SKRIPSI

PENERAPAN *MULTISCALE GEOGRAPHICALLY WEIGHTED REGRESSION* (MGWR) DALAM PEMODELAN SPASIAL TINGKAT PENGANGGURAN TERDIDIK DI PROVINSI SUMATERA UTARA

SARAH TRI PUJI FLORENSIA SIHOMBING
NPM 22083010020

DOSEN PEMBIMBING
Aviolla Terza Damaliana, S.Si., M.Stat.
Wahyu Syaifullah J.S, S.Kom., M.Kom.

KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA

2026

LEMBAR PENGESAHAN

PENERAPAN *MULTISCALE GEOGRAPHICALLY WEIGHTED REGRESSION* (MGWR) DALAM PEMODELAN SPASIAL TINGKAT PENGANGGURAN TERDIDIK DI PROVINSI SUMATERA UTARA

Oleh:
SARAH TRI PUJI FLORENSIA SIHOMBING
NPM. 22083010020

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur pada Tanggal 27 Oktober 2026:

Menyetujui,

Aviolla Terza Damaliana, S.Si., M.Stat. NIP. 19940802 2022 03 2 015	.....	**(Pembimbing I)**
Wahyu Syaifullah J.S, S.Kom., M.Kom. NIP. 19860825 202121 1 003	.....	**(Pembimbing II)**
Prismahardi Aji Riyantoko, S.Si., M.Si., PhD NIP. 19931011 202406 1 001	.....	**(Ketua Penguji)**
Muhammad Nasrudin, M.Stat. NIP. 19960909 202406 1 002	.....	**(Penguji I)**

Mengetahui,

Dekan Fakultas Ilmu Komputer

Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

PENERAPAN *MULTISCALE GEOGRAPHICALLY WEIGHTED REGRESSION* (MGWR) DALAM PEMODELAN SPASIAL TINGKAT PENGANGGURAN TERDIDIK DI PROVINSI SUMATERA UTARA

Oleh:
SARAH TRI PUJI FLORENSIA SIHOMBING
NPM. 22083010020

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

Plt.Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer

Dr. Faisal Muttaqin, S.Kom., M.T.
NIP. 19851231 202121 1 009

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama Mahasiswa	:	Sarah Tri Puji Florensia Sihombing
NPM	:	22083010020
Program	:	Sarjana (S1)
Program Studi	:	Sains Data
Fakultas	:	Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 20 Oktober 2026
Yang Membuat Pernyataan,



SARAH T.P.F SIHOMBING
NPM. 22083010020

ABSTRAK

Nama Mahasiswa / NPM	:	Sarah Tri Puji Florensia Sihombing / 22083010020
Judul Skripsi	:	Penerapan *Multiscale Geographically Weighted Regression* (MGWR) dalam Pemodelan Spasial Tingkat Pengangguran Terdidik di Provinsi Sumatera Utara
Dosen Pembimbing	:	1. Aviolla Terza Damaliana, S.Si., M.Stat. 2. Wahyu Syaifullah J.S, S.Kom., M.Kom.

