

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa karena atas berkat dan rahmat-Nya, laporan skripsi berjudul "*PERBANDINGAN PERFORMA RANDOM FOREST DAN SUPPORT VECTOR MACHINE PADA DATA BINER DAN DATA MULTI-CLASS*" dapat diselesaikan dengan baik dan tepat waktu.

Laporan skripsi ini disusun untuk memenuhi persyaratan bagi mahasiswa yang wajib ditempuh sesuai dengan kurikulum Program Studi Matematika Fakultas Sains dan Teknologi, Universitas Pelita Harapan, Tangerang. Skripsi ini juga bermanfaat bagi penulis untuk menerapkan pengetahuan yang telah didapat dan memperoleh pengalaman baru yang tidak dapat diperoleh dari perkuliahan.

Laporan skripsi ini berhasil diselesaikan dengan dukungan dari banyak pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Eric Jobiliong, Ph.D., selaku Dekan Fakultas Sains dan Teknologi.
2. Ibu Dr. Nuri Arum Anugrahati, selaku Wakil Dekan Fakultas Sains dan Teknologi.
3. Bapak Laurence, M.T., selaku Direktur Administrasi dan Kemahasiswaan Fakultas Sains dan Teknologi.
4. Bapak Kie Van Ivanky Saputra, Ph.D., selaku Ketua Program Studi Matematika dan pembimbing skripsi yang telah memberikan masukan dan arahan selama proses perkuliahan.
5. Ibu Lina Cahyadi, S.Si., M.Si., selaku pembimbing skripsi dan juga pembimbing akademik yang terus memberikan bimbingan, arahan, dan dukungan dalam pengerjaan laporan skripsi dan selama masa perkuliahan.
6. Bapak Ferry Vincenttius Ferdinand, S.Si., S.Inf., M.Pd., M.M., selaku dosen Program Studi Matematika yang telah banyak membantu penulis dalam menjalani dunia perkuliahan serta memberikan masukan dalam skripsi ini.
7. Seluruh dosen yang telah mendidik dan mengajarkan berbagai ilmu kepada penulis selama masa perkuliahan.

