

**PENDEKATAN BARU MENGGUNAKAN *ENSEMBLE BAGGING* PADA
ALGORITMA C4.5 DALAM IDENTIFIKASI KECENDERUNGAN
DEPRESI SITUASIONAL**

Tiya Jayanti
(160411100008)

Dosen Pembimbing 1 : Mula'ab, S.Si., M.Kom.

Dosen Pembimbing 2 : Eka Mala Sari Rochman, S.Kom., M.Kom.

ABSTRAK

Depresi dikategorikan sebagai ancaman kesehatan yang dapat menyebabkan kematian. Hal tersebut umumnya disebabkan oleh banyak faktor, salah satunya adalah karakteristik demografi. Faktor ini dapat memicu kecenderungan depresi situasional pada seseorang akibat dari situasi hidup yang kurang menyenangkan. Depresi situasional dapat terjadi pada masyarakat desa. Namun hingga saat ini, penanganan terkait depresi situasional masih kurang dilakukan. Sedangkan di lini lapangan kurang tersedianya data untuk melakukan klasifikasi kecenderungan depresi situasional. Oleh karena itu, data yang digunakan dalam penelitian ini diperoleh dari sumber data *kaggle*. Metode klasifikasi yang tepat seperti mudah direpresentasikan, tentunya dapat membantu mendeteksi kecenderungan depresi situasional. Penelitian ini menggunakan metode klasifikasi pohon keputusan *Algoritma C4.5* dengan pendekatan *Ensemble Bagging* untuk meningkatkan performa metode klasifikasi, agar dihasilkan akurasi yang lebih baik. Pengujian dilakukan dengan *10-Fold Cross Validation* dan diperoleh hasil akurasi tertinggi 0.84 pada *Ensemble Bagging* sedangkan menggunakan metode klasifikasi tunggal *Algoritma C4.5* diperoleh akurasi tertinggi 0.78. Hal ini menunjukkan bahwa menggunakan pendekatan *Ensemble Bagging* dapat meningkatkan performa metode klasifikasi tunggal *Algoritma C4.5*.

Kata Kunci: Depresi Situasional, Klasifikasi Numerik Diskrit, *Ensemble Bagging*

DAFTAR ISI

JUDUL SKRIPSI	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERNYATAAN ORISINALITAS	iii
ABSTRAK	iv
KATA PENGANTAR.....	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xi
DAFTAR KODE PROGRAM	xii
DAFTAR LAMPIRAN	xiii
DAFTAR GAMBAR LAMPIRAN.....	xiv
DAFTAR TABEL LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah	3
1.2.1 Masalah Penelitian.....	3
1.2.2 Metode Usulan.....	3
1.2.3 Pertanyaan Penelitian.....	3
1.3 Tujuan dan Manfaat Penelitian	4
1.3.1 Tujuan Penelitian	4
1.3.2 Manfaat Penelitian	4
1.4 Batasan-batasan.....	4
1.5 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA.....	6
2.1 Depresi Situasional	6
2.2 <i>Data Mining</i>	9
2.2.1 <i>Preprocessing</i>	10
2.2.2 Proses <i>Mining</i>	12
2.3.3 Evaluasi.....	16

2.3	Penelitian Sebelumnya.....	18
BAB III	METODE USULAN.....	23
3.1	Arsitektur Sistem	23
3.2	Algoritma	25
3.3	<i>Dataset</i>	28
3.4	Tahapan Penelitian.....	29
3.4.1	Studi Literatur	30
3.4.2	Analisis dan Perancangan Sistem	30
3.4.3	Pembuatan Sistem.....	30
3.4.4	Uji Coba Sistem.....	31
3.4.5	Analisis dan Evaluasi.....	31
3.4.6	Dokumentasi	31
3.5	Pengujian.....	31
BAB IV	HASIL DAN PEMBAHASAN	33
4.1	Lingkungan Pengembangan.....	33
4.2	Pembuatan Sistem.....	33
4.2.1	<i>Source Code</i> Metode <i>Ensemble Bagging</i> Pada Pohon Keputusan <i>Algoritma C4.5</i>	35
4.2.2	<i>Source Code</i> Metode Klasifikasi Tunggal Pohon Keputusan <i>Algoritma C4.5</i>	36
4.2.3	Hasil Implementasi Metode <i>Ensemble Bagging</i> pada Pohon Keputusan <i>Algoritma C4.5</i>	36
4.2.4	Hasil Implementasi Metode Klasifikasi Tunggal Pohon Keputusan <i>Algoritma C4.5</i>	39
4.3	Hasil Pengujian	40
4.3.1	Hasil Pengujian <i>Ensemble Bagging</i>	41
4.3.2	Perbandingan Hasil Pengujian <i>Ensemble Bagging</i> pada <i>Algoritma</i> <i>C4.5</i> dan Metode Tunggal <i>Algoritma C4.5</i>	46
4.4	Analisa Hasil Pengujian.....	47
BAB V	PENUTUP	48
5.1	Kesimpulan	48

