

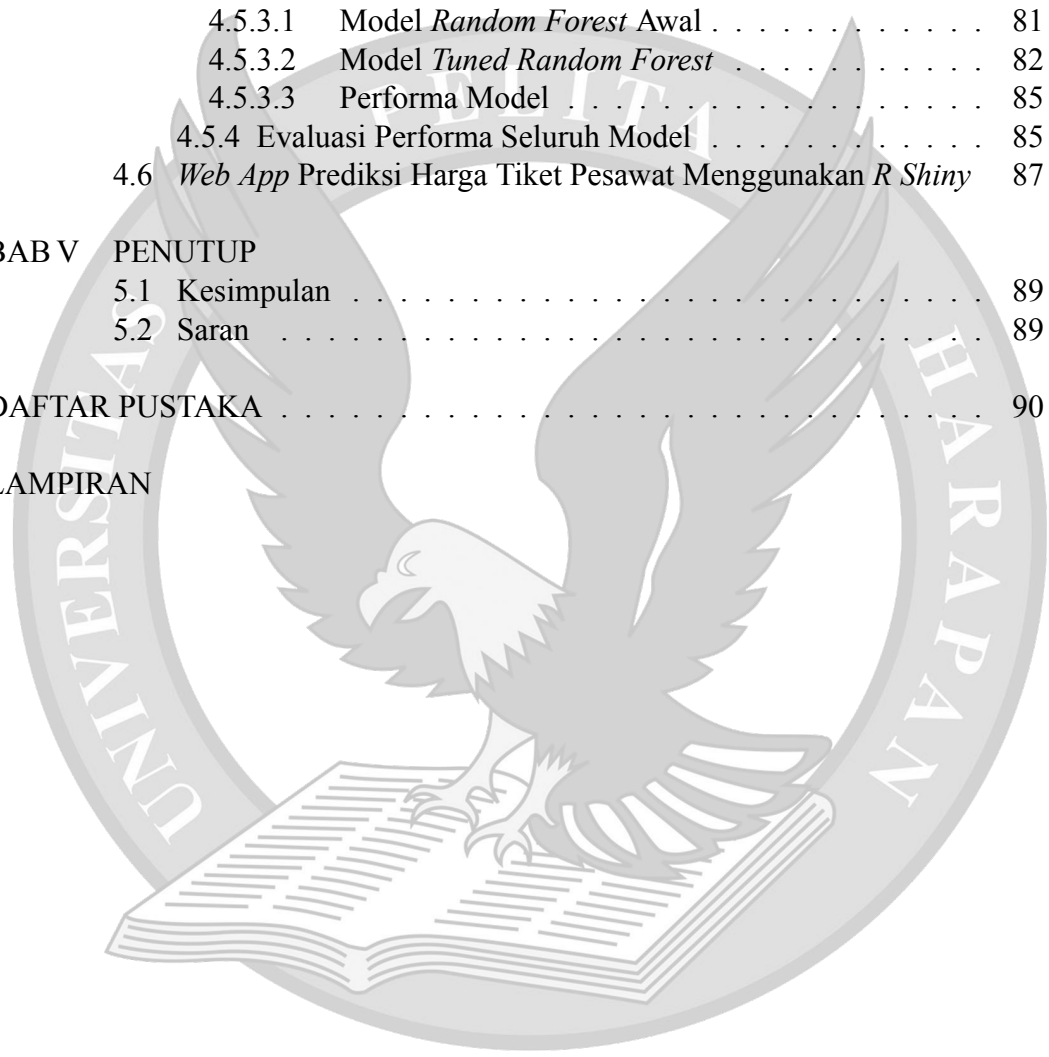
DAFTAR ISI

	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	ix
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR LAMPIRAN	xvi
BABI PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
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
BAB II LANDASAN TEORI	
2.1 <i>Machine Learning</i>	6
2.1.1 <i>Big Data</i>	6
2.1.2 Pengertian <i>Machine Learning</i> dan Masalah yang Dapat Diselesaikan	7
2.1.3 <i>Supervised Learning</i> dan <i>Unsupervised Learning</i>	8
2.1.4 Proses <i>Machine Learning</i>	9
2.2 <i>Tree-Based Models</i>	10
2.2.1 Pengertian <i>Tree-Based Models</i>	10
2.2.2 Terminologi	11
2.2.3 Struktur <i>Tree-Based Models</i>	12
2.3 <i>Regression Tree</i>	13
2.3.1 Proses Membangun Model <i>Regression Tree</i>	13
2.4 <i>The Bias-Variance Trade Off</i>	15
2.5 <i>Overfitting</i> dan <i>Underfitting</i>	17
2.5.1 <i>Overfitting</i>	17
2.5.2 <i>Underfitting</i>	18
2.5.3 <i>Fit</i> yang Baik pada <i>Machine Learning</i>	18

2.6	Meningkatkan Akurasi <i>Regression Tree</i>	19
2.6.1	<i>Pruning</i>	19
2.6.2	<i>Bagging</i>	21
2.6.3	<i>Random Forest</i>	22
2.6.4	<i>Hyperparameter Optimization</i>	23
2.7	<i>k-Fold Cross-Validation</i>	23
2.8	Evaluasi Model	24
2.8.1	<i>Root Mean Squared Error</i>	24
