

SKRIPSI

**PEMODELAN BANYAKNYA MYCOBACTERIUM
TUBERCULOSIS PADA CITRA HAPUSAN DAHAK DENGAN
PENDEKATAN MODEL REGRESI BINOMIAL NEGATIF
BERDASARKAN ESTIMATOR POLINOMIAL LOKAL
DENGAN PEREDUKSI DIMENSI *GENETIC ALGORITHM***



BIYOSO PRADYO PURNOMO

**PROGRAM STUDI S-1 STATISTIKA
DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA
SURABAYA**

2022

SKRIPSI

**PEMODELAN BANYAKNYA MYCOBACTERIUM
TUBERCULOSIS PADA CITRA HAPUSAN DAHAK DENGAN
PENDEKATAN MODEL REGRESI BINOMIAL NEGATIF
BERDASARKAN ESTIMATOR POLINOMIAL LOKAL
DENGAN PEREDUKSI DIMENSI *GENETIC ALGORITHM***



BIYOSO PRADYO PURNOMO

**PROGRAM STUDI S-1 STATISTIKA
DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA
SURABAYA**

2022

**PEMODELAN BANYAKNYA MYCOBACTERIUM TUBERCULOSIS
PADA CITRA HAPUSAN DAHAK DENGAN PENDEKATAN MODEL
REGRESI BINOMIAL NEGATIF BERDASARKAN ESTIMATOR
POLINOMIAL LOKAL DENGAN PEREDUKSI DIMENSI *GENETIC*
*ALGORITHM***

SKRIPSI

Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana Statistika
pada Fakultas Sains dan Teknologi
Universitas Airlangga

Oleh:

BIYOSO PRADNYO PURNOMO
NIM. 081711833027

Disetujui Oleh:

Pembimbing I



Dr. Nur Chamidah, M.Si
NIP. 197206021998022001

Pembimbing II



Dr. Toha Saifudin, S.Si, M.Si.
NIP. 197501061999031002

LEMBAR PENGESAHAN NASKAH SKRIPSI

Judul : Pemodelan Banyaknya *Mycobacterium Tuberculosis* Pada
Citra Hapusan Dahak Dengan Pendekatan Model Regresi
Binomial Negatif Berdasarkan Estimator Polinomial Lokal
Dengan Pereduksi Dimensi *Genetic Algorithm*

Penyusun : Biyoso Pradnyo Purnomo

NIM : 081711833027

Pembimbing I : Dr. Nur Chamidah, M.Si.

Pembimbing II : Dr. Toha Saifudin, S.Si, M.Si.

Tanggal Seminar : 27 Januari 2022

Disetujui oleh:

Pembimbing I,



Dr. Nur Chamidah, M.Si.

NIP. 197206021998022001

Pembimbing II,



Dr. Toha Saifudin, S.Si, M.Si

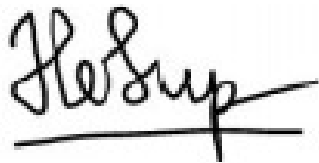
NIP. 197501061999031002

Mengetahui,

Ketua Departemen Matematika

Fakultas Sains dan Teknologi

Universitas Airlangga



Dr. Herry Suprajitno, S.Si., M.Si.

NIP. 196804041994031020

Koordinator Program Studi S1 Statistika

Fakultas Sains dan Teknologi

Universitas Airlangga



Drs. H. Sediono, M.Si.

NIP. 196107121987011001

PEDOMAN PENGGUNAAN SKRIPSI

Skripsi ini tidak dipublikasikan, namun tersedia di perpustakaan dalam lingkungan Universitas Airlangga, diperkenankan untuk dipakai sebagai referensi kepustakaan, tetapi pengutip harus seizin penulis dan harus menyebutkan sumbernya sesuai kebiasaan ilmiah. Dokumen skripsi ini merupakan hak milik Universitas Airlangga.

