

**FRAMEWORK HYBRID FEATURE SELECTION  
BERBASIS PCA–ADAPTIVE CORRELATION  
THRESHOLD UNTUK OPTIMASI DAN  
GENERALISASI KLASIFIKASI PADA DATASET  
BREAST CANCER WISCONSIN**

**TESIS**



**PROGRAM STUDI ILMU KOMPUTER (S2)  
PROGRAM PASCASARJANA  
UNIVERSITAS PENDIDIKAN GANESHA**



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

# TESIS

## DIAJUKAN UNTUK MELENGKAPI TUGAS DAN MEMENUHI SYARAT-SYARAT UNTUK MENCAPAI GELAR MAGISTER KOMPUTER



|               |                                                                      |
|---------------|----------------------------------------------------------------------|
| Pembimbing I  | Dr. I Gede Aris Gunadi, S.Si. M.Kom.<br>NIP.197703182008121004       |
| Pembimbing II | Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs.<br>NIP.198307252008011008 |



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

Tesis oleh I Wayan Dodi Putra Artawan telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister Komputer di Ilmu Komputer (S2), Program Pascasarjana, Universitas Pendidikan Ganesha  
Pada tanggal 06 Agustus 2026

### Dewan Penguji

|         |                                                                        |
|---------|------------------------------------------------------------------------|
| Ketua   | Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs.<br>NIP.198307252008011008   |
| Anggota | Dr. I Gede Aris Gunadi, S.Si. M.Kom.<br>NIP.197703182008121004         |
| Anggota | Prof. Dr. Gede Indrawan, S.T., M.T.<br>NIP.197601022003121001          |
| Anggota | Prof. Dr. Luh Joni Erawati Dewi, S.T., M.Pd.<br>NIP.197606252001122001 |

Mengetahui Direktur Program Pascasarjana  
Universitas Pendidikan Ganesha,



Prof. Dr. I Nyoman Jampel, M.Pd.

NIP.195910101986031003



- UU ITE No. 11 Tahun 2008 Pasal 5 Ayat 1 "Informasi Elektronik dan/atau hasil cetaknya merupakan alat bukti hukum yang sah"
- Dokumen ini telah ditandatangani secara elektronik menggunakan sertifikat elektronik yang diterbitkan BSrE - BSSN, validitas dokumen elektronik ini bisa dicek menggunakan aplikasi mobile VeryDS oleh BSrE
- Cetakan dokumen ini merupakan salinan dari file dokumen bertandatangan elektronik yang keabsahannya dapat diakses melalui scan QRCode yang terdapat pada sertifikat ini.

## LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister Pendidikan dari Pascasarjana Universitas Pendidikan Ganesha seluruhnya merupakan hasil karya saya sendiri. Bagian-bagian tertentu dalam penulisan tesis yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila di kemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan yang berlaku di wilayah Negara Kesatuan Republik Indonesia.

Singaraja, 7 Agustus 2026

Yang memberi pernyataan,



I Wayan Dodi Putra Artawan

## PRAKATA

Puji dan syukur penulis panjatkan ke hadapan Tuhan Yang Maha Esa atas segala rahmat dan anugerah-Nya, sehingga tesis yang berjudul “Framework Hybrid Feature Selection Berbasis PCA–Adaptive Correlation Threshold untuk Optimasi dan Generalisasi Klasifikasi pada Dataset Breast Cancer Wisconsin” dapat diselesaikan sesuai dengan yang direncanakan. Proposal tesis ini disusun sebagai salah satu persyaratan untuk memperoleh gelar Magister Komputer pada Program Studi Magister Ilmu Komputer, Program Pascasarjana Universitas Pendidikan Ganesha.