Pengangguran terdidik merupakan permasalahan ketenagakerjaan yang berkaitan dengan kemampuan pasar kerja dalam menyerap tenaga kerja berpendidikan. Penelitian ini bertujuan menganalisis pola spasial tingkat pengangguran terdidik, mengidentifikasi hubungan faktor sosial-ekonomi antarwilayah, serta membandingkan kinerja model *Multiscale Geographically Weighted Regression* (MGWR) dengan *Ordinary Least Squares* (OLS). Data yang digunakan merupakan data sekunder Badan Pusat Statistik tahun 2024 pada 33 kabupaten/kota di Provinsi Sumatera Utara. Analisis dilakukan menggunakan *Local Indicators of Spatial Association* (LISA) dan *Multiscale Geographically Weighted Regression* (MGWR) dengan matriks pembobot spasial *K-Nearest Neighbor* (KNN) sebanyak lima tetangga terdekat. Hasil analisis menunjukkan autokorelasi spasial positif yang signifikan berdasarkan nilai Moran's I sebesar 0,3002 dan *p-value* sebesar 0,002. Analisis LISA mengidentifikasi enam wilayah dengan pola spasial signifikan, yaitu tiga wilayah *High-High*, dan dua wilayah *Low-Low* dan satu wilayah *High-Low*. Hasil MGWR menunjukkan bahwa PDRB per Kapita memiliki koefisien negatif yang signifikan pada sembilan kabupaten/kota, sedangkan Upah Minimum Kabupaten/Kota memiliki koefisien positif yang signifikan pada satu kabupaten/kota. Dibandingkan OLS, MGWR menghasilkan RSS lebih rendah serta nilai R^2 dan $Adjusted\ R^2$ lebih tinggi, meskipun nilai AICc-nya lebih besar yang menunjukkan kompleksitas model yang lebih tinggi. Hasil penelitian diimplementasikan dalam dashboard interaktif berbasis R Shiny untuk memvisualisasikan pola spasial dan variasi hubungan antarvariabel sebagai bahan pertimbangan dalam perumusan kebijakan ketenagakerjaan berbasis wilayah.

Kata Kunci: Pengangguran terdidik, LISA, MGWR, R Shiny.

ABSTRACT

Student Name / NPM	:	Sarah Tri Puji Florensia Sihombing / 22083010020
Undergraduate thesis title	:	*Application of Multiscale Geographically Weighted Regression (MGWR) for Spatial Modeling of Educated Unemployment in North Sumatera Province*
Advisors	:	*1. Aviolla Terza Damaliana, S.Si., M.Stat* *2. Wahyu Syaifullah J.S, S.Kom., M.Kom*

Educated unemployment is a labor market issue associated with the labor market's capacity to absorb educated workers. This study aims to analyze the spatial patterns of educated unemployment, identify the relationships between socioeconomic factors and educated unemployment across regions, and compare the performance of Multiscale Geographically Weighted Regression (MGWR) and Ordinary Least Squares (OLS). This study uses secondary data from the Central Statistics Agency (BPS) for 2024, covering 33 regencies/cities in North Sumatera Province. The analysis employs Local Indicator of Spatial Association (LISA) and Multiscale Geographically Weighted Regression (MGWR), using a K-Nearest Neighbors (KNN) spatial weight matrix with five nearest neighbors. The results indicate significant positive spatial autocorrelation, with a Moran's I value of 0.3002 and a p-value of 0.002. The LISA analysis identifies six regions with significant local spatial patterns, comprising three High-High regions, two Low-Low regions, and one High-Low region. The MGWR results show that GRDP per Capita has a statistically significant negative coefficient in nine regencies/cities, whereas the Regency/City has a statistically significant positive coefficient in one regency/city. Compared with OLS, MGWR produces a lower Residual Sum of Squares (RSS) and higher R^2 *and* $\text{Adjusted } R^2$ *values, although its AICc value is higher. The research findings are implemented in an interactive R Shiny dashboard to visualize spatial patterns and variations in relationships between variables, providing information to support the formulation of region-based employment policies.*