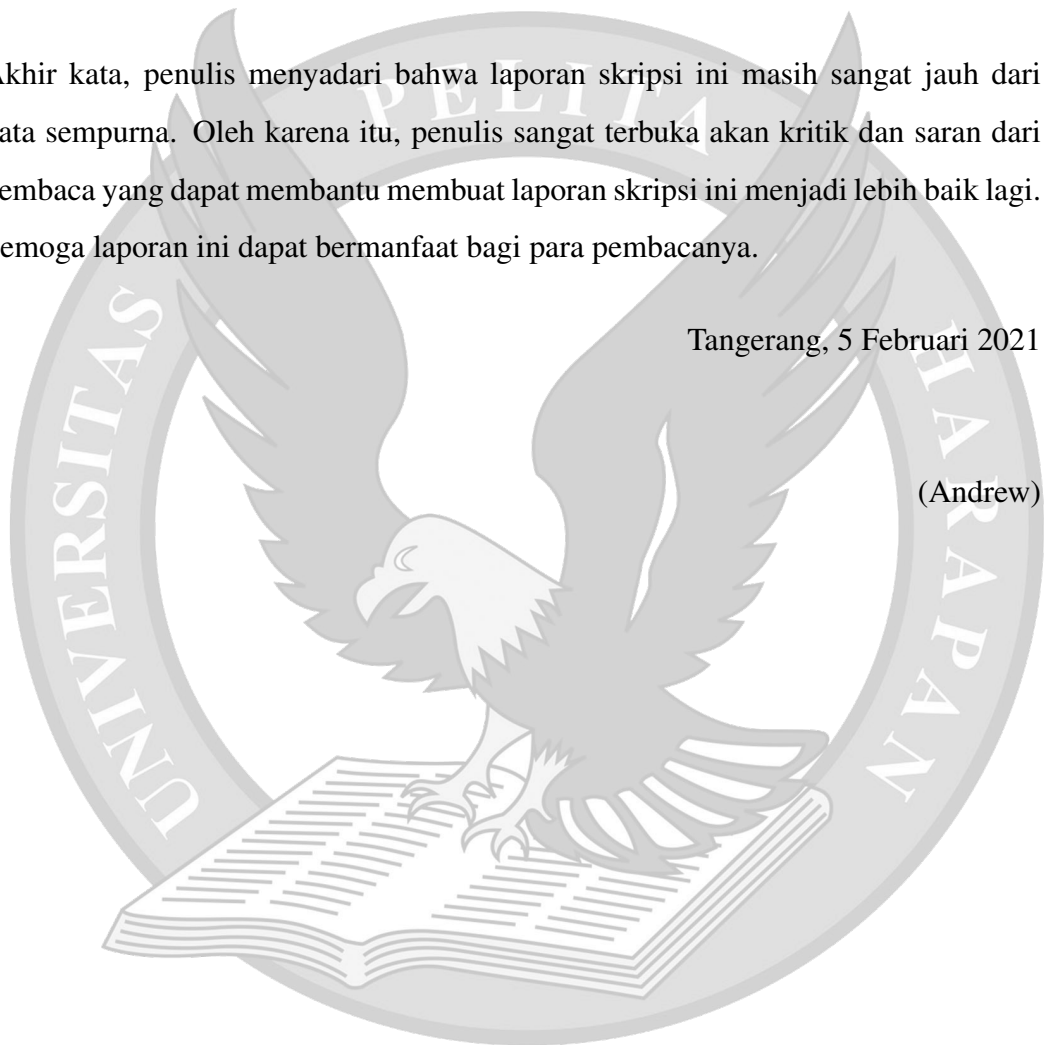
8. Orang tua dan keluarga yang telah memberikan banyak dukungan dan semangat dari awal hingga akhir perkuliahan.
9. Felia yang telah menemani dan mendukung seluruh proses perkuliahan dan penulisan skripsi.
10. Venska, Simon, dan Theda yang telah menjadi tempat mencurahkan isi hati.
11. Crossing River, yaitu Deandra, Feona, Fernando, Inggrid, Kevin, Melyssa, Nicholas, Noviana, dan Yunia yang telah banyak membantu dalam menjalani hari-hari bersama selama perkuliahan.
12. Mathies 1, yaitu Bernath, Lidya, Marcelina, Monic, Nicholas, Vanessa, dan William yang telah menjadi mentee yang sangat membanggakan.
13. Ivana yang telah menemani penulisan skripsi pada masa sulit.
14. Kevin dan Josephine, HM sekaligus partner skripsifightr yang telah menjadi tempat bertukar pikiran dalam menjalani proses mentoring.
15. Baygon, yaitu Averina, Ive, Jose, dan Shella yang telah menjadi rekan dalam melakukan hal yang tidak masuk akal.
16. BPH emlaif, yaitu Jason, Kristian, dan Lavenia, yang telah mau susah bersama dalam mengundang abang.
17. FH1, yaitu Dionysius, Felix, Gabrielle, Hannie, Herlinna, James, Jennifer, Nitta, Sindy, dan Vanesia yang telah melayani bersama dalam mentoring.
18. Mentoring 2019/2020, yaitu Cathrine, Felicia A., Felicia O., Heidy, Ika, Jeany, Jefry, Jessica, Lysandra, Metta, Michael, Novellen, Rosita, dan Rahel yang telah mewarnai masa perkuliahan.
19. HMM-UPH 2018/2019, yaitu Angel, Bella, Calvin, Carlos, Edward, Jessica, Josephine, Licya, Monica, Reynaldi, Stefanie, khususnya anggota bidang eksternal yaitu Cathlyn, Revata, dan Xenia yang jarang membalas grup.
20. HMM-UPH 2017/2018, yaitu Desce, Iyvon, Marisa, Regina, Stella, Winnie, khususnya anggota bidang eksternal yaitu Yoga, Isabel, dan Josephine, yang belum diteraktir ayce.
21. Seluruh anggota kepanitiaan MATRIX, STEALT, HMM Training 2019, dan M-Life 2020 yang telah menerima penulis sebagai ketua yang apa adanya.