Dalam proses penyusunan proposal tesis ini, penulis menyadari bahwa keberhasilannya tidak terlepas dari dukungan, bantuan, serta bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Dr. I Gede Aris Gunadi, S.Si. M.Kom., selaku Pembimbing I yang dengan penuh kesabaran telah memberikan bimbingan, arahan, serta motivasi yang sangat berarti bagi penulis dalam menghadapi berbagai tantangan selama proses penyusunan proposal tesis ini;
2. Dr. Ir. I Made Gede Sunarya, S.Kom., M.Cs., selaku Pembimbing II yang telah memberikan masukan, arahan, serta dorongan semangat kepada penulis sehingga proposal tesis ini dapat tersusun dengan baik;
3. Prof. Dr. Gede Indrawan, S.T., M.T. dan Prof. Dr. Luh Joni Erawati Dewi, S.T., M.Pd. sebagai penguji yang telah banyak memberikan masukan-masukan yang bermanfaat untuk penyempurnaan tesis ini;
4. Koordinator Program Studi Magister Ilmu Komputer beserta seluruh staf dosen pengajar yang telah memberikan ilmu pengetahuan, bimbingan, serta motivasi selama penulis menempuh pendidikan di Program Pascasarjana Universitas Pendidikan Ganesha;
5. Direktur Program Pascasarjana Universitas Pendidikan Ganesha beserta seluruh staf yang telah memberikan fasilitas dan dukungan selama proses penyusunan proposal tesis ini;

6. Rektor Universitas Pendidikan Ganesha, yang telah memberikan kesempatan serta dukungan kepada penulis untuk menempuh pendidikan pada Program Pascasarjana Universitas Pendidikan Ganesha;
7. Rekan-rekan mahasiswa Program Studi Ilmu Komputer yang telah memberikan dukungan, semangat, serta berbagi pengalaman selama menjalani proses perkuliahan dan penyusunan proposal tesis ini;
8. Keluarga, sahabat, dan berbagai pihak terkait yang telah memberikan dukungan baik secara langsung maupun tidak langsung dalam penyelesaian tesis ini.

Semoga segala bantuan dan dukungan yang telah diberikan oleh berbagai pihak mendapatkan balasan yang setimpal dari Tuhan Yang Maha Esa berupa kesehatan, kebahagiaan, serta keberkahan dalam menjalani kehidupan.

Penulis menyadari bahwa proposal tesis ini masih memiliki keterbatasan dan belum sepenuhnya sempurna. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang konstruktif demi penyempurnaan proposal tesis ini di masa yang akan datang. Penulis berharap proposal tesis ini dapat memberikan kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam bidang machine learning dan data mining, serta bermanfaat bagi masyarakat akademis.

Singaraja, 8 Agustus 2026

Penulis

## DAFTAR ISI

|                                                                       |       |
|-----------------------------------------------------------------------|-------|
| LEMBAR PERSETUJUAN PEMBIMBING .....                                   | ii    |
| LEMBAR PERSETUJUAN TIM PENGUJI .....                                  | iii   |
| LEMBAR PERNYATAAN .....                                               | iv    |
| PRAKATA .....                                                         | v     |
| ABSTRAK .....                                                         | vii   |
| ABSTRACT .....                                                        | viii  |
| DAFTAR ISI .....                                                      | ix    |
| DAFTAR TABEL .....                                                    | xv    |
| DAFTAR GAMBAR .....                                                   | xviii |
| DAFTAR LAMPIRAN .....                                                 | xx    |
| BAB I PENDAHULUAN .....                                               | 1     |
| 1.1 Latar Belakang .....                                              | 1     |
| 1.2 Identifikasi Masalah .....                                        | 4     |
| 1.3 Batasan Masalah .....                                             | 4     |
| 1.4 Rumusan Masalah .....                                             | 5     |
| 1.5 Tujuan Penelitian .....                                           | 5     |
| 1.6 Manfaat Penelitian .....                                          | 6     |
| BAB II LANDASAN TEORI .....                                           | 8     |
| 2.1 Kajian Teori .....                                                | 8     |
| 2.1.1 Artificial Intelligence dan Machine Learning .....              | 8     |
| 2.1.2 Machine Learning dalam Bidang Medis .....                       | 10    |
| 2.1.3 Permasalahan Data Berdimensi Tinggi (High-Dimensional Data) ... | 11    |
| 2.1.4 Feature Selection .....                                         | 12    |

