

**ANALISIS SPASIAL TERHADAP MODEL
KLASIFIKASI *RANDOM FOREST* UNTUK
PEMETAAN KETERSEDIAAN FASILITAS
OLAHRAGA DI DKI JAKARTA**

SKRIPSI SARJANA SISTEM INFORMASI

Oleh:

Hansen Candra

217006516063



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI
DAN INFORMATIKA
UNIVERSITAS NASIONAL**

2024

**ANALISIS SPASIAL TERHADAP MODEL
KLASIFIKASI *RANDOM FOREST* UNTUK
PEMETAAN KETERSEDIAAN FASILITAS
OLAHRAGA DI DKI JAKARTA**

SKRIPSI SARJANA SISTEM INFORMASI

Karya ilmiah sebagai salah satu syarat untuk memperoleh gelar Sarjana Sistem
Informasi dari Fakultas Teknologi Komunikasi dan Informatika

Oleh:

Hansen Candra

217006516063



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNOLOGI KOMUNIKASI
DAN INFORMATIKA
UNIVERSITAS NASIONAL**

2024

HALAMAN PENGESAHAN

TUGAS AKHIR

ANALISIS SPASIAL TERHADAP MODEL KLASIFIKASI *RANDOM FOREST* UNTUK PEMETAAN KETERSEDIAAN FASILITAS OLAHRAGA DI DKI JAKARTA



Dosen Pembimbing

(Dr. Andrianingsih, S.Kom., MMSI.)

PERNYATAAN KEASLIAN TUGAS AKHIR

Saya menyatakan dengan sesungguhnya bahwa Tugas Akhir dengan judul :

ANALISIS SPASIAL TERHADAP MODEL KLASIFIKASI RANDOM FOREST UNTUK PEMETAAN KETERSEDIAAN FASILITAS OLAHRAGA DI DKI

JAKARTA

Yang dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional, sebagaimana yang saya ketahui adalah bukan merupakan tiruan atau publikasi dari Tugas Akhir yang pernah diajukan atau dipakai untuk mendapatkan gelar di lingkungan Universitas Nasional maupun perguruan tinggi atau instansi lainnya, kecuali pada bagian – bagian tertentu yang menjadi sumber informasi atau acuan yang dicantumkan sebagaimana mestinya.

Jakarta, 28 Februari 2025



Hansen Candra

217006516063

LEMBAR PERSETUJUAN REVIEW AKHIR

Tugas Akhir dengan judul :

ANALISIS SPASIAL TERHADAP MODEL KLASIFIKASI RANDOM FOREST UNTUK PEMETAAN KETERSEDIAAN FASILITAS OLAHRAGA DI DKI JAKARTA

Dibuat untuk melengkapi salah satu persyaratan menjadi Sarjana Komputer pada Program Studi Sistem Informasi, Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional. Tugas Akhir ini diujikan pada Sidang Review Akhir Semester Ganjil 2024-2025 pada tanggal 28 Februari Tahun 2025

Dosen Pembimbing 1

Dr. Andrianingsih, S.Kom., MMSI.)

NIDN: 0303097902

Ketua Program Studi

Dr. Andrianingsih, S.Kom., MMSI.)

0303097902

LEMBAR PERSETUJUAN JUDUL YANG TIDAK ATAU YANG DIREVISI

Nama : Hansen Candra
 NPM : 217006516063
 Fakultas/Akademi : Fakultas Teknologi Komunikasi dan Informatika
 Program Studi : Sistem Informasi
 Tanggal Sidang : 26 Februari 2025

JUDUL DALAM BAHASA INDONESIA :

ANALISIS SPASIAL TERHADAP MODEL KLASIFIKASI RANDOM FOREST
 UNTUK PEMETAAN KETERSEDIAAN FASILITAS OLAHRAHA DI DKI
 JAKARTA

JUDUL DALAM BAHASA INGGRIS :

ANALYSIS OF RANDOM FOREST CLASSIFICATION MODEL FOR
 AVAILABILITY MAPPING OF SPORTS FACILITIES IN JAKARTA

TANDA TANGAN DAN TANGGAL

Pembimbing 1	Ka. Prodi	Mahasiswa
TGL : 28. Februari 2025.	TGL : 28. Februari 2025.	TGL : 28. Februari 2025.
		