22. Matematika 2016 yang terkenal sebagai angkatan yang penuh drama (dramathies).
23. Keluarga Mathies yang telah menjadi keluarga baru selama menjalani dunia perkuliahan.
24. Pihak lain yang telah membantu penulis secara langsung maupun tidak langsung dalam proses perkuliahan dan penulisan skripsi.

Akhir kata, penulis menyadari bahwa laporan skripsi ini masih sangat jauh dari kata sempurna. Oleh karena itu, penulis sangat terbuka akan kritik dan saran dari pembaca yang dapat membantu membuat laporan skripsi ini menjadi lebih baik lagi. Semoga laporan ini dapat bermanfaat bagi para pembacanya.

Tangerang, 5 Februari 2021

(Andrew)



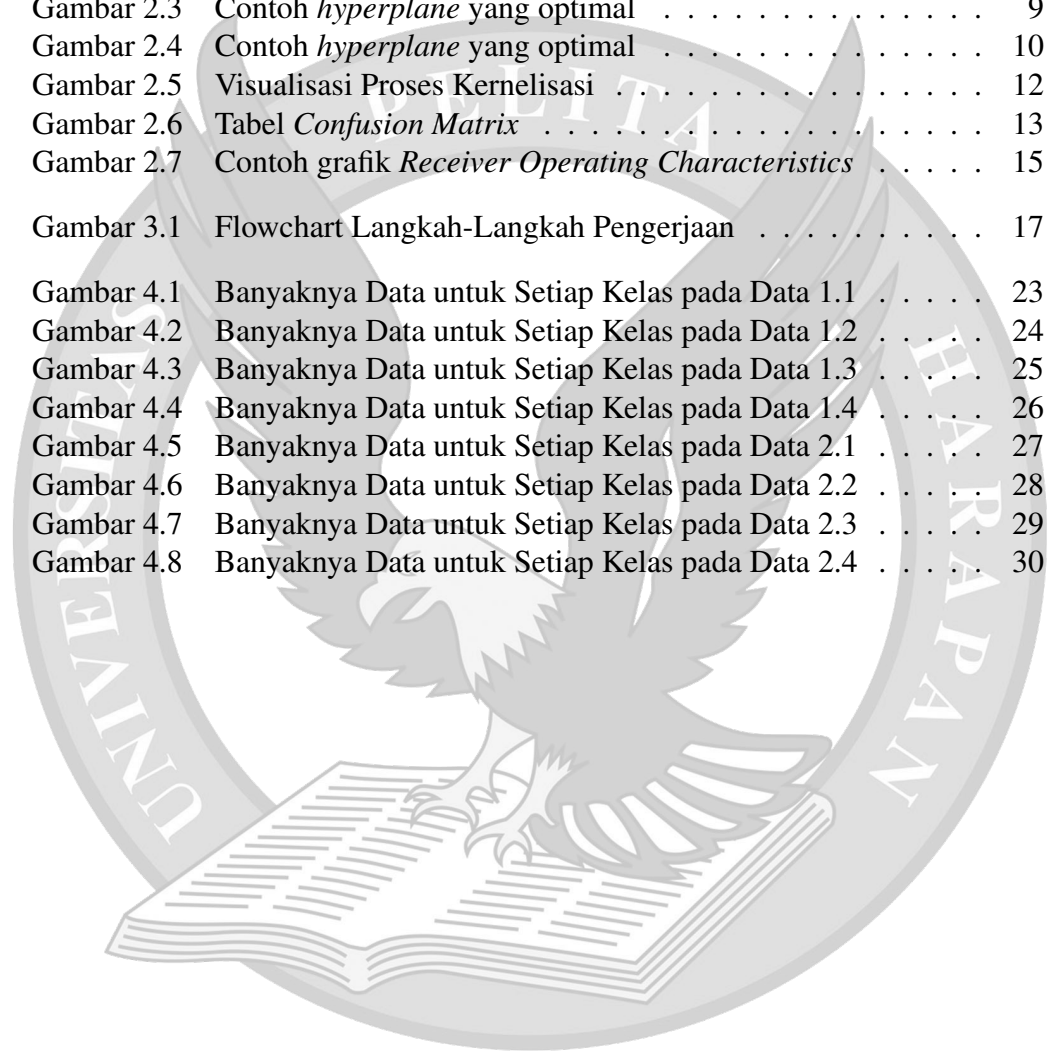
# DAFTAR ISI

	halaman
HALAMAN JUDUL	
PERNYATAAN DAN PERSETUJUAN UNGGAH TUGAS AKHIR	
PERSETUJUAN DOSEN PEMBIMBING SKRIPSI	
PERSETUJUAN TIM PENGUJI SKRIPSI	
ABSTRAK . . . . .	v
<i>ABSTRACT</i> . . . . .	vi
KATA PENGANTAR . . . . .	vii
DAFTAR ISI . . . . .	x
DAFTAR GAMBAR . . . . .	xii
DAFTAR TABEL . . . . .	xiii
DAFTAR LAMPIRAN . . . . .	xv
<b>BAB I PENDAHULUAN</b>	
1.1 Latar Belakang . . . . .	1
1.2 Rumusan Masalah . . . . .	3
1.3 Tujuan Penelitian . . . . .	3
1.4 Batasan Masalah . . . . .	3
1.5 Manfaat Penelitian . . . . .	4
1.5.1 Manfaat Teoritis . . . . .	4
1.5.2 Manfaat Praktis . . . . .	4
1.6 Sistematika Penulisan . . . . .	4
<b>BAB II LANDASAN TEORI</b>	
2.1 <i>Random Forest</i> . . . . .	6
2.1.1 <i>Decision Tree</i> . . . . .	6
2.1.2 <i>Bagging</i> . . . . .	7
2.2 <i>Support Vector Machine</i> . . . . .	8
2.2.1 Maksimalisasi Margin . . . . .	9
2.2.2 Kernelisasi . . . . .	12
2.3 Evaluasi Model Klasifikasi . . . . .	13