|                                |                                                           |     |
|--------------------------------|-----------------------------------------------------------|-----|
| 2.1.5                          | Feature Extraction .....                                  | 17  |
| 2.1.6                          | Principal Component Analysis (PCA) .....                  | 18  |
| 2.1.7                          | Correlation-Based Feature Selection.....                  | 24  |
| 2.1.8                          | Adaptive Correlation Threshold (ACT).....                 | 26  |
| 2.1.9                          | Hybrid Feature Selection.....                             | 31  |
| 2.1.10                         | Framework Hybrid PCA–Adaptive Correlation Threshold ..... | 35  |
| 2.1.11                         | Random Forest .....                                       | 38  |
| 2.1.12                         | Support Vector Machine (SVM).....                         | 45  |
| 2.1.13                         | k-Nearest Neighbor (k-NN) .....                           | 56  |
| 2.1.14                         | Komparatif Algoritma Klasifikasi.....                     | 61  |
| 2.1.15                         | Breast Cancer Wisconsin Dataset .....                     | 62  |
| 2.1.16                         | Evaluasi Model Klasifikasi .....                          | 70  |
| 2.1.17                         | Tools Penelitian.....                                     | 80  |
| 2.2                            | Penelitian Terdahulu .....                                | 84  |
| 2.3                            | State of the Art dan Research Gap .....                   | 89  |
| 2.3.1                          | State of the Art .....                                    | 89  |
| 2.3.2                          | Research Gap .....                                        | 93  |
| 2.4                            | Kerangka Berpikir.....                                    | 99  |
| 2.5                            | Hipotesis Penelitian.....                                 | 101 |
| 2.5.1                          | Hipotesis Utama ( $H_1$ ) .....                           | 102 |
| 2.5.2                          | Hipotesis Pendukung.....                                  | 102 |
| 2.5.3                          | Hipotesis Statistik.....                                  | 103 |
| BAB III METODE PENELITIAN..... |                                                           | 104 |
| 3.1                            | Desain Penelitian.....                                    | 104 |
| 3.2                            | Dataset Penelitian.....                                   | 106 |
| 3.2.1                          | Breast Cancer Wisconsin Dataset .....                     | 107 |
| 3.2.2                          | Karakteristik Dataset.....                                | 108 |
| 3.2.3                          | Variabel Penelitian .....                                 | 111 |
| 3.3                            | Framework Penelitian .....                                | 114 |

|       |                                                   |     |
|-------|---------------------------------------------------|-----|
| 3.3.1 | Gambaran Umum Framework.....                      | 116 |
| 3.3.2 | Tahapan Framework.....                            | 118 |
| 3.4   | Tahapan Penelitian.....                           | 120 |
| 3.4.1 | Data Collection.....                              | 122 |
| 3.4.2 | Data Preprocessing.....                           | 123 |
| 3.4.3 | Adaptive Correlation Threshold.....               | 124 |
| 3.4.4 | Principal Component Analysis.....                 | 126 |
| 3.4.5 | Hybrid Feature Selection.....                     | 128 |
| 3.4.6 | Classification.....                               | 130 |
| 3.4.7 | Evaluation .....                                  | 131 |
| 3.4.8 | Algoritma Framework Hybrid Feature Selection..... | 132 |
| 3.5   | Algoritma Klasifikasi.....                        | 136 |
| 3.5.1 | Random Forest.....                                | 137 |
| 3.5.2 | Support Vector Machine.....                       | 139 |
| 3.5.3 | k-Nearest Neighbor.....                           | 140 |
| 3.6   | Skenario Eksperimen.....                          | 142 |
| 3.6.1 | Baseline.....                                     | 143 |
| 3.6.2 | Adaptive Correlation Threshold (ACT).....         | 144 |
| 3.6.3 | Principal Component Analysis (PCA).....           | 145 |
| 3.6.4 | Hybrid PCA–ACT.....                               | 146 |
| 3.7   | Parameter Eksperimen .....                        | 149 |
| 3.7.1 | Parameter Dataset.....                            | 150 |
| 3.7.2 | Parameter Preprocessing .....                     | 150 |
| 3.7.3 | Parameter Adaptive Correlation Threshold .....    | 150 |
| 3.7.4 | Parameter Principal Component Analysis .....      | 151 |
| 3.7.5 | Parameter Random Forest.....                      | 151 |
| 3.7.6 | Parameter Support Vector Machine .....            | 151 |
| 3.7.7 | Parameter k-Nearest Neighbor.....                 | 152 |
| 3.7.8 | Parameter Evaluasi.....                           | 152 |