2.8.2	<i>Mean Absolute Error</i>	25
2.8.3	<i>Mean Absolute Percentage Error</i>	25
2.8.4	Koefisien Determinasi	26
2.9	Tinjauan Pustaka	26
BAB III METODOLOGI		
3.1	Pengumpulan Data	29
3.2	Analisis Data	32
3.3	Pembagian Data	32
3.4	Model Prediksi dengan <i>Regression Tree</i> , <i>Bagging Regression Tree</i> , dan <i>Random Forest</i>	33
3.5	Memprediksi Data	35
3.6	Evaluasi Model	35
3.7	<i>Web App</i> dengan <i>R Shiny</i>	36
BAB IV ANALISIS DAN PEMBAHASAN		
4.1	Data	37
4.2	Analisis Data	39
4.2.1	Rute Jakarta-Surabaya (CGK-SUB)	40
4.2.1.1	Harga Tiket Pesawat Berdasarkan Jenis Maskapai Penerbangan pada Rute Jakarta-Surabaya	40
4.2.1.2	Harga Tiket Pesawat Berdasarkan Jam Keberangkatan pada Rute Jakarta-Surabaya	40
4.2.1.3	Harga Tiket Pesawat Berdasarkan Jam Kedatangan pada Rute Jakarta-Surabaya	41
4.2.1.4	Harga Tiket Pesawat Berdasarkan Jumlah Transit pada Rute Jakarta-Surabaya	42
4.2.1.5	Harga Tiket Pesawat Berdasarkan Hari Penerbangan pada Rute Jakarta-Surabaya	43
4.2.1.6	Harga Tiket Pesawat Berdasarkan Hari Libur pada Rute Jakarta-Surabaya	44
4.2.1.7	Harga Tiket Pesawat Berdasarkan Jumlah Hari Tersisa Sebelum Keberangkatan pada Rute Jakarta-Surabaya	45
4.2.1.8	Harga Tiket Pesawat Berdasarkan Durasi Penerbangan pada Rute Jakarta-Surabaya	46