2.3.1 <i>Confusion Matrix</i> . . . . .	13
2.3.2 <i>Receiver Operating Characteristics</i> dan <i>Area Under Curve</i> . . . . .	15
2.4 Tinjauan Pustaka . . . . .	16
<b>BAB III METODOLOGI</b>	
3.1 Pengumpulan dan Pembagian Data . . . . .	18
3.2 Pembuatan Model <i>Random Forest</i> . . . . .	18
3.3 Pembuatan Model <i>Support Vector Machine</i> . . . . .	18
3.4 Seleksi Atribut . . . . .	19
3.4.1 Eliminasi Mundur . . . . .	19
3.4.2 Seleksi Maju . . . . .	19

3.5	Pengujian dan Evaluasi Model . . . . .	20
3.6	Penarikan Kesimpulan . . . . .	20
<b>BAB IV ANALISIS DAN PEMBAHASAN</b>		
4.1	Data . . . . .	21
4.1.1	Data Biner . . . . .	21
4.1.2	Data <i>Multi-class</i> . . . . .	26
4.2	Hasil Analisis Data Biner . . . . .	31
4.2.1	Hasil Analisis Data 1.1 . . . . .	31
4.2.2	Hasil Analisis Data 1.2 . . . . .	33
4.2.3	Hasil Analisis Data 1.3 . . . . .	35
4.2.4	Hasil Analisis Data 1.4 . . . . .	37
4.3	Hasil Analisis Data <i>Multi-Class</i> . . . . .	39
4.3.1	Hasil Analisis Data 2.1 . . . . .	39
4.3.2	Hasil Analisis Data 2.2 . . . . .	41
4.3.3	Hasil Analisis Data 2.3 . . . . .	43
4.3.4	Hasil Analisis Data 2.4 . . . . .	45
4.4	Rangkuman Hasil Evaluasi . . . . .	47
4.5	Kepentingan Atribut Pada Model <i>Random Forest</i> . . . . .	48
4.6	Seleksi Atribut <i>Support Vector Machine</i> . . . . .	52
4.7	Perbandingan Kepentingan Atribut Model <i>Random Forest</i> dengan Seleksi Atribut model <i>Support Vector Machine</i> . . . . .	53
4.8	Rangkuman Hasil Evaluasi Setelah Seleksi Atribut . . . . .	55
<b>BAB V PENUTUP</b>		
5.1	Kesimpulan . . . . .	56
5.2	Saran . . . . .	57
<b>DAFTAR PUSTAKA</b>		