Keywords: *Educated unemployment, LISA, MGWR, R Shiny.*

KATA PENGANTAR

Puji syukur ke hadirat Tuhan Yang Maha Esa atas segala berkat dan karunia-Nya kepada penulis sehingga skripsi dengan judul **"PENERAPAN *MULTISCALE GEOGRAPHICALLY WEIGHTED REGRESSION* (MGWR) DALAM PEMODELAN SPASIAL TINGKAT PENGANGGURAN TERDIDIK DI PROVINSI SUMATERA UTARA"** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Ibu Aviolla Terza Damaliana, S.Si., M.Stat selaku Dosen Pembimbing 1 dan Bapak Wahyu Syaifullah J.S, S.Kom., M.Kom selaku Dosen Pembimbing 2 yang telah meluangkan waktunya untuk memberikan bimbingan, nasihat serta motivasi kepada penulis. Selain itu, selama penyusunan skripsi penulis juga banyak menerima bantuan dari berbagai pihak. Untuk itu penulis mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional "Veteran" Jawa Timur.
2. Bapak Dr. Faisal Muttaqin, S.Kom., M.T selaku Pelaksana Tugas Koordinator Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur.
3. Bapak Amri Muhaimin, S.Stat., M.Stat., M.S selaku Dosen Wali yang telah membimbing penulis selama perkuliahan berlangsung.
4. Dosen-dosen Program Studi Sains Data UPN "Veteran" Jawa Timur yang telah memberikan ilmu yang sangat bermanfaat dalam penulisan skripsi ini.
5. Diri sendiri yang telah bertahan dalam pengerjaan skripsi ini dan tetap kuat untuk menjalani rintangan yang ada selama penulisan berlangsung
6. Kedua orang tua dan kedua saudara serta keluarga besar penulis yang selalu memberikan semangat kepada penulis setiap harinya dan memberikan semangat berupa materi maupun non-materi.
7. Nathanael yang telah memotivasi penulis serta membersamai setiap proses yang ada dalam pengerjaan skripsi hingga selesai dari suka maupun duka.

8. Seluruh teman-teman Program Studi Sains Data Angkatan 2022 yang selalu menemani berjuang sejak memasuki masa perkuliahan.
9. Seluruh teman-teman dari Komisi Pemuda dan Remaja GKI Rungkut Asri yang telah memberikan semangat serta membersamai kehidupan penulis selama di Surabaya.
10. Teman-teman SMA penulis (Yenny, Gracia, Sesil dan Noel) yang telah memberikan dukungan selama proses pengerjaan skripsi ini.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, Oktober 2026

Penulis

DAFTAR ISI

HALAMAN SAMPUL ..... **i**
LEMBAR PENGESAHAN ..... **iv**
LEMBAR PERSETUJUAN ..... **iv**
SURAT PERNYATAAN BEBAS PLAGIASI ..... **vi**
ABSTRAK ..... **v**
ABSTRACT ..... **vi**
KATA PENGANTAR ..... **vii**
DAFTAR ISI ..... **ix**
DAFTAR GAMBAR ..... **xiii**
DAFTAR TABEL ..... **xiv**
DAFTAR LAMPIRAN ..... **xv**
DAFTAR NOTASI ..... **xvi**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Rumusan Masalah 7
1.3 Batasan Masalah 8
1.4 Tujuan Penelitian 8
1.5 Manfaat Penelitian 9
BAB II TINJAUAN PUSTAKA **10**
2.1 Penelitian Terdahulu 10
2.2 Kerangka Teori 16
2.2.1 Tingkat Pengangguran Terdidik 16
2.2.2 Faktor Sosio-Ekonomi 19
2.2.3 Regresi Linear 28
2.2.4 Uji Simultan 29
2.2.5 Uji Parsial 30
2.2.6 Uji Asumsi Klasik 31
2.2.7 Uji Heterogenitas Spasial 33
2.2.8 Uji Autokorelasi Spasial Global 34
2.2.9 Analisis Autokorelasi Spasial Lokal (LISA) 36
2.2.10 Fungsi Pembobot 40

2.2.11 Bandwidth Optimal 42
2.2.12 *Geographically Weighted Regression (GWR)* 43
2.2.13 *Multiscale Geographically Weighted Regression* (MGWR) 46
2.2.14 Pengujian Signifikansi Model MGWR 50
2.2.15 Evaluasi Model MGWR 52
2.2.16 *User Interface* 54

BAB III METODOLOGI PENELITIAN 55
3.1. Variabel Penelitian dan Sumber Data 55
3.2 Langkah Analisis 56
3.3 Desain Sistem 61

