

***MACHINE LEARNING***  
**DATA OSEANOGRAFI INDONESIA**  
**UNTUK PREDIKSI POTENSI SUMBER DAYA LAUT**

**Tesis**  
**untuk memenuhi sebagian persyaratan**  
**mencapai derajat Sarjana S-2 Program Studi**  
**Magister Sistem Informasi**



**Denny Arbahri**  
**3000321413024**

**SEKOLAH PASCASARJANA**  
**UNIVERSITAS DIPONEGORO**  
**SEMARANG**  
**2024**

**HALAMAN PENGESAHAN**

**TESIS**

***MACHINE LEARNING DATA OSEANOGRAFI INDONESIA  
UNTUK PREDIKSI POTENSI SUMBER DAYA LAUT***

**Oleh:**

**Denny Arbahri  
30000321413024**

Telah diujikan dan dinyatakan lulus ujian tesis pada tanggal 18 Maret 2024 oleh tim penguji Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro.

Semarang, 18 Maret 2024  
Mengetahui,

Pembimbing I

Dr. Oky Dwi Nurhayati, ST., MT  
NIP. 197910022009122001

Penguji I

Prof. Dr. Ir. R. Rizal Isnanto, S.T., M.M.,  
M.T., IPU, ASEAN Eng.  
NIP. 197007272000121001

Pembimbing II

Dr. Ir. Imam Mudita, M.Eng.Sc.  
NIP. 196306021990121001

Penguji II

Dr. Kasiyati, S.Si., M.Si.  
NIP. 197705262005012001

Mengetahui,  
Dekan Sekolah Pasca Sarjana  
Universitas Diponegoro



Dr. R.B. Sularto, S.H., M.Hum  
NIP. 196701011991031005

Ketua Program Studi  
Magister Sistem Informasi UNDIP

Drs. Bayu Surarso, M.Sc., Ph.D.  
NIP. 196311051988031001

**PERNYATAAN PERSETUJUAN  
PUBLIKASI TESIS UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Diponegoro, saya yang bertanda tangan di bawah ini:

Nama : Denny Arbahri  
NIM : 30000321413024  
Program Studi : Magister Sistem Informasi  
Program : Sekolah Pascasarjana  
Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Diponegoro Hak Bebas Royalti Noneksklusif atas karya ilmiah saya yang berjudul:

***MACHINE LEARNING DATA OSEANOGRAFI INDONESIA  
UNTUK PREDIKSI POTENSI SUMBER DAYA LAUT***

beserta perangkat yang ada. Dengan Hak bebas Royalti Noneksklusif ini Program Studi Magister Sistem Informasi Sekolah Pascasarjana Universitas Diponegoro berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tesis saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta.

Dibuat di : Semarang  
Pada tanggal : 21 Februari 2024  
Yang menyatakan



Denny Arbahri  
NIM. 30000321413024

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam tesis ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar akademik di suatu perguruan tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Semarang, 21 Februari 2024



Denny Arbahri

NIM. 30000321413024



SEKOLAH PASCASARJANA

## KATA PENGANTAR

Bismillahirrahmanirrahiim

Puji syukur penulis panjatkan kehadirat Allah Subhanawata“ala atas karunia rahmat dan hidayah Nya, sehingga penulis dapat menyelesaikan tesis ini dengan lancar. Shalawat dan salam penulis panjatkan keharibaan Rasulullah Shalallahu Alaihi Wassalam serta keluarga dan para sahabatnya.

Selanjutnya penulis ingin menyampaikan rasa terima kasih yang tak terhingga kepada semua pihak yang telah membantu kelancaran tesis ini, baik berupa dorongan moril maupun materil, karena dengan dukungan beliau-beliau penulis dapat menyelesaikan tesis ini. Penulis ingin menyampaikan penghargaan yang sebesar-besarnya kepada yang terhormat:

1. Ibunda tercinta, dengan ungkapan terima kasih dan penghargaan yang sangat spesial penulis haturkan dengan penuh hormat dan rendah hati,
2. Ibu Dr. Oky Dwi Nurhayati, S.T., M.T. selaku dosen pembimbing I tesis,
3. Bapak Dr. Ir. Imam Mudita, M.Eng.Sc. selaku pembimbing II tesis,
4. Bapak Dr. R.B. Sularto, S.H., M.Hum. selaku Dekan Pascasarjana Universitas Diponegoro,
5. Bapak Drs. Bayu Surarso, M.Sc., P.hD selaku Ketua Prodi Magister Sistem Informasi Universitas Diponegoro,
6. Bapak/Ibu dosen Magister Sistem Informasi Universitas Diponegoro yang telah banyak berjasa memberikan ilmunya kepada penulis,
7. Bapak/Ibu pimpinan Badan Riset Dan Inovasi Nasional yang telah memberikan kebijakan pendanaan dalam mendukung sekolah penulis,
8. Arifiyana selaku analist oseanografer yang telah berjasa membantu penulis dalam mengkonversi data mentah oseanografi menjadi data yang dapat digunakan untuk penelitian ini. Dan Bapak/Ibu yang belum bisa disebutkan di sini karena keterbatasan untuk mengingatnya satu-persatu.