# DAFTAR GAMBAR

	halaman
Gambar 1.1 Sejarah perkembangan pengolahan data . . . . .	1
Gambar 1.2 Persentase Upaya Dalam Proses <i>Data Mining</i> . . . . .	2
Gambar 2.1 Penamaan <i>node</i> pada <i>Decision Tree</i> . . . . .	6
Gambar 2.2 Contoh <i>hyperplane</i> pada model SVM . . . . .	8
Gambar 2.3 Contoh <i>hyperplane</i> yang optimal . . . . .	9
Gambar 2.4 Contoh <i>hyperplane</i> yang optimal . . . . .	10
Gambar 2.5 Visualisasi Proses Kernelisasi . . . . .	12
Gambar 2.6 Tabel <i>Confusion Matrix</i> . . . . .	13
Gambar 2.7 Contoh grafik <i>Receiver Operating Characteristics</i> . . . . .	15
Gambar 3.1 Flowchart Langkah-Langkah Pengerjaan . . . . .	17
Gambar 4.1 Banyaknya Data untuk Setiap Kelas pada Data 1.1 . . . . .	23
Gambar 4.2 Banyaknya Data untuk Setiap Kelas pada Data 1.2 . . . . .	24
Gambar 4.3 Banyaknya Data untuk Setiap Kelas pada Data 1.3 . . . . .	25
Gambar 4.4 Banyaknya Data untuk Setiap Kelas pada Data 1.4 . . . . .	26
Gambar 4.5 Banyaknya Data untuk Setiap Kelas pada Data 2.1 . . . . .	27
Gambar 4.6 Banyaknya Data untuk Setiap Kelas pada Data 2.2 . . . . .	28
Gambar 4.7 Banyaknya Data untuk Setiap Kelas pada Data 2.3 . . . . .	29
Gambar 4.8 Banyaknya Data untuk Setiap Kelas pada Data 2.4 . . . . .	30



## DAFTAR TABEL

	halaman
Tabel 4.1 Rangkuman Data Penelitian . . . . .	21
Tabel 4.2 Penjelasan Atribut pada Data 1.1 . . . . .	22
Tabel 4.3 Penjelasan Atribut pada Data 1.2 . . . . .	23
Tabel 4.4 Penjelasan Atribut pada Data 1.3 . . . . .	24
Tabel 4.5 Penjelasan Atribut pada Data 1.4 . . . . .	25
Tabel 4.6 Penjelasan Atribut pada Data 2.1 . . . . .	27
Tabel 4.7 Penjelasan Atribut pada Data 2.2 . . . . .	28
Tabel 4.8 Penjelasan Atribut pada Data 2.3 . . . . .	29
Tabel 4.9 Penjelasan Atribut pada Data 2.4 . . . . .	30
Tabel 4.10 Jumlah Data Pelatihan dan Tes pada Data 1.1 . . . . .	31
Tabel 4.11 Hasil Evaluasi Model Support Vector Machine pada Data 1.1 . . . . .	32
Tabel 4.12 Hasil Evaluasi Model Random Forest pada Data 1.1 . . . . .	32
Tabel 4.13 Rangkuman Hasil Evaluasi Pada Data 1.1 . . . . .	33
Tabel 4.14 Jumlah Data Pelatihan dan Tes pada Data 1.2 . . . . .	33
Tabel 4.15 Hasil Evaluasi Model Support Vector Machine pada Data 1.2 . . . . .	34
Tabel 4.16 Hasil Evaluasi Model Random Forest pada Data 1.2 . . . . .	34
Tabel 4.17 Rangkuman Hasil Evaluasi Pada Data 1.2 . . . . .	35
Tabel 4.18 Jumlah Data Pelatihan dan Tes pada Data 1.3 . . . . .	35
Tabel 4.19 Hasil Evaluasi Model Support Vector Machine pada Data 1.3 . . . . .	36
Tabel 4.20 Hasil Evaluasi Model Random Forest pada Data 1.3 . . . . .	36
Tabel 4.21 Rangkuman Hasil Evaluasi Pada Data 1.3 . . . . .	37
Tabel 4.22 Jumlah Data Pelatihan dan Tes pada Data 1.4 . . . . .	37
Tabel 4.23 Hasil Evaluasi Model Support Vector Machine pada Data 1.4 . . . . .	38
Tabel 4.24 Hasil Evaluasi Model Random Forest pada Data 1.4 . . . . .	38
Tabel 4.25 Rangkuman Hasil Evaluasi Pada Data 1.4 . . . . .	39
Tabel 4.26 Jumlah Data Pelatihan dan Tes pada Data 2.1 . . . . .	39
Tabel 4.27 Hasil Evaluasi Model Support Vector Machine pada Data 2.1 . . . . .	40
Tabel 4.28 Hasil Evaluasi Model Random Forest pada Data 2.1 . . . . .	40
Tabel 4.29 Rangkuman Hasil Evaluasi Pada Data 2.1 . . . . .	41
Tabel 4.30 Jumlah Data Pelatihan dan Tes pada Data 2.2 . . . . .	41
Tabel 4.31 Hasil Evaluasi Model Support Vector Machine pada Data 2.2 . . . . .	42
Tabel 4.32 Hasil Evaluasi Model Random Forest pada Data 2.2 . . . . .	42
Tabel 4.33 Rangkuman Hasil Evaluasi Pada Data 2.2 . . . . .	43
Tabel 4.34 Jumlah Data Pelatihan dan Tes pada Data 2.3 . . . . .	43
Tabel 4.35 Hasil Evaluasi Model Support Vector Machine pada Data 2.3 . . . . .	44
Tabel 4.36 Hasil Evaluasi Model Random Forest pada Data 2.3 . . . . .	44
Tabel 4.37 Rangkuman Hasil Evaluasi Pada Data 2.3 . . . . .	45
Tabel 4.38 Jumlah Data Pelatihan dan Tes pada Data 2.4 . . . . .	45
Tabel 4.39 Hasil Evaluasi Model Support Vector Machine pada Data 2.4 . . . . .	46
Tabel 4.40 Hasil Evaluasi Model Random Forest pada Data 2.4 . . . . .	46