4.2.1.9	Harga Tiket Pesawat Berdasarkan Hari Data Penerbangan Didapatkan pada Rute Jakarta-Surabaya	47
4.2.1.10	Harga Tiket Pesawat Berdasarkan Hari Libur Data Penerbangan Didapatkan pada Rute Jakarta-Surabaya	48
4.2.2	Rute Jakarta-Manado (CGK-MDC)	49
4.2.2.1	Harga Tiket Pesawat Berdasarkan Jenis Maskapai Penerbangan pada Rute Jakarta-Manado	49
4.2.2.2	Harga Tiket Pesawat Berdasarkan Jam Keberangkatan pada Rute Jakarta-Manado	50
4.2.2.3	Harga Tiket Pesawat Berdasarkan Jam Kedatangan pada Rute Jakarta-Manado	51
4.2.2.4	Harga Tiket Pesawat Berdasarkan Jumlah Transit pada Rute Jakarta-Manado	52
4.2.2.5	Harga Tiket Pesawat Berdasarkan Hari Penerbangan pada Rute Jakarta-Manado	53
4.2.2.6	Harga Tiket Pesawat Berdasarkan Hari Libur pada Rute Jakarta-Manado	53
4.2.2.7	Harga Tiket Pesawat Berdasarkan Jumlah Hari Tersisa Sebelum Keberangkatan pada Rute Jakarta-Manado	54
4.2.2.8	Harga Tiket Pesawat Berdasarkan Durasi Penerbangan pada Rute Jakarta-Manado	55
4.2.2.9	Harga Tiket Pesawat Berdasarkan Hari Data Penerbangan Didapatkan pada Rute Jakarta-Manado	56
4.2.2.10	Harga Tiket Pesawat Berdasarkan Hari Libur Data Penerbangan Didapatkan pada Rute Jakarta-Manado	56
4.3	Pembagian Data	57
4.4	Model Prediksi untuk Rute Jakarta-Surabaya	57
4.4.1	<i>Regression Tree</i>	57
4.4.1.1	Model <i>Regression Tree</i> Awal	57
4.4.1.2	Model <i>Tuned Regression Tree</i>	60
4.4.1.3	Performa Model	62
4.4.2	<i>Bagging Regression Tree</i>	62
4.4.2.1	Model <i>Bagging Regression Tree</i>	62
4.4.2.2	Performa Model	65
4.4.3	<i>Random Forest</i>	65
4.4.3.1	Model <i>Random Forest</i> Awal	65
4.4.3.2	Model <i>Tuned Random Forest</i>	66
4.4.3.3	Performa Model	70
4.4.4	Evaluasi Performa Seluruh Model	70

4.5	Model Prediksi untuk Rute Jakarta-Manado	72
4.5.1	<i>Regression Tree</i>	72
4.5.1.1	Model <i>Regression Tree</i> Awal	72
4.5.1.2	Model <i>Tuned Regression Tree</i>	75
4.5.1.3	Performa Model	77
4.5.2	<i>Bagging Regression Tree</i>	78
4.5.2.1	Model <i>Bagging Regression Tree</i>	78
4.5.2.2	Performa Model	80
4.5.3	<i>Random Forest</i>	81
4.5.3.1	Model <i>Random Forest</i> Awal	81
4.5.3.2	Model <i>Tuned Random Forest</i>	82
4.5.3.3	Performa Model	85
4.5.4	Evaluasi Performa Seluruh Model	85
4.6	<i>Web App</i> Prediksi Harga Tiket Pesawat Menggunakan <i>R Shiny</i>	87
BAB V PENUTUP		
5.1	Kesimpulan	89
5.2	Saran	89
DAFTAR PUSTAKA		90
LAMPIRAN		