KATA PENGANTAR

Puji syukur kami panjatkan kepada Tuhan Yang Maha Esa, atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan laporan ini dengan judul "**Analisis Spasial Terhadap Model Klasifikasi Random Forest Untuk Pemetaan Ketersediaan Fasilitas Olahraga Di DKI Jakarta**". Penulisan laporan ini disusun untuk memenuhi salah satu persyaratan dalam kelulusan, serta sebagai sarana pembelajaran dan pengembangan kompetensi.

Dalam proses penyusunan laporan ini, penulis telah mendapatkan banyak dukungan, bimbingan, dan masukan dari berbagai pihak, terutama kepada dosen pembimbing Tugas Akhir yaitu Ibu Dr. Andrianingsih, S.Kom., MMSI. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Dr. Agung Triayudi, S.Kom., M.Kom. selaku Dekan Fakultas Teknologi Komunikasi dan Informatika Universitas Nasional.
2. Ibu Dr. Andrianingsih, S.Kom., MMSI. selaku Ketua Program Studi Sistem Informasi sekaligus dosen pembimbing skripsi, yang telah memberikan arahan, dukungan, serta bimbingan dengan penuh kesabaran dan dedikasi hingga terselesaikannya skripsi ini.
3. Kedua orang tua tercinta, yang telah memberikan doa, motivasi, dukungan, dan pengorbanan tanpa batas, penulis sangat berterima kasih atas segala yang telah diberikan selama ini.
4. Penulis juga memberikan apresiasi kepada diri sendiri atas usaha dan menjaga semangat, konsistensi dalam beribadah, serta kepercayaan diri, penulis mampu menghadapi tantangan hingga mencapai tahap ini.
5. Semua pihak yang tidak dapat disebutkan satu per-satu, baik dosen dan teman seangkatan yang telah memberikan dukungan secara langsung maupun tidak langsung selama proses penyusunan skripsi ini.

Penulis menyadari bahwa laporan ini masih memiliki keterbatasan dan jauh dari kesempurnaan. Oleh karena itu, penulis membuka diri untuk menerima saran dan kritik yang membangun demi perbaikan di masa mendatang. Semoga laporan

ini dapat bermanfaat bagi pembaca dan memberikan kontribusi positif. Akhir kata, penulis mengucapkan terima kasih dan berharap laporan ini dapat diterima dengan baik.

Jakarta, 28 Februari 2024




Hansen Candra

ABSTRAK

Dalam rangka menyambut Indonesia Emas 2045, pemerintah berkomitmen untuk memajukan sektor olahraga di tingkat nasional. Dukungan diberikan secara menyeluruh untuk mendorong perkembangan olahraga di Indonesia. Penelitian ini bertujuan untuk menganalisis distribusi fasilitas olahraga di Provinsi DKI Jakarta menggunakan pemodelan spasial dan algoritma Machine Learning Random Forest. Tujuan utama dari penelitian ini adalah untuk mengelompokkan wilayah berdasarkan tingkat ketersediaan fasilitas olahraga menjadi kategori rendah, cukup, dan tinggi, serta mengevaluasi akurasi algoritma Random Forest dalam melakukan klasifikasi tersebut. Metodologi yang digunakan dalam penelitian ini adalah CRISP-DM (Cross-Industry Standard Process for Data Mining) yang terdiri dari Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, dan Deployment. Data yang digunakan meliputi data spasial wilayah kecamatan di DKI Jakarta dan data atribut fasilitas olahraga. Algoritma Random Forest dipilih karena kemampuannya dalam mengklasifikasikan data yang kompleks dan memberikan informasi penting mengenai feature importance. Hasil penelitian menunjukkan bahwa distribusi fasilitas olahraga di DKI Jakarta tidak merata, dengan kategori rendah lebih banyak ditemukan di wilayah Jakarta Pusat dan Jakarta Utara, sedangkan kategori tinggi tersebar di wilayah lainnya. Algoritma Random Forest menunjukkan akurasi sebesar 89% dalam mengklasifikasikan fasilitas olahraga, dengan precision dan recall yang tinggi pada kategori tinggi, namun precision yang lebih rendah pada kategori rendah disebabkan oleh ketidakseimbangan jumlah data.