SURAT PERNYATAAN TENTANG ORISINALITAS

Yang bertanda tangan di bawah ini, saya:

Nama : Biyoso Pradnyo Purnomo

NIM : 081711833027

Program Studi : Statistika

Fakultas : Sains dan Teknologi

Jenjang : Sarjana (S1)

Menyatakan bahwa saya tidak melakukan tindakan plagiat dalam penulisan skripsi saya yang berjudul

**PEMODELAN BANYAKNYA MYCOBACTERIUM
TUBERCULOSIS PADA CITRA HAPUSAN DAHAK DENGAN
PENDEKATAN MODEL REGRESI BINOMIAL NEGATIF
BERDASARKAN ESTIMATOR POLINOMIAL LOKAL
DENGAN PEREDUKSI DIMENSI *GENETIC ALGORITHM***

Apabila suatu saat nanti terbukti melakukan tindakan plagiat, maka saya akan menerima sanksi yang telah ditetapkan. Demikian surat pernyataan ini saya buat dengan sebenar-benarnya.

Makassar, 28 Januari 2022

Penulis,



Biyoso Pradnyo Purnomo

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat dan rahmat-Nya penulis dapat menyelesaikan skripsi yang berjudul “Pemodelan Banyaknya Mycobacterium Tuberculosis Pada Citra Hapusan Dahak Dengan Pendekatan Model Regresi Binomial Negatif Berdasarkan Estimator Polinomial Lokal Dengan Pereduksi Dimensi *Genetic Algorithm*” dengan baik.

Atas penyelesaian skripsi ini, penulis menyampaikan terima kasih kepada:

1. Kedua orang tua beserta seluruh keluarga yang senantiasa memberi dukungan dalam bentuk doa maupun motivasi.
2. Dr. Nur Chamidah M.Si selaku dosen pembimbing I dan Dr. Toha Saifuddin S. Si., M. Si., selaku dosen pembimbing II yang membimbing dan memberi saran dalam penyusunan skripsi ini.
3. Teman-teman statistika angkatan 2017 dan semua pihak yang tidak bisa disebutkan satu persatu.
4. Bethari Putri Purnama, adik khayalan dari dunia seberang yang setia menemani ketika susah maupun senang.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, sehingga penulis mengharapkan kritik dan saran dari semua pihak.

Makassar, 28 Januari 2022

Penulis,

Biyoso Pradnyo Purnomo

Biyoso Pradnyo Purnomo, 2022. **Pemodelan Banyaknya *Mycobacterium Tuberculosis* Pada Citra Hapusan Dahak Dengan Pendekatan Model Regresi Binomial Negatif Berdasarkan Estimator Polinomial Lokal Dengan Pereduksi Dimensi *Genetic Algorithm***. Skripsi dibawah bimbingan Dr. Nur Chamidah M.Si dan Dr. Toha Saifuddin S. Si., M. Si., Program Studi S1-Statistika, Departemen Matematika, Fakultas Sains dan Teknologi, Universitas Airlangga, Surabaya.

ABSTRAK

Tuberkulosis merupakan penyakit menular akut yang disebabkan oleh bakteri *Mycobacterium tuberculosis*. Salah satu cara untuk menanggulangi penyakit Tuberkulosis adalah dengan pendeteksian kasus sedini mungkin. Hal ini dimaksudkan untuk mencegah penularan dari penderita terhadap orang lain dan meningkatkan keefektifan pengobatan bagi sang penderita. Metode utama untuk mendeteksi Tuberkulosis adalah dengan mendeteksi bakteri penyebab penyakit tersebut di dalam dahak penderitanya. Sayangnya, metode ini memiliki sensitivitas yang rendah, bahkan WHO sendiri mencatat bahwa secara global, hanya 45,6% dari kasus baru Tuberkulosis yang terdeteksi positif melalui metode ini. Penelitian ini bertujuan untuk membuat model yang mampu mengidentifikasi banyaknya *Mycobacterium tuberculosis* pada sampel dahak menggunakan pendekatan model regresi binomial negatif berdasarkan estimator polinomial lokal. Tahapannya dimulai dengan pengolahan citra, kemudian reduksi dimensi menggunakan *Genetic Algorithm*, kemudian mengestimasi banyaknya *Mycobacterium tuberculosis* menggunakan model regresi binomial negatif berdasarkan estimator polinomial lokal. Hasil yang didapatkan adalah model dengan orde 212 menghasilkan hasil estimasi yang paling mendekati kenyataannya. Hal ini didukung dengan nilai MAPE terkecil diantara model lainnya, yaitu sebesar 4,03%.