DAFTAR GAMBAR

	halaman
Gambar 2.1	Komponen Utama <i>Big Data</i> 6
Gambar 2.2	<i>Supervised Learning</i> 8
Gambar 2.3	<i>Unsupervised Learning</i> 9
Gambar 2.4	Proses <i>Machine Learning</i> 9
Gambar 2.5	Proses Segmentasi Data dan Bentuk Diagram Pohon pada <i>Tree-Based Models</i> 11
Gambar 2.6	Struktur Dasar <i>Tree-Based Models</i> 12
Gambar 2.7	<i>Overfitting</i> pada <i>Decision Trees</i> 19
Gambar 3.1	<i>Flowchart</i> Langkah-Langkah Pengerjaan 28
Gambar 3.2	Contoh <i>Splitting</i> Pertama 34
Gambar 4.1	Contoh Data Penerbangan Rute Jakarta-Surabaya 37
Gambar 4.2	Contoh Data Penerbangan Rute Jakarta-Manado 38
Gambar 4.3	Harga Tiket Pesawat Berdasarkan Jenis Maskapai Penerbangan pada Rute Jakarta-Surabaya 40
Gambar 4.4	Harga Tiket Pesawat Berdasarkan Jam Keberangkatan pada Rute Jakarta-Surabaya 41
Gambar 4.5	Harga Tiket Pesawat Berdasarkan Jam Kedatangan pada Rute Jakarta-Surabaya 42
Gambar 4.6	Harga Tiket Pesawat Berdasarkan Jumlah Transit pada Rute Jakarta-Surabaya 43
Gambar 4.7	Harga Tiket Pesawat Berdasarkan Hari Penerbangan pada Rute Jakarta-Surabaya 44
Gambar 4.8	Harga Tiket Pesawat Berdasarkan Hari Libur pada Rute Jakarta-Surabaya 45
Gambar 4.9	Harga Tiket Pesawat Berdasarkan Jumlah Hari Tersisa Sebelum Keberangkatan pada Rute Jakarta-Surabaya 46
Gambar 4.10	Harga Tiket Pesawat Berdasarkan Durasi Penerbangan pada Rute Jakarta-Surabaya 47
Gambar 4.11	Harga Tiket Pesawat Berdasarkan Hari Data Penerbangan Didapatkan pada Rute Jakarta-Surabaya 48
Gambar 4.12	Harga Tiket Pesawat Berdasarkan Hari Libur Data Penerbangan Didapatkan pada Rute Jakarta-Surabaya 49
Gambar 4.13	Harga Tiket Pesawat Berdasarkan Jenis Maskapai Penerbangan pada Rute Jakarta-Manado 50
Gambar 4.14	Harga Tiket Pesawat Berdasarkan Jam Keberangkatan pada Rute Jakarta-Manado 51
Gambar 4.15	Harga Tiket Pesawat Berdasarkan Jam Kedatangan pada Rute Jakarta-Manado 52

Gambar 4.16	Harga Tiket Pesawat Berdasarkan Jumlah Transit pada Rute Jakarta-Manado	52
Gambar 4.17	Harga Tiket Pesawat Berdasarkan Hari Keberangkatan pada Rute Jakarta-Manado	53
Gambar 4.18	Harga Tiket Pesawat Berdasarkan Hari Libur pada Rute Jakarta-Manado	54
Gambar 4.19	Harga Tiket Pesawat Berdasarkan Jumlah Hari Tersisa Sebelum Keberangkatan pada Rute Jakarta-Manado	55
Gambar 4.20	Harga Tiket Pesawat Berdasarkan Durasi Penerbangan pada Rute Jakarta-Manado	55
Gambar 4.21	Harga Tiket Pesawat Berdasarkan Hari Data Penerbangan Didapatkan pada Rute Jakarta-Manado	56
Gambar 4.22	Harga Tiket Pesawat Berdasarkan Hari Libur Data Penerbangan Didapatkan pada Rute Jakarta-Manado	57
Gambar 4.23	<i>Plot Model Regression Tree</i> Awal untuk Rute Jakarta-Surabaya	59
Gambar 4.24	<i>Plot CP Model Regression Tree</i> Awal untuk Rute Jakarta-Surabaya	60
Gambar 4.25	<i>Plot Jumlah Model Pohon</i> terhadap Nilai Error pada Model <i>Bagging Regression Tree</i> untuk Rute Jakarta-Surabaya	63
Gambar 4.26	Urutan <i>Variable Importance</i> pada Model <i>Bagging Regression Tree</i> untuk Rute Jakarta-Surabaya	64
Gambar 4.27	<i>Plot Hasil Parameter Tuning</i> terhadap Parameter <i>mtry</i> pada Model <i>Random Forest</i> untuk Rute Jakarta-Surabaya	68
Gambar 4.28	Visualisasi Performa Model untuk Rute Jakarta-Surabaya	71
Gambar 4.29	Nilai Varians Model Prediksi untuk Rute Jakarta-Surabaya	72
Gambar 4.30	<i>Plot Model Regression Tree</i> Awal untuk Rute Jakarta-Manado	74
Gambar 4.31	<i>Plot CP Model Regression Tree</i> Awal untuk Rute Jakarta-Manado	75
Gambar 4.32	<i>Plot Jumlah Model Pohon</i> terhadap Nilai Error pada Model <i>Bagging Regression Tree</i> untuk Rute Jakarta-Manado	78
Gambar 4.33	Urutan <i>Variable Importance</i> pada Model <i>Bagging Regression Tree</i> untuk Rute Jakarta-Manado	80
Gambar 4.34	<i>Plot Hasil Parameter Tuning</i> terhadap Parameter <i>mtry</i> pada Model <i>Random Forest</i> untuk Rute Jakarta-Manado	83
Gambar 4.35	Visualisasi Performa Model untuk Rute Jakarta-Manado	86
Gambar 4.36	Nilai Varians Model Prediksi untuk Rute Jakarta-Manado	87
Gambar 4.37	<i>Navigation Bar Web App</i> Prediksi Harga Tiket Pesawat	87
Gambar 4.38	<i>Web App</i> Prediksi Harga Tiket Pesawat	88