|                                              |                                                                       |     |
|----------------------------------------------|-----------------------------------------------------------------------|-----|
| 3.8                                          | Metode Evaluasi.....                                                  | 153 |
| 3.8.1                                        | Confusion Matrix .....                                                | 154 |
| 3.8.2                                        | Accuracy .....                                                        | 155 |
| 3.8.3                                        | Precision.....                                                        | 155 |
| 3.8.4                                        | Recall.....                                                           | 156 |
| 3.8.5                                        | F1-Score .....                                                        | 156 |
| 3.8.6                                        | Receiver Operating Characteristic (ROC) .....                         | 156 |
| 3.8.7                                        | Area Under Curve (AUC) .....                                          | 156 |
| 3.8.8                                        | Analisis Komparatif .....                                             | 157 |
| 3.9                                          | Metode Analisis Hasil .....                                           | 158 |
| 3.9.1                                        | Analisis Performa Model Klasifikasi .....                             | 160 |
| 3.9.2                                        | Analisis Efektivitas Seleksi Fitur .....                              | 161 |
| 3.9.3                                        | Analisis Generalisasi Framework .....                                 | 162 |
| 3.9.4                                        | Analisis Efisiensi Framework .....                                    | 163 |
| 3.9.5                                        | Analisis Komparatif Keseluruhan .....                                 | 163 |
| BAB IV HASIL PENELITIAN DAN PEMBAHASAN ..... |                                                                       | 165 |
| 4.1                                          | Gambaran Umum Penelitian .....                                        | 165 |
| 4.1.1                                        | Deskripsi Objek Penelitian .....                                      | 165 |
| 4.1.2                                        | Lingkungan Eksperimen .....                                           | 166 |
| 4.1.3                                        | Arsitektur Alur Eksperimen dan Tahapan Penelitian.....                | 167 |
| 4.2                                          | Hasil Tahap Prapemrosesan Data (Preprocessing) .....                  | 168 |
| 4.2.1                                        | Pemeriksaan Struktur Dataset dan Variabel Target .....                | 168 |
| 4.2.2                                        | Transformasi Skala Fitur dan Mitigasi Data Leakage.....               | 169 |
| 4.2.3                                        | Analisis Matriks Korelasi Awal (Pearson Correlation Heatmap)...       | 170 |
| 4.3                                          | Hasil Penerapan Skenario 1: Model Baseline (Tanpa Reduksi Fitur). 171 |     |
| 4.3.1                                        | Karakteristik Ruang Fitur Asli (30 Fitur).....                        | 171 |
| 4.3.2                                        | Performa Klasifikasi Algoritma .....                                  | 172 |
| 4.3.3                                        | Analisis Dampak Multikolinearitas terhadap Kompleksitas Model         | 173 |
| 4.4                                          | Penerapan Skenario 2: Seleksi Fitur Menggunakan (ACT Saja) .....      | 174 |

