



**PENGEMBANGAN POHON KLASIFIKASI MENGGUNAKAN
CART ARCING PADA KEJADIAN DIABETES MELITUS
DI RSUD DR.M.M. DUNDA KABUPATEN GORONTALO**

LAPORAN PENELITIAN

OLEH

Dr. RAMA HIOLA, Dra. M.Kes

NIP.19540324198103 2 001

**JURUSAN KESEHATAN MASYARAKAT
FAKULTAS ILMU-ILMU KESEHATAN DAN KEOLAHRAGAAN
UNIVERSITAS NEGERI GORONTALO**

2013

HALAMAN PENGESAHAN

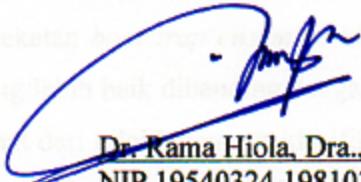
1. a. Judul : Pengembangan Pohon Klasifikasi Menggunakan *CART ARCING* Pada Kejadian Diabetes Mellitus di RSUD Dr.M.M. Dunda Kabupaten Gorontalo
- b. Bidang Ilmu : Terapan Ilmu Kesehatan
- c. Kategori Penelitian : III (Pengembangan Dasar Keilmuan)
2. Pelaksana :
 - a. Nama Lengkap : Dr. Rama Hiola, Dra., M.Kes
 - b. Jenis Kelamin : Perempuan
 - c. Pangkat/Golongan/NIP : Pembina Utama Muda/IVC/19540324 198103 2 001
 - d. Disiplin Ilmu : Biostatistika
 - e. Jabatan Fungsional : Lektor Kepala
 - f. Fakultas/Jurusan : FIKK/Kesehatan Masyarakat
 - g. Perguruan Tinggi : Universitas Negeri Gorontalo
 - h. Alamat Kantor : Kampus III, Jl. Prof. Ario Katili No.44 Gorontalo
 - i. Alamat Rumah : Jl. Jend.Pol.A.Sudjarwo No.222
 - j. No. Telepon : 0435-824528, 085240021922
3. Lokasi Penelitian : RSUD Dr.M.M. Dunda Kabupaten Gorontalo
4. Lama Penelitian : 3 (Tiga) Bulan
5. Biaya yang diperlukan : Rp. 4.000.000,-
6. Sumber Biaya : Dana PNBK Fakultas Tahun 2013

Gorontalo, Oktober 2013

Mengetahui,
Dekan FIKK,

Ketua Peneliti,


Dra. Rany Hiola, M.Kes
NIP 19530913 198302 2 001


Dr. Rama Hiola, Dra., M.Kes
NIP 19540324 198103 2 001

Menyetujui,
Ketua Lembaga Penelitian
Universitas Negeri Gorontalo



Dr. Fitriyane Lihawa, M.Si
NIP 19691209 199303 2 001

DAFTAR ISI

| | |
|---|----------|
| Sampul Luar | |
| Sampul Dalam..... | i |
| Halaman Pengesahan | ii |
| Daftar Isi | iii |
| Daftar Gambar..... | vi |
| Daftar Tabel | vii |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Permasalahan | 4 |
| 1.3 Tujuan Penelitian | 5 |
| 1.3.1 Tujuan Umum..... | 5 |
| 1.3.2 Tujuan Khusus..... | 5 |
| 1.4 Manfaat Penelitian | 5 |
| 1.4.1 Manfaat Teoritis | 6 |
| 1.4.2 Manfaat Praktis..... | 6 |
| BAB II TINJAUAN PUSTAKA | 7 |
| 2.1 Classification and Regression Trees (<i>CART</i>) | 7 |
| 2.1.1 Proses Pembentukan pohon klasifikasi | 9 |
| 2.1.2 Pemangkasan Pohon Klasifikasi | 12 |
| 2.1.3 Pohon Klasifikasi Optimal | 14 |
| 2.2 Adaptive Resampling and Combining (<i>Arcing</i>) | 16 |

Metode statistika yang dapat digunakan untuk pengklasifikasian adalah analisis diskriminan dan regresi logistik. Pendekatan analisis diskriminan harus memenuhi asumsi yaitu data harus berdistribusi normal multivariat dan matrik kovarian yang sama untuk setiap populasi. Metoda regresi logistik memiliki beberapa kelemahan yaitu nilai yang dihasilkan dari model regresi logistik berupa nilai probabilitas dirasa kurang praktis dan memerlukan data yang lengkap, serta sensitif terhadap *outlier*.

Metode lain yang dapat digunakan untuk pengklasifikasian adalah pohon klasifikasi, fleksibel, dapat diaplikasikan pada struktur data apapun, praktis dalam penggunaan dan interpretasi. Untuk melihat seberapa akurat ringkasan suatu data digunakan *bootstrap*. Kajian yang dikembangkan adalah Apakah prosedur pohon klasifikasi yang maksimal dengan melibatkan dugaan parameter dapat dikembangkan melalui *Bootstrap Ensemble*? Tujuan penelitian adalah untuk menunjukkan bahwa *Bootstrap Ensemble* dapat diaplikasikan pada model pohon klasifikasi kasus DM pasien di RSUD Dr. M.M dunda Kabupaten Gorontalo.

Hasil penelitian menunjukkan bahwa *bootstrap* dapat diterapkan pada klasifikasi pohon, dan estimasi *bootstrap ensemble* pohon klasifikasi dilakukan dengan *least square*. Fleksibilitas dari pendekatan *bootstrap ensemble* klasifikasi pohon memberikan *performance* kinerja yang lebih baik dibanding dengan regresi logistik dan diskriminan. Hal ini dapat dilihat dari nilai ketepatan klasifikasi baik data *in-sample* dan *out-sample* pada pengelompokan pasien DM di RSUD. M.M. Dunda Kabupaten Gorontalo. Pasien DM dipengaruhi oleh faktor riwayat keluarga, obesitas, umur dan jenis kelamin, dengan nilai prediksi pada klasifikasi 94,84 persen. Model dengan menggunakan *Ensemble Bagging* Pohon Klasifikasi (50) dan *Ensemble Arcing* Pohon Klasifikasi (150) merupakan model terbaik karena menghasilkan nilai ketepatan klasifikasi paling tinggi 97,80 persen.