DAFTAR TABEL

	halaman
Tabel 2.1 Interpretasi Nilai MAPE	26
Tabel 3.1 Data Mentah	29
Tabel 3.2 Data Bersih	30
Tabel 3.3 Variabel pada Data Penerbangan	32
Tabel 3.4 Contoh Data Penerbangan	33
Tabel 3.5 Contoh Pembagian Data Bagian Pertama	33
Tabel 3.6 Contoh Pembagian Data Bagian Kedua	34
Tabel 3.7 Contoh Data Penerbangan dan Hasil Prediksinya	35
Tabel 4.1 Kelas dan Elemen untuk Setiap Variabel pada Data Penerbangan	39
Tabel 4.2 Tingkat Kesalahan Prediksi pada Setiap <i>Terminal Nodes</i> Model <i>Regression Tree</i> untuk Rute Jakarta-Surabaya	62
Tabel 4.3 Performa Model <i>Regression Tree</i> untuk Rute Jakarta-Surabaya	62
Tabel 4.4 Performa Model <i>Bagging Regression Tree</i> untuk Rute Jakarta-Surabaya	65
Tabel 4.5 <i>Parameter Tuning</i> terhadap Parameter <i>ntree</i> pada Model <i>Random Forest</i> untuk Rute Jakarta-Surabaya	68
Tabel 4.6 <i>Parameter Tuning</i> terhadap Parameter <i>maxnodes</i> pada Model <i>Random Forest</i> untuk Rute Jakarta-Surabaya	69
Tabel 4.7 Performa Model <i>Random Forest</i> untuk Rute Jakarta-Surabaya	70
Tabel 4.8 Performa Model untuk Rute Jakarta-Surabaya	70
Tabel 4.9 Nilai Bias dan Varians Model Prediksi untuk Rute Jakarta-Surabaya	71
Tabel 4.10 Tingkat Kesalahan Prediksi pada Setiap <i>Terminal Nodes</i> Model <i>Regression Tree</i> untuk Rute Jakarta-Manado	77
Tabel 4.11 Performa Model <i>Regression Tree</i> untuk Rute Jakarta-Manado	77
Tabel 4.12 Performa Model <i>Bagging Regression Tree</i> untuk Rute Jakarta-Manado	80
Tabel 4.13 <i>Parameter Tuning</i> terhadap Parameter <i>ntree</i> pada Model <i>Random Forest</i> untuk Rute Jakarta-Manado	83
Tabel 4.14 <i>Parameter Tuning</i> terhadap Parameter <i>maxnodes</i> pada Model <i>Random Forest</i> untuk Rute Jakarta-Manado	84
Tabel 4.15 Performa Model <i>Random Forest</i> untuk Rute Jakarta-Manado .	85
Tabel 4.16 Performa Model untuk Rute Jakarta-Manado	85
Tabel 4.17 Nilai Bias dan Varians Model Prediksi untuk Rute Jakarta-Manado	86

DAFTAR LAMPIRAN

	halaman
Lampiran A Data	
A.1 Data Penerbangan Rute Jakarta-Surabaya	A-1
A.2 Data Penerbangan Rute Jakarta-Manado	A-4