|       |                                                                     |     |
|-------|---------------------------------------------------------------------|-----|
| 4.4.1 | Penentuan Nilai Ambang Batas Adaptif .....                          | 174 |
| 4.4.2 | Hasil Eliminasi Fitur Redundan (Retensi 13 Fitur).....              | 175 |
| 4.4.3 | Evaluasi Kinerja Klasifikasi Menggunakan Algoritma .....            | 177 |
| 4.5   | Penerapan Skenario 3: Reduksi Dimensi Menggunakan (PCA Saja). 179   |     |
| 4.5.1 | Analisis Dekomposisi Nilai.....                                     | 179 |
| 4.5.2 | Pemilihan Berdasarkan Kriteria Variansi $\geq 90\%$ .....           | 180 |
| 4.5.3 | Visualisasi Proyeksi Spasial (Scree Plot dan Scatter Plot 2D) ..... | 180 |
| 4.5.4 | Evaluasi Kinerja Klasifikasi Menggunakan Algoritma .....            | 182 |
| 4.6   | Penerapan Skenario 4: Framework Hybrid PCA-ACT .....                | 184 |
| 4.6.1 | Arsitektur dan Alur Kerja Framework Hybrid PCA-ACT.....             | 184 |
| 4.6.2 | Hasil Seleksi 8 Komponen Utama .....                                | 185 |
| 4.6.3 | Analisis Rasio Sampel per Fitur dan Curse of Dimensionality.....    | 186 |
| 4.7   | Evaluasi Performa dan Komparasi Menyeluruh Kinerja Klasifikasi..    | 187 |
| 4.7.1 | Analisis Metrik Akurasi (Accuracy) .....                            | 187 |
| 4.7.2 | Analisis Metrik Presisi dan Relevansi Klinis (Precision) .....      | 188 |
| 4.7.3 | Analisis Rasio Sampel per Fitur dan Curse of Dimensionality.....    | 188 |
| 4.7.4 | Analisis Keseimbangan Model Berbasis F1-Score .....                 | 189 |
| 4.7.5 | Evaluasi Kurva ROC dan Daya Diskriminasi Model (AUC-ROC) 190        |     |
| 4.7.6 | Analisis Kompleksitas Waktu Komputasi dan Efisiensi Inferensi .     | 191 |
| 4.8   | Analisis Komparatif Antar-Skenario Pengujian.....                   | 192 |
| 4.8.1 | Komparasi Kuantitatif Jumlah Fitur dan Tingkat Reduksi Dimensi      | 192 |
| 4.8.2 | Komparasi Performa Klasifikasi .....                                | 192 |
| 4.8.3 | Pengujian Menggunakan Paired Sample t-Test .....                    | 193 |
| 4.9   | Pembahasan dan Kontribusi Ilmiah .....                              | 195 |
| 4.9.1 | Interpretasi Hasil Uji Hipotesis dan Pembuktian Performa .....      | 196 |
| 4.9.2 | Pemenuhan Kesenjangan Penelitian (Research Gap) .....               | 199 |
| 4.9.3 | Posisi Performa terhadap Studi Literatur .....                      | 200 |
| 4.9.4 | Penegasan Kebaruan (Novelty) Framework yang Diusulkan .....         | 202 |
| 4.10  | Analisis Ancaman Terhadap Validitas (Threats to Validity).....      | 203 |

|                      |                                               |     |
|----------------------|-----------------------------------------------|-----|
| 4.10.1               | Validitas Internal (Internal Validity).....   | 203 |
| 4.10.2               | Validitas Eksternal (External Validity).....  | 204 |
| 4.10.3               | Validitas Konstruk (Construct Validity) ..... | 206 |
| BAB V PENUTUP.....   |                                               | 208 |
| 5.1                  | Rangkuman .....                               | 208 |
| 5.2                  | Kesimpulan .....                              | 209 |
| 5.3                  | Saran.....                                    | 210 |
| DAFTAR PUSTAKA ..... |                                               | 212 |
| LAMPIRAN.....        |                                               | 217 |