Tabel 4.41	Rangkuman Hasil Evaluasi Pada Data 2.4 . . . . .	47
Tabel 4.42	Rangkuman Hasil Evaluasi Seluruh Data . . . . .	47
Tabel 4.43	Rangkuman Rata-rata Hasil Evaluasi Seluruh Data . . . . .	48
Tabel 4.44	Kepentingan Atribut Random Forest pada Data 1.1 . . . . .	48
Tabel 4.45	Kepentingan Atribut Random Forest pada Data 1.2 . . . . .	49
Tabel 4.46	Kepentingan Atribut Random Forest pada Data 1.3 . . . . .	49
Tabel 4.47	Kepentingan Atribut Random Forest pada Data 1.4 . . . . .	50
Tabel 4.48	Kepentingan Atribut Random Forest pada Data 2.1 . . . . .	50
Tabel 4.49	Kepentingan Atribut Random Forest pada Data 2.2 . . . . .	51
Tabel 4.50	Kepentingan Atribut Random Forest pada Data 2.3 . . . . .	51
Tabel 4.51	Kepentingan Atribut Random Forest pada Data 2.4 . . . . .	52
Tabel 4.52	Hasil Pemilihan Atribut pada <i>Support Vector Machine</i> . . . . .	52
Tabel 4.53	Perbandingan Kepentingan Atribut Model <i>Random Forest</i> dengan Seleksi Atribut model <i>Support Vector Machine</i> pada Data 1.3 . . . . .	53
Tabel 4.54	Perbandingan Kepentingan Atribut Model <i>Random Forest</i> dengan Seleksi Atribut model <i>Support Vector Machine</i> pada Data 2.1 . . . . .	54
Tabel 4.55	Perbandingan Kepentingan Atribut Model <i>Random Forest</i> dengan Seleksi Atribut model <i>Support Vector Machine</i> pada Data 2.4 . . . . .	54
Tabel 4.56	Rangkuman Hasil Evaluasi Seluruh Data Setelah Pemilihan Variabel . . . . .	55
Tabel 4.57	Rangkuman Jumlah Data dan Jumlah Atribut dengan Metode Terbaik . . . . .	55