Kata Kunci : *Tuberkulosis, Mycobacterium tuberculosis, Pengolahan Citra, Genetic Algorithm, Regresi Binomial Negatif, Estimator Polinomial Lokal*

Biyoso Pradnyo Purnomo, 2022. **Modeling the Number of Mycobacterium Tuberculosis in Sputum Smear Image Using Negative Binomial Regression Model Approach Based on Local Polynomial Estimator with Dimension Reduction Genetic Algorithm.** This final project is under supervised by Dr. Nur Chamidah M.Si and Dr. Toha Saifuddin S. Si., M. Si., Bachelor Program of Statistics, Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya.

ABSTRACT

Tuberculosis is an acute infectious disease caused by *Mycobacterium tuberculosis*. A common method to deal with tuberculosis is to detect cases as early as possible. The main method for detecting tuberculosis is by identifying the number of *Mycobacterium tuberculosis* in sputum. Unfortunately, this method has low sensitivity and the results depend on the technician. WHO itself noted that globally, only 45.6% of new tuberculosis cases were detected positive through this method. This study aims to create a model that is capable to identify the number of *Mycobacterium tuberculosis* in sputum samples using negative binomial regression model approach based on local polynomial estimator. The steps begin with image processing, then dimension reduction using Genetic Algorithm, then estimating the number of *Mycobacterium tuberculosis* using negative binomial regression model based on a local polynomial estimator. The results obtained are the model with the ordo of 212 produces estimation results that are closest to the reality. This is supported by the smallest MAPE among other models, which is 4.03%.

Keywords : *Tuberculosis, Mycobacterium tuberculosis, Image Processing, Genetic Algorithm, Negative Binomial Regression, Local Polynomial Estimator*

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR PENGESAHAN	iv
PEDOMAN PENGGUNAAN SKRIPSI	v
SURAT PERNYATAAN TENTANG ORISINALITAS	vi
KATA PENGANTAR	vii
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	6
1.4 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Tuberkulosis	7
2.2 Diagnosa Tuberkulosis	7
2.3 <i>Digital Image Processing</i>	10
2.4 <i>Genetic Algorithm</i>	13
2.4.1 Inisiasi Parameter	14
2.4.2 Fungsi Evaluasi	15
2.4.3 Seleksi	16
2.4.4 Proses <i>Crossover</i>	17
2.4.5 Proses Mutasi	18
2.5 Model Respon <i>Count</i>	19
2.6 Regresi Poisson	20
2.7 Penyimpangan Asumsi pada Regresi Poisson	21
2.8 Fungsi Gamma dan Fungsi Beta	22
2.8 Regresi Binomial Negatif	24
2.9 Regresi Nonparametrik	28

2.10 Fungsi Kernel dan Kernel Multivariat	29
2.11 Estimator Polinomial Lokal	34
2.12 Model Aditif	36
2.13 Metode <i>Locally Weighted Likelihood</i>	36
2.14 Metode <i>Nelder-Mead</i>	37
2.15 Pemilihan <i>Bandwidth Optimal</i>	39
2.16 Statistik Uji <i>Deviance</i>	40
2.17 MAPE	41
2.18 <i>Software</i> Python	41
2.19 Bahasa Pemrograman R	42
BAB III METODE PENELITIAN	43
3.1 Data dan Sumber Data	43
3.2 Variabel Penelitian	43
3.3 Langkah-langkah Analisis Data	43
BAB IV PEMBAHASAN	52
4.1 Pengolahan Citra Hapusan Dahak	52
4.1.1 Pembacaan Data Citra Hapusan Dahak	52
4.1.2 Proses <i>Grayscale</i>	53
4.1.3 Proses Ekualisasi	54
4.1.4 Proses Penajaman	54
4.1.5 Proses <i>Multi-Otsu Thresholding</i>	55
4.1.6 Proses Segmentasi	56
4.1.7 Proses Normalisasi	56
4.1.8 Penyimpanan Hasil Pengolahan Citra	57
4.2 Reduksi Dimensi Menggunakan <i>Genetic Algorithm</i>	58
4.2.1 Inisialisasi Parameter	58
4.2.2 Membangkitkan Populasi Awal	58
4.2.3 Evaluasi	59
4.2.4 Seleksi	60
4.2.5 Proses <i>Crossover</i>	62
4.2.6 Proses Mutasi	66
4.2.7 Membentuk Populasi Baru	68

