

LAPORAN ANALISIS PERBANDINGAN ARSITEKTUR MLP BACKPROPAGATION

Eksperimen Multi-Layer Perceptron dengan Berbagai Hidden Layer Configurations
untuk Prediksi Harga Rumah California Housing Dataset

Experiment Type	MLP Architecture Comparison
Dataset	California Housing Dataset
Algorithm	Multi-Layer Perceptron (MLPRegressor)
Activation Function	ReLU
Solver	Adam Optimizer
Max Iterations	2000
Early Stopping	Enabled
Architecture Tested	3 configurations
Date	October 25, 2025

EXECUTIVE SUMMARY

Eksperimen ini mengevaluasi tiga arsitektur MLP yang berbeda untuk prediksi harga rumah. Hasil menunjukkan bahwa **Percobaan 3** dengan arsitektur (16, 16, 32) memberikan performa terbaik dengan RMSE 73713.10, menghasilkan improvement sebesar 3.70% dibandingkan arsitektur terburuk. Semua model berhasil converge tanpa mencapai maximum iterations, menunjukkan stabilitas training yang baik.

1. METODOLOGI EKSPERIMEN

1.1 Setup Eksperimen

Eksperimen ini menggunakan scikit-learn MLPRegressor untuk mengevaluasi performa berbagai arsitektur neural network dalam memprediksi harga rumah. Dataset California Housing digunakan sebagai benchmark dengan 5 fitur input yang telah diseleksi.

1.2 Dataset Processing

- **Dataset:** California Housing Dataset
- **Feature Selection:** 5 fitur numerik terpilih (median_income, total_rooms, total_bedrooms, population, households)
- **Missing Value Handling:** Mean imputation menggunakan SimpleImputer
- **Train-Test Split:** 80%-20% dengan random_state=1
- **Target Variable:** median_house_value

1.3 Model Configuration

Semua model menggunakan konfigurasi yang sama kecuali hidden layer architecture:

Parameter	Value	Description
Activation Function	ReLU	Rectified Linear Unit activation
Solver	Adam	Adaptive moment estimation optimizer
Random State	1	For reproducibility
Max Iterations	2000	Maximum training iterations
Early Stopping	True	Stop when no improvement
N Iter No Change	20	Patience for early stopping
Alpha	Default (0.0001)	L2 regularization parameter

1.4 Arsitektur yang Diuji

Tiga arsitektur MLP yang berbeda diuji untuk mengevaluasi pengaruh kompleksitas terhadap performa:

Percobaan	Architecture	Total Layers	Parameters*	Complexity
Percobaan 1	(16,)	3 layers	~113	Simple
Percobaan 2	(16, 16)	4 layers	~369	Medium
Percobaan 3	(16, 16, 32)	5 layers	~913	Complex

*Parameter count approximate: Input(5) → Hidden layers → Output(1)

2. HASIL EKSPERIMEN

2.1 Overview Hasil

Berikut adalah hasil lengkap dari ketiga eksperimen arsitektur MLP:

Metric	Percobaan 1 (16,)	Percobaan 2 (16,16)	Percobaan 3 (16,16,32)
Training Time (s)	34.06	13.99	20.13
Iterations	1647	461	406
Converged	Yes	Yes	Yes
Convergence Warnings	0	0	0
MAE	55890.88	55158.61	54199.45
MSE	5859245905	5631533633	5433621217
RMSE	76545.71	75043.54	73713.10

2.2 Performance Ranking

Berdasarkan RMSE (Root Mean Square Error) sebagai metric utama:

Rank	Experiment	Architecture	RMSE	Improvement
1	Percobaan 3	(16, 16, 32)	73713.10	Baseline
2	Percobaan 2	(16, 16)	75043.54	1.96%
3	Percobaan 1	(16,)	76545.71	0.00%

3. ANALISIS DAN PEMBAHASAN

3.1 Analisis Performa

Percobaan 3 dengan arsitektur (16, 16, 32) menunjukkan performa terbaik dengan RMSE 73713.10, memberikan improvement sebesar 3.70% dibandingkan Percobaan 1. Trend menunjukkan bahwa arsitektur yang lebih kompleks cenderung memberikan performa yang lebih baik, namun dengan trade-off computational cost yang berbeda.

3.2 Analisis Efisiensi Training

Percobaan 2 menunjukkan training time tercepat (13.99s) sedangkan **Percobaan 1** membutuhkan waktu terlama (34.06s). Menariknya, model tercepat bukan yang paling sederhana, menunjukkan bahwa arsitektur yang tepat dapat mencapai convergence lebih efisien.

3.3 Analisis Convergence

Semua model berhasil converge tanpa mencapai maximum iterations (2000), menunjukkan:

Experiment	Iterations Used	Max Iterations	Convergence %	Status
Percobaan 1	1647	2000	82.3%	✓ Converged
Percobaan 2	461	2000	23.1%	✓ Converged
Percobaan 3	406	2000	20.3%	✓ Converged

3.4 Key Insights

- **Architecture Complexity:** Peningkatan kompleksitas arsitektur memberikan improvement yang konsisten
- **Training Stability:** Semua model menunjukkan training yang stabil tanpa convergence issues
- **Efficiency Trade-off:** Percobaan 2 memberikan balance terbaik antara speed dan accuracy
- **Scalability:** Penambahan layer dan neurons memberikan diminishing returns yang reasonable
- **Generalization:** Tidak ada indikasi overfitting atau instability dalam training process

4. KESIMPULAN DAN REKOMENDASI

4.1 Kesimpulan Utama

Berdasarkan eksperimen yang telah dilakukan, dapat disimpulkan bahwa:

1. **Arsitektur Optimal:** Percobaan 3 dengan konfigurasi (16, 16, 32) memberikan performa terbaik dengan RMSE 73713.10.
2. **Performance Improvement:** Arsitektur terbaik memberikan improvement sebesar 3.70% dibandingkan arsitektur paling sederhana.
3. **Training Stability:** Semua arsitektur menunjukkan convergence yang baik tanpa mencapai maximum iterations, menandakan konfigurasi hyperparameter yang tepat.
4. **Complexity vs Performance:** Terdapat correlation positif antara complexity arsitektur dengan performance, namun dengan diminishing returns.
5. **Computational Efficiency:** Percobaan 2 (16,16) menunjukkan balance terbaik antara training time dan accuracy.

4.2 Rekomendasi

Untuk Production Implementation:

- **Best Performance:** Gunakan arsitektur (16, 16, 32) untuk maximum accuracy
- **Balanced Approach:** Pertimbangkan arsitektur (16, 16) untuk balance speed-accuracy
- **Resource Constrained:** Arsitektur (16,) tetap viable dengan performance yang acceptable

Untuk Further Research:

- Eksplorasi arsitektur yang lebih dalam dengan regularization techniques
- Hyperparameter tuning untuk learning rate, alpha, dan batch size
- Cross-validation untuk validasi yang lebih robust
- Comparison dengan algoritma ensemble methods
- Feature engineering untuk improve overall performance

REKOMENDASI FINAL

Untuk implementasi prediksi harga rumah California Housing:

- **Production Ready:** Gunakan Percobaan 3 - (16, 16, 32)
- **RMSE Target:** 73713.10
- **Training Time:** ~20.1 seconds
- **Reliability:** Converged dalam 406 iterations
- **Confidence Level:** High (no convergence warnings)

5. VISUALISASI HASIL

MLP Architecture Comparison Results

