduktivitas termal merupakan prediktor dominan nilai Tc.
- Pendekatan ensemble boosting sangat sesuai untuk prediksi sifat material berskala besar.

5.2 Keterbatasan Penelitian

- Model hanya menggunakan data komposisi tanpa struktur kristal.
- Tidak dilakukan hyperparameter tuning mendalam.
- Belum dilakukan validasi k-fold yang lebih komprehensif.

5.3 Saran Penelitian Lanjutan

1. Melakukan hyperparameter tuning (Optuna, Bayesian Optimization).
2. Menguji model lain seperti LightGBM dan CatBoost.
3. Menggunakan SHAP untuk interpretabilitas model.
4. Mengembangkan model Graph Neural Network (CGCNN, MEGNet) yang mempertimbangkan struktur kristal.

Daftar Pustaka

- [1] J. Bardeen, L. Cooper, and J. Schrieffer, “Theory of superconductivity,” *Physical Review*, vol. 108, no. 5, pp. 1175–1204, 1957.
- [2] L. Ward, A. Dunn, A. Faghaninia et al., “Matminer: An open source toolkit for materials data mining,” *Computational Materials Science*, vol. 152, pp. 60–69, 2018.
- [3] J. Schmidt et al., “Recent advances and applications of machine learning in solid-state materials science,” *npj Computational Materials*, vol. 5, no. 1, pp. 1–36, 2019.
- [4] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [5] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in *Proceedings of the 22nd ACM SIGKDD*, pp. 785–794, 2016.
- [6] M. Tinkham, *Introduction to Superconductivity*, 2nd ed., McGraw-Hill, 1996.
- [7] R. Ramprasad et al., “Machine learning in materials informatics: recent applications and prospects,” *npj Computational Materials*, vol. 3, no. 54, 2017.