BAB IV HASIL DAN PEMBAHASAN 63
4.1 Pengumpulan dan Penyusunan Data 63
4.2 Preprocessing Data 64
4.2.1 Pemeriksaan *Missing Values* dan Baris Duplikat 64
4.2.2 Integrasi Data Tabular dengan Data Spasial 65
4.3 Analisis Deskriptif 71
4.3.1 Statistika Deskriptif 72
4.3.2 Eksplorasi Persebaran Tingkat Pengangguran Terdidik 75
4.3.3 Eksplorasi Hubungan Antarvariabel 77
4.4 Pemodelan Regresi Linear Berganda 81
4.4.1 Estimasi Parameter 81
4.4.2 Uji Asumsi Klasik 84
4.4.2.1 Uji Normalitas 84
4.4.2.2 Uji Multikolinearitas 86
4.4.3 Uji Signifikansi Model 88
4.4.3.1 Uji Simultan 88
4.4.2.2 Uji Parsial 90
4.5 Uji Aspek Spasial 92
4.5.1 Pembentukan Matriks Pembobot Spasial 92
4.5.2 Uji Heterogenitas Spasial 95
4.5.3 Uji Autokorelasi Spasial Global 97

4.5.4 Analisis Autokorelasi Spasial Lokal (LISA) 100
4.6 Pemodelan *Multiscale Geographically Weighted Regression* (MGWR) 104
4.6.1 Standarisasi Data 105
4.6.2 Perhitungan Matriks Jarak 107
4.6.3 Pemilihan Bandwidth Optimal 114
4.6.4 Estimasi Parameter MGWR 117
4.6.5 Perhitungan Manual Estimasi Parameter MGWR 120
4.6.6 Uji Signifikansi Parameter 133
4.7 Evaluasi Model 137
4.8 Rekomendasi Kebijakan 140
4.9 Penerapan Metode MGWR melalui *User Interface* 143
*4.*9.1 Halaman *Home* 143
4.9.2 Halaman Unggah Data 144
4.9.3 Halaman Konfigurasi 145
4.9.4 Halaman Hasil Pemodelan 147

BAB V PENUTUP 152
5.1 Kesimpulan 152
5.2 Saran Pengembangan 154

DAFTAR PUSTAKA 157
LAMPIRAN 163

DAFTAR GAMBAR

Gambar 3. 1 Diagram Alir Penelitian 56
Gambar 3. 2 Desain Sistem GUI..... 61
Gambar 4. 1 Peta Persebaran Tingkat Pengangguran Terdidik Sumatera Utara .. 76
Gambar 4. 2 Hubungan Antarvariabel 79
Gambar 4. 3 Peta Klaster LISA..... 103
Gambar 4. 4 Peta Kelompok Variabel Signifikan..... 141
Gambar 4. 5 Tampilan Halaman *Home* 144
Gambar 4. 6 Tampilan Halaman Sebelum Data Diunggah..... 145
Gambar 4. 7 Tampilan Halaman Setelah Data Diunggah..... 145
Gambar 4. 8 Tampilan Halaman Konfigurasi..... 146
Gambar 4. 9 Tampilan *Tab* Ringkasan OLS..... 147
Gambar 4. 10 Tampilan *Tab* Uji Asumsi dan Spasial..... 148
Gambar 4. 11 Tampilan *tab* Hasil MGWR..... 149
Gambar 4. 12 Tampilan *tab* Interpretasi dan Unduh..... 150

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu 10
Tabel 2. 2 Indikator Warna pada LISA 39
Tabel 3. 1 Variabel Penelitian 55
Tabel 3. 2 Struktur Data 56
Tabel 4. 1 Data Penelitian 63
Tabel 4. 2 Hasil Integrasi Data Tabular dan Data Spasial 71
Tabel 4. 3 Hasil Statistika Deskriptif 73
Tabel 4. 4 Hasil Estimasi Parameter Regresi Linear Berganda 82
Tabel 4. 5 Hasil Uji Normalitas *Shapiro-Wilk* 85
Tabel 4. 6 Hasil Uji Multikolinearitas 87
Tabel 4. 7 Hasil Uji Simultan 89
Tabel 4. 8 Hasil Uji Parsial 91
Tabel 4. 9 Matriks Pembobot Spasial 94
Tabel 4. 10 Hasil Uji Heterogenitas Spasial 96
Tabel 4. 11 Hasil Uji Autokorelasi Spasial Global 99
Tabel 4. 12 Hasil Perhitungan *Local Moran's I* 101
Tabel 4. 13 Data Hasil Standarisasi 106
Tabel 4. 14 Matriks Jarak *Haversine* (km) 113
Tabel 4. 15 Hasil *Bandwidth* Optimal 116
Tabel 4. 16 Hasil Estimasi Parameter MGWR 119
Tabel 4. 17 Data Terstandarisasi 121
Tabel 4. 18 Matriks Jarak 122
Tabel 4. 19 Hasil Pembaruan Parameter Iterasi Pertama 130
Tabel 4. 20 *Bandwidth* Optimal 132
Tabel 4. 21 Jumlah Wilayah Signifikan Setiap Variabel Model MGWR 135
Tabel 4. 22 Klasifikasi Wilayah Berdasarkan Variabel Signifikan 135
Tabel 4. 23 Perbandingan Model OLS dan MGWR 138