## DAFTAR TABEL

|                                                                                   |     |
|-----------------------------------------------------------------------------------|-----|
| Tabel 2.1 Perbandingan Pendekatan Feature Selection .....                         | 35  |
| Tabel 2.2 Usulan Framework.....                                                   | 38  |
| Tabel 2.3 Hubungan dengan penelitian.....                                         | 38  |
| Tabel 2.4 Fungsi <i>Kernel</i> .....                                              | 50  |
| Tabel 2.5 kelebihan <i>Support Vector Machine</i> .....                           | 51  |
| Tabel 2.6 Kelemahan <i>Support Vector Machine</i> .....                           | 53  |
| Tabel 2.7 Parameter Terpenting Dalam Algoritma k-NN.....                          | 59  |
| Tabel 2.8 Perbandingan Algoritma Klasifikasi.....                                 | 62  |
| Tabel 2.9 Karakteristik <i>Breast Cancer Wisconsin (Diagnostic) Dataset</i> ..... | 65  |
| Tabel 2.10 Informasi Umum Dataset.....                                            | 65  |
| Tabel 2.11 10 Karakteristik Dasar Inti Sel.....                                   | 66  |
| Tabel 2.12 Distribusi Kelas.....                                                  | 67  |
| Tabel 2.13 Kelompok Fitur pada <i>Breast Cancer Wisconsin Dataset</i> .....       | 68  |
| Tabel 2.14 <i>Confusion Matrix</i> .....                                          | 72  |
| Tabel 2.15 Interpretasi Nilai <i>AUC</i> .....                                    | 78  |
| Tabel 2.16 Perbandingan Metrik Evaluasi Klasifikasi.....                          | 79  |
| Tabel 2.17 Library Python yang Digunakan pada Penelitian .....                    | 83  |
| Tabel 2.18 Sintesis Penelitian Terdahulu.....                                     | 85  |
| Tabel 2.19 <i>Research Gap</i> Penelitian.....                                    | 95  |
| Tabel 2.20 Matriks Novelty Penelitian .....                                       | 96  |
| Tabel 3.1 Ringkasan Desain Penelitian.....                                        | 106 |
| Tabel 3.2 Informasi <i>Breast Cancer Wisconsin Dataset</i> .....                  | 108 |
| Tabel 3.3 Kelompok Fitur <i>Breast Cancer Wisconsin Dataset</i> .....             | 110 |
| Tabel 3.4 Variabel Penelitian.....                                                | 113 |
| Tabel 3.5 Tahapan Framework Penelitian .....                                      | 120 |
| Tabel 3.6 Ringkasan <i>Dataset</i> .....                                          | 123 |
| Tabel 3.7 Tahapan Principal Component Analysis.....                               | 128 |
| Tabel 3.8 Skenario Eksperimen .....                                               | 132 |
| Tabel 3.9 <i>Input–Process–Output Framework</i> .....                             | 135 |