Kata Kunci: *Klasifikasi Fasilitas Olahraga, Analisis Spasial, Machine Learning, Sistem Informasi Geografi, Random Forest, CRISP-DM.*

ABSTRACT

In order to welcome the Golden Indonesia 2045, the government is committed to advancing the sports sector at the national level. Support is provided as a whole to encourage the development of sports in Indonesia. This research aims to analyze the distribution of sports facilities in DKI Jakarta Province using spatial modeling and Machine Learning Random Forest algorithm. The main objective of this research is to classify areas based on the level of availability of sports facilities into low, sufficient, and high categories, and evaluate the accuracy of the Random Forest algorithm in performing the classification. The methodology used in this research is CRISP-DM (Cross-Industry Standard Process for Data Mining) which consists of Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. The data used includes spatial data of sub-district areas in DKI Jakarta and sports facility attribute data. The Random Forest algorithm was chosen because of its ability to classify complex data and provide important information about feature importance. The results showed that the distribution of sports facilities in DKI Jakarta is uneven, with low categories found more in the Central Jakarta and North Jakarta areas, while high categories are scattered in other areas. The Random Forest algorithm showed 89% accuracy in classifying sports facilities, with high precision and recall in the high category, but lower precision in the low category due to the imbalance in the amount of data.

Keywords: Sports Facility Classification, Spatial Analysis, Machine Learning, Geographic Information System, Random Forest, CRISP-DM.

DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR	vii
ABSTRAK	ix
ABSTRACT.....	x
DAFTAR ISI.....	xi
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Tujuan	6
1.4 Manfaat Penelitian	7
1.5 Batasan Masalah.....	7
BAB II TINJAUAN PUSTAKA.....	9
2.1 Landasan Teori.....	9
2.2 <i>Research Positioning</i>	16
2.3 Technology Information System.....	18
2.3.1 <i>Machine Learning</i>	18
2.3.2 Analisis Spasial.....	19
2.3.3 CRISP-DM	20
2.3.4 Klasifikasi.....	21
2.3.5 Algoritma Random Forest	21
2.3.6 Confusion Matrix.....	22
2.3.7 Python.....	24
2.3.8 Jupyter	24
2.3.9 QGIS.....	24
BAB III METODOLOGI PENELITIAN.....	26
3.1 Lokasi Penelitian.....	26
3.2 Roadmap Penelitian	27

3.3	Penentuan Objek Penelitian	28
3.4	Sumber Data.....	29
3.5	Tahapan Penelitian.....	29
3.6	Analisis Spasial	30
3.7	Metode CRISP-DM.....	31
3.7.1	Business Understanding (Pemahaman Bisnis).....	31
3.7.2	Data Understanding.....	31
3.7.3	Data Preparation (Persiapan Data)	32
3.7.4	Modeling (Pemodelan).....	32
3.7.5	Evaluation (Evaluasi).....	33
3.7.6	Deployment (Penerapan).....	33
BAB IV HASIL DAN PEMBAHASAN		34
4.1	Implementasi Tahap Penelitian	34
4.2	Business Understanding	34
4.3	Data Understanding	37
4.4	Data Preparation	39
4.5	Modeling.....	49
4.6	Evaluation.....	51
4.7	Pemodelan Spasial.....	52
4.7.1	Peta Kecamatan Provinsi DKI Jakarta	52
4.7.2	Join Attributes by Field Value	53
4.7.3	Peta Jumlah Penduduk Kecamatan Provinsi DKI Jakarta.....	56
4.7.4	Peta Klasifikasi Kategori Fasilitas Olahraga DKI Jakarta	57
4.8	Perhitungan Manual.....	59
4.9	Deployment	61
BAB V KESIMPULAN DAN SARAN.....		67
5.1	Kesimpulan.....	67
5.2	Saran	68
DAFTAR PUSTAKA		69
LAMPIRAN.....		72