DAFTAR LAMPIRAN

Lampiran 1. Hasil Integrasi Data Tabular dan Data Spasial 163
Lampiran 2. Kode Program Analisis 164
Lampiran 3. Kode Program GUI 165

DAFTAR NOTASI

Y	:	Tingkat Pengangguran Terdidik
X_1	:	Indeks Pembangunan Manusia (IPM)
X_2	:	Produk Domestik Regional Bruto (PDRB) per Kapita
X_3	:	Kepadatan Penduduk
X_4	:	Upah Minimum Kabupaten/Kota (UMK)
X_5	:	Tingkat Kemiskinan
Y_i	:	Nilai variabel dependen pada wilayah ke-*i*
X_{ik}	:	Nilai variabel independent ke-*k* pada wilayah ke-*i*
β_0	:	Konstanta atau intersep regresi global
β_k	:	Koefisien regresi global untuk variabel ke-*k*
$\beta_k(u_i, v_i)$	:	Koefisien regresi lokal untuk variabel ke-*k* pada lokasi ke-*i*
$\hat{\beta}_{k,i}$	:	Estimasi koefisien lokal parameter ke-*k* pada wilayah ke-*i*
$\hat{Y}_i$	:	Nilai prediksi variabel dependen pada wilayah ke-*i*
ε_i	:	Residual atau galat pada wilayah ke-*i*
$\mathbf{X}$	:	Matriks variabel independen
y	:	Vektor variabel dependen
β	:	Vektor koefisien regresi
u_i, v_i	:	Koordinat geografis wilayah ke-*i*
x_{ij}	:	Nilai observasi ke-*i* pada variabel ke-*j* sebelum standardisasi
$\bar{x}_j$	:	Rata-rata variabel ke-*j*
s_j	:	Simpangan baku variabel ke-*j*
z_{ij}	:	Nilai standardisasi observasi ke-*i* pada variabel ke-*j*
d_{ij}	:	Jarak geografis antara wilayah ke-*i* dan ke-*j*
φ_i	:	Latitude wilayah ke-*i* dalam radian
λ_i	:	Longitude wilayah ke-*i* dalam radian
R	:	Jari-jari bumi dalam perhitungan jarak Haversine
w_{ij}	:	Bobot spasial wilayah ke-*j* terhadap wilayah ke-*i*
$\mathbf{W}$	:	Matriks pembobot spasial

b_i	:	Batas jarak *bandwidth* pada lokasi ke-*i* untuk kernel adaptif
bw_k	:	*Bandwidth* untuk parameter ke-*k* pada model MGWR
H_0	:	Hipotesis nol
H_1	:	Hipotesis alternatif
α	:	Taraf signifikansi pengujian
t_{ik}	:	Statistik uji *t* untuk parameter ke-*k* pada lokasi ke-*i*
t_{kritis}	:	Nilai kritis distribusi *t*
I	:	Indeks Moran global
I_i	:	Indeks Moran lokal pada wilayah ke-*i*
$E(I)$	:	Nilai harapan indeks Moran global
σ^2	:	Varians residual
W_{SW}	:	Statistik uji *Shapiro-Wilk*