|                                                                               |     |
|-------------------------------------------------------------------------------|-----|
| Tabel 3.10 Parameter <i>Random Forest</i> .....                               | 138 |
| Tabel 3.11 Parameter <i>Support Vector Machine</i> .....                      | 140 |
| Tabel 3.12 Parameter <i>k-Nearest Neighbor</i> .....                          | 141 |
| Tabel 3.13 Perbandingan Algoritma Klasifikasi.....                            | 141 |
| Tabel 3.14 Ringkasan Skenario Eksperimen .....                                | 148 |
| Tabel 3.15 Perbandingan Output Setiap Skenario .....                          | 148 |
| Tabel 3.16 Matriks Desain Eksperimen .....                                    | 148 |
| Tabel 3.17 Parameter <i>Dataset</i> .....                                     | 150 |
| Tabel 3.18 Parameter <i>Preprocessing</i> .....                               | 150 |
| Tabel 3.19 Parameter ACT .....                                                | 151 |
| Tabel 3.20 Parameter PCA.....                                                 | 151 |
| Tabel 3.21 Parameter <i>Random Forest</i> .....                               | 151 |
| Tabel 3.22 Parameter SVM.....                                                 | 151 |
| Tabel 3.23 Parameter k-NN .....                                               | 152 |
| Tabel 3.24 Parameter Evaluasi.....                                            | 152 |
| Tabel 3.25 Ringkasan Seluruh Parameter Eksperimen.....                        | 152 |
| Tabel 3.26 Justifikasi Parameter Eksperimen.....                              | 152 |
| Tabel 3.27 <i>Confision Matrix</i> .....                                      | 155 |
| Tabel 3.28 Interpretasi Nilai AUC .....                                       | 157 |
| Tabel 3.29 Indikator Keberhasilan Penelitian.....                             | 158 |
| Tabel 3.30 Rencana Analisis Performa.....                                     | 160 |
| Tabel 3.31 Analisis Jumlah Fitur .....                                        | 161 |
| Tabel 3.32 Analisis Generalisasi.....                                         | 162 |
| Tabel 3.33 Analisis Efisiensi <i>Framework</i> .....                          | 163 |
| Tabel 3.34 Kriteria Keberhasilan Penelitian .....                             | 164 |
| Tabel 4.1 Ringkasan Fitur Utama Berdasarkan Karakteristik Inti Sel.....       | 166 |
| Tabel 4.2 Perangkat Lunak Penelitian .....                                    | 167 |
| Tabel 4.3 Ringkasan Pemeriksaan Struktur Awal Dataset.....                    | 169 |
| Tabel 4.4 Hasil Performa Klasifikasi Skenario <i>Baseline</i> (30 Fitur)..... | 173 |
| Tabel 4.5 Ringkasan Hasil Eliminasi Fitur Skenario ACT .....                  | 175 |
| Tabel 4.6 Pemetaan Fitur .....                                                | 176 |

|                                                                                   |     |
|-----------------------------------------------------------------------------------|-----|
| Tabel 4.7 Hasil Performa Klasifikasi Skenario 2 (ACT Saja - 13 Fitur) .....       | 177 |
| Tabel 4.8 Distribusi Variansi 10 Komponen Utama.....                              | 180 |
| Tabel 4.9 Hasil Performa Klasifikasi Skenario 3 (PCA Saja) .....                  | 182 |
| Tabel 4.10 Perbandingan Tahapan Kompresi Dimensi Antar-Skenario .....             | 186 |
| Tabel 4.11 Komparasi Metrik Akurasi Seluruh Skenario (%).....                     | 188 |
| Tabel 4.12 Komparasi Metrik Presisi ( <i>Precision</i> ) (%).....                 | 188 |
| Tabel 4.13 Komparasi Metrik Recall Dan False Negative (FN).....                   | 189 |
| Tabel 4.14 Perbandingan F1-Score (%) Antar-Skenario dan Algoritma .....           | 190 |
| Tabel 4.15 Komparasi Waktu Komputasi Pelatihan dan Inferensi (Model SVM)<br>..... | 191 |
| Tabel 4.16 Komparasi Kuantitatif Reduksi Dimensi Seluruh Skenario .....           | 192 |
| Tabel 4.17 Komparasi Performa Klasifikasi (SVM Kernel RBF) .....                  | 193 |
| Tabel 4.18 Ringkasan Hasil <i>Paired Sample t-Test</i> Komparasi Model.....       | 198 |
| Tabel 4.19 Posisi Performa Framework terhadap SOTA Literatur Terkini.....         | 201 |
| Tabel 4.20 Perbandingan F1-Score (%) Antar-Skenario dan Algoritma .....           | 204 |