DAFTAR GAMBAR

Gambar 1. 1 Provisni DKI Jakarta	2
Gambar 2. 1 Reasearch Positioning	16
Gambar 2. 2 Penelitian Terkait	17
Gambar 2. 3 Tahapan Machine Learning.....	19
Gambar 2. 4 Pemetaan Menggunakan QGIS	20
Gambar 2. 5 Tahapan CRISP-DM	21
Gambar 2. 6 Tahapan Random Forest.....	22
Gambar 2.7 QGIS.....	25
Gambar 3. 1 Peta Provinsi DKI Jakarta	26
Gambar 3. 2 RoadMap Penelitian	27
Gambar 3. 3 Tahapan Penelitian	30
Gambar 4. 1 Kementiran Pemuda Dan Olahraga RI.....	35
Gambar 4.2 Dataset Fasilitas Olahraga.....	35
Gambar 4. 3 Gambar Data Spasial administrasi Kecamatan DKI Jakarta.....	36
Gambar 4. 4 Gambar Import Library	37
Gambar 4. 5 Dataset Head	38
Gambar 4. 6 Statistik Deskriptif	38
Gambar 4. 7 Informasi Dataset	38
Gambar 4. 8 Missing Value	39
Gambar 4. 9 Data Baris Duplikat.....	39
Gambar 4. 10 Baris Duplikat yang sudah di hapus.....	40
Gambar 4. 11 Grafik Jumlah Penduduk Perkota.....	40
Gambar 4. 12 Jumlah Penduduk Perkecamatan	41
Gambar 4. 13 Distribusi Fasilitas Olahraga Jenis per Kota/Kabupaten.....	42
Gambar 4. 14 Distribusi Fasilitas Olahraga Jenis per Kecamatan	43
Gambar 4. 15 Grafik Pairplot.....	44
Gambar 4. 16 Feature Engineer	45
Gambar 4. 17 hasil klasifikasi kecamatan.....	45
Gambar 4. 18 Intersection Atribut kategori_rasio_fasilitas	46

Gambar 4. 19 Hasil kategori_rasio_fasilitas	46
Gambar 4. 20 Code Encoding	47
Gambar 4. 21 Hasil Encoding	47
Gambar 4. 22 Hasil Encoding	48
Gambar 4. 23 Hasil Scaling	48
Gambar 4. 24 Distribusi Kategori Fasilitas Olahraga	49
Gambar 4. 25 Distribusi Kategori Fasilitas Olahraga (Bar Chart).....	49
Gambar 4. 26 Distribusi Kategori Fasilitas Olahraga (Bar Chart).....	50
Gambar 4. 27 Confusion Matrix Random Forest.....	51
Gambar 4. 28 Peta per Kecamatan Provinsi DKI Jakarta	52
Gambar 4. 29 Join Attributes by Field Value	54
Gambar 4. 30 Table Setelah Join Attributes Value.....	55
Gambar 4. 31 Jumlah penduduk per kecamatan	56
Gambar 4. 32 klasifikasi fasilitas olahraga per kecamatan	58
Gambar 4. 33 QGIS2WEB Plugin	61
Gambar 4. 34 Layers and Group	62
Gambar 4. 35 Appearance.....	63
Gambar 4. 36 Export.....	64
Gambar 4. 37 Website Pemetaan Fasilitas Olahraga	65

DAFTAR TABEL

Tabel 1. 1 Landasan Teori.....	9
Tabel 2. 1 Confusion Matrix	23
Tabel 3. 1 Parameter Sumber Data	29
Tabel 4. 1 Dataset Sampel.....	59
Tabel 4. 2 Konversi Data	59