4.2.8 Hasil Reduksi Dimensi.....	72
4.3 Pengestimasian dan Analisis Model Data Hasil Reduksi Dimensi	73
4.2.1 Pendeteksian Overdispersi	74
4.2.2 Pemodelan Menggunakan Metode Parametrik	74
4.2.3 Pemodelan Menggunakan Metode Nonparametrik.....	76
4.2.4 Validasi dan Analisis Model	78
BAB V KESIMPULAN DAN SARAN.....	82
5.1 Kesimpulan	82
5.2 Saran.....	83
DAFTAR PUSTAKA	84
LAMPIRAN.....	88

DAFTAR TABEL

Nomor	Judul Tabel	Halaman
2.1	Diagnosa Pasien Tuberkulosis	8
2.2	Model Diskret dan Variabel Responnya	19
2.3	Tingkat Akurasi Prediksi Berdasarkan MAPE	41
2.4	Beberapa <i>package</i> yang digunakan untuk pengolahan citra	42
4.1	Contoh Rincian Hasil Pembacaan Citra	53
4.2	Contoh Rincian Hasil Pengolahan Citra	56
4.3	Populasi Hasil Bangkitan	59
4.4	Nilai MSE dan <i>Fitness</i> masing-masing individu	60
4.5	Nilai MSE, <i>Fitness</i> , dan <i>Fitness</i> Kumulatif tiap individu	61
4.6	Calon Induk <i>Crossover</i>	63
4.7	Induk <i>Crossover</i> yang terpilih	63
4.8	Induk 1 dan 2 Proses <i>Crossover</i>	65
4.9	Induk 1 dan 2 Proses <i>Crossover</i> Serta Hasil <i>Crossover</i>	65
4.10	Calon Induk Mutasi	66
4.11	Induk Mutasi yang terpilih	67
4.12	Populasi Baru dan Nilai Tujuannya	68
4.13	Populasi Baru Diurutkan Berdasarkan Nilai Tujuannya	70
4.14	Tabel Individu-Individu Terbaik	71
4.15	Individu-Individu Terbaik Pada Tiap Orde dan Iterasi	72
4.16	Hasil Pemodelan Model Parametrik	75
4.17	Hasil Pemodelan Model Nonparametrik	76
4.18	Validasi Hasil Estimasi Menggunakan Data <i>Outsample</i>	78
4.19	Validasi Hasil Estimasi Keseluruhan	80

DAFTAR GAMBAR

Nomor	Judul Gambar	Halaman
2.1	Hasil Radiografi Dada Pasien Tuberkulosis	9
2.2	Contoh Lapang Pandang Dahak Penderita TB	10
2.3	<i>Single Point Crossover</i>	17
2.4	<i>Two Point Crossover</i>	17
3.1	Prosedur Pengolahan Citra	46
3.2	Prosedur Pereduksian Dimensi	48
3.3	Prosedur Pemodelan	51
4.1	Contoh Gambar Hasil Pembacaan Citra	53
4.2	Contoh Citra Hasil Proses <i>Grayscale</i>	54
4.3	Contoh Citra Hasil Proses Ekualisasi	54
4.4	Contoh Citra Hasil Proses Penajaman	55
4.5	Contoh Citra Hasil Proses <i>Multi-Otsu Thresholding</i>	55
4.6	Contoh Citra Hasil Proses Normalisasi	56
4.7	Contoh Citra Hasil Proses Segmentasi	57
4.8	Ilustrasi Roda Rolet (<i>Roulette Wheel</i>)	61
4.9	Hasil Pendeteksian Overdispersi	74