## DAFTAR GAMBAR

|                                                                                 |     |
|---------------------------------------------------------------------------------|-----|
| Gambar 2.1 Alur Hybrid Feature Selection .....                                  | 33  |
| Gambar 2.2 <i>Framework</i> Konseptual <i>Hybrid PCA–ACT</i> .....              | 37  |
| Gambar 2.3 Alur Kerja <i>Random Forest</i> .....                                | 41  |
| Gambar 2.4 Alur Kerja Support Vector Machine .....                              | 47  |
| Gambar 2.5 Konsep <i>Support Vector Machine</i> .....                           | 55  |
| Gambar 2.6 Alur Kerja <i>k-Nearest Neighbor</i> .....                           | 58  |
| Gambar 2.7 Hubungan Struktur Fitur dengan <i>Framework Hybrid PCA–ACT</i> ..... | 70  |
| Gambar 2.8 Ilustrasi <i>ROC Curve</i> .....                                     | 77  |
| Gambar 2.9 5 Tahapan Perkembangan Penelitian .....                              | 90  |
| Gambar 2.10 Kerangka Berpikir.....                                              | 101 |
| Gambar 3.1 Desain Penelitian.....                                               | 104 |
| Gambar 3.2 Struktur Pembentukan Fitur Dataset.....                              | 111 |
| Gambar 3.3 <i>Framework</i> Penelitian.....                                     | 118 |
| Gambar 3.4 Tahapan Penelitian.....                                              | 122 |
| Gambar 3.5 Preprocessing .....                                                  | 124 |
| Gambar 3.6 Diagram ACT .....                                                    | 125 |
| Gambar 3.7 <i>Flowchart Principal Component Analysis</i> .....                  | 127 |
| Gambar 3.8 Diagram Hybrid Feature Selection.....                                | 129 |
| Gambar 3.9 Diagram Klasifikasi.....                                             | 131 |
| Gambar 3.10 Algoritma Framework Penelitian.....                                 | 135 |
| Gambar 3.11 Tahapan Klasifikasi.....                                            | 137 |
| Gambar 3.12 <i>Flowchart Random Forest</i> .....                                | 138 |
| Gambar 3.13 <i>Flowchart SVM</i> .....                                          | 139 |
| Gambar 3.14 <i>Flowchart k-NN</i> .....                                         | 140 |
| Gambar 3.15 Diagram integrasi klasifikasi.....                                  | 141 |
| Gambar 3.16 Skenario Eksperimen.....                                            | 143 |
| Gambar 3.17 Diagram <i>Baseline</i> .....                                       | 144 |
| Gambar 3.18 Diagram ACT .....                                                   | 145 |
| Gambar 3.19 Diagram PCA .....                                                   | 146 |

|                                                               |     |
|---------------------------------------------------------------|-----|
| Gambar 3.20 Diagram Hybrid PCA-ACT .....                      | 147 |
| Gambar 3.21 Parameter Eksperimen.....                         | 149 |
| Gambar 3.22 Alur Evaluasi Penelitian.....                     | 154 |
| Gambar 3.23 Kerangka Analisis Hasil.....                      | 158 |
| Gambar 3.24 Metode Analisis Hasil .....                       | 159 |
| Gambar 3.25 Analisis Seleksi Fitur .....                      | 161 |
| Gambar 3.26 Analisis Generalisasi .....                       | 162 |
| Gambar 3.27 Kerangka Analisis Keseluruhan .....               | 164 |
| Gambar 4.1 Visualisasi Heatmap Kolerasi 30x30 Fitur Asli..... | 171 |
| Gambar 4.2 Scree Plot dan Variansi Kumulatif.....             | 181 |
| Gambar 4.3 Proyeksi Spasial 2D: PC1 vs PC2.....               | 182 |
| Gambar 4.4 Visualisasi Perbandingan Kurva ROC .....           | 190 |
| Gambar 4.5 Proses Evaluasi Paired t-Test .....                | 195 |



## DAFTAR LAMPIRAN

|                                                              |     |
|--------------------------------------------------------------|-----|
| Lampiran 1 Dataset Breast Cancer Wisconsin (Diagnostic)..... | 217 |
| Lampiran 2 Tautan Kode Program.....                          | 218 |