5.1	Saran	48
REFERENSI		49
LAMPIRAN-LAMPIRAN		56

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Data <i>Mining</i>	10
Gambar 2. 2 Ilustrasi 10-Fold <i>Cross Validation</i>	12
Gambar 2. 3 Skema <i>Bootstrapping</i> teknik <i>Bagging</i>	13
Gambar 3. 1 Diagram IPO	23
Gambar 3. 2 Skema <i>Ensemble Bagging</i>	24
Gambar 3. 3 <i>Flowchart</i> Pelatihan	25
Gambar 3. 4 <i>Flowchart</i> Pengujian	27
Gambar 3. 5 Tahapan Penelitian	30
Gambar 4. 1 Percobaan 1 - <i>Tree Bootstrap</i> 1	37
Gambar 4. 2 Percobaan 1 - <i>Tree Bootstrap</i> 2	38
Gambar 4. 3 Percobaan 1 - <i>Tree Bootstrap</i> 3	39
Gambar 4. 4 Percobaan 1 - <i>Tree C4.5</i>	40

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	17
Tabel 2. 2 Rangkuman Penelitian Sebelumnya	20
Tabel 3. 1 Keterangan <i>Dataset</i>	28
Tabel 4. 1 Lingkungan Pengembangan.....	33
Tabel 4. 2 Hasil Uji Coba Skenario 1	41
Tabel 4. 3 Hasil Uji Coba Skenario 2	41
Tabel 4. 4 Hasil Uji Coba Skenario 3	42
Tabel 4. 5 Hasil Uji Coba Skenario 4	42
Tabel 4. 6 Hasil Uji Coba Skenario 5	43
Tabel 4. 7 Hasil Uji Coba Skenario 6	43
Tabel 4. 8 Hasil Uji Coba Skenario 7	44
Tabel 4. 9 Hasil Uji Coba Skenario 8	44
Tabel 4. 10 Hasil Uji Coba Skenario 9	45
Tabel 4. 11 Hasil Uji Coba Skenario 10	45
Tabel 4. 12 Hasil Uji Coba <i>Ensemble Bagging</i> pada <i>Algoritma C4.5</i>	46
Tabel 4. 13 Uji Coba Metode <i>Algoritma C4.5</i>	46

DAFTAR KODE PROGRAM

Kode Program 4. 1 <i>Source Code</i> Penanganan <i>Missing Value</i>	34
Kode Program 4. 2 <i>Source Code</i> Diskritisasi.....	34
Kode Program 4. 3 <i>Source Code K-Fold Cross Validation</i>	35
Kode Program 4. 4 <i>Source Code Ensemble Bagging</i> pada <i>Algoritma C4.5</i>	36
Kode Program 4. 5 <i>Source Code Algoritma C4.5</i>	36

DAFTAR LAMPIRAN

LAMPIRAN I	Data Depresi Situasional.....	56
LAMPIRAN II	Contoh Perhitungan Manual.....	66
LAMPIRAN III	Kode Program.....	80
LAMPIRAN IV	Hasil Eksperimen.....	85

DAFTAR GAMBAR LAMPIRAN

Gambar 1 Hasil Pohon Keputusan Data <i>Bootstrap</i> 1	78
Gambar 2 Hasil Pohon Keputusan Data <i>Bootstrap</i> 2	78
Gambar 3 Hasil Pohon Keputusan Data <i>Bootstrap</i> 3	79
Gambar 4 <i>Tree Bootstrap</i> 1	104
Gambar 5 <i>Tree Bootstrap</i> 2	112
Gambar 6 <i>Tree Bootstrap</i> 3	121

DAFTAR TABEL LAMPIRAN

Tabel 1 Data 1-100 Dari Fitur Ke-1 Sampai Ke-10	56
Tabel 2 Data 1-100 dari Fitur Ke-11 Sampai Ke-20	61
Tabel 3 Data Awal	66
Tabel 4 Hasil Penanganan Data Kosong	66
Tabel 5 Data Setelah Proses <i>Discretization</i>	67
Tabel 6 Data <i>Bootstrap</i> 1	68
Tabel 7 Data <i>Bootstrap</i> 2	68
Tabel 8 Data <i>Bootstrap</i> 3	68
Tabel 9 Perhitungan <i>Node</i> 1 Data <i>Bootstrap</i> 1	69
Tabel 10 Perhitungan <i>Node</i> 1.1 Data <i>Bootstrap</i> 1	70
Tabel 11 Perhitungan <i>Node</i> 1.1.1 Data <i>Bootstrap</i> 1	71
Tabel 12 Perhitungan <i>Node</i> 1 Data <i>Bootstrap</i> 2	72
Tabel 13 Perhitungan <i>Node</i> 1.1 Data <i>Bootstrap</i> 2	73
Tabel 14 Perhitungan <i>Node</i> 1.1.1 Data <i>Bootstrap</i> 2	74
Tabel 15 Perhitungan <i>Node</i> 1 Data <i>Bootstrap</i> 3	74
Tabel 16 Perhitungan <i>Node</i> 1.1 Data <i>Bootstrap</i> 3	75
Tabel 17 Perhitungan <i>Node</i> 1.1.1 Data <i>Bootstrap</i> 3	76
Tabel 18 Perhitungan <i>Node</i> 1.1.1.1 Data <i>Bootstrap</i> 3	77
Tabel 19 Hasil Prediksi	79
Tabel 20 Hasil Eksperimen Diskritisasi Data	85
Tabel 21 Eksperimen <i>Index</i> Data <i>Test Fold</i> 1	86
Tabel 22 Eksperimen <i>Index</i> Data <i>Test Fold</i> 1	86
Tabel 23 Eksperimen <i>Index</i> Data <i>Test Fold</i> 1	89
Tabel 24 Eksperimen <i>Index</i> Data <i>Test Fold</i> 1	92
Tabel 25 Hasil <i>Majority Voting</i>	122