Semarang, 17 November 2023

DENNY ARBAHRI



## DAFTAR ISI

	Halaman
Halaman Judul .....	i
Halaman Pengesahan .....	ii
Halaman Pernyataan Persetujuan Publikasi .....	iii
Halaman Pernyataan.....	iv
Kata Pengantar .....	v
Daftar Isi .....	vi
Daftar Gambar .....	vii
Daftar Tabel .....	viii
Daftar Arti Lambang dan Singkatan .....	ix
Daftar Lampiran.....	x
Abstrak.....	xi
Abstract .....	xii
<b>BAB I PENDAHULUAN</b>	
1.1. Latar Belakang .....	1
1.2. Tujuan Penelitian .....	6
1.3. Manfaat Penelitian .....	6
<b>BAB II TINJAUAN PUSTAKA DAN DASAR TEORI</b>	
2.1. Tinjauan Pustaka .....	7
2.2. Dasar Teori .....	9
2.2.1. <i>Machine Learning</i> (ML) .....	9
2.2.2. <i>Orange3</i> untuk <i>Unsupervised Learning</i> .....	20
2.2.3. <i>Oceanografi Conductivity, Temperature, Depth</i> (CTD).....	21
2.2.4. Sumber Daya Laut.....	23
<b>BAB III METODE PENELITIAN</b>	
3.1. Bahan dan Alat Penelitian .....	25
3.2. Prosedur Penelitian .....	27
3.3. Metode Penelitian .....	31
<b>BAB IV HASIL PENELITIAN DAN PEMBAHASAN</b>	
4.1. Hasil Penelitian.....	35
4.2. Pembahasan .....	39
<b>BAB V KESIMPULAN DAN SARAN</b>	
5.1. Kesimpulan.....	45
5.2. Saran.....	46
Daftar Pustaka .....	47
LAMPIRAN.....	51

## DAFTAR GAMBAR

	Halaman
2.1. Jumlah dan tipe variabel serta bentuk kurva regresi .....	12
3.1. Diagram alur kerja penelitian yang menggambarkan tahapan metodologi penelitian .....	26
3.2. Diagram alur proses pengumpulan data oseanografi CTD .....	28
3.3. Diagram prosedur penggunaan <i>orange3</i> dalam penelitian ini .....	30
3.4. Distribusi data CTD yang tersebar dengan baik .....	31
3.5. Ranking data CTD .....	32
3.6. Korelasi antara variabel prediktor dengan variabel target.....	32
4.1. Hasil pelatihan dan uji data CTD.....	34
4.2. Hasil prediksi model DT – LR – RF.....	36
4.3. Alur proses ML dengan dukungan <i>orange3</i> melakukan prediksi. ....	38
4.4. Simpul akar hasil pemodelan DT.....	39

SEMARANG  
SEKOLAH PASCASARJANA

## DAFTAR TABEL

	Halaman
2.1. Ringkasan tinjauan literatur penelitian terdahulu.....	6
3.1. Data CTD sebagai data pokok dalam penelitian ini.....	25
4.1. PREDIKTOR: X1 – X2 – X3 .....	40
4.2. $RSS_{0.5}$ hasil perhitungan untuk prediktor X1, X2, X3.....	40
4.3. Deskripsi Koefisien Korelasi Pearson.....	42
4.4. Korelasi fitur-fitur antara penelitian ini dengan penelitian terdahulu .....	42



SEKOLAH PASCASARJANA



## DAFTAR ARTI LAMBANG DAN SINGKATAN

### DAFTAR SINGKATAN

Singkatan	Kepanjangan Singkatan
AI	<i>Artificial Intelligent</i>
BPPT	Badan Pengkajian dan Penerapan Teknologi
CTD	<i>Conductivity, Temperature, Depth</i>
Cond	<i>Conductivity</i>
cm	<i>centimeter</i>
Dept	<i>Depth</i>
DT	<i>Decision Tree</i>
id	<i>idntity</i>
KNN	<i>K-Nearest Neighbor</i>
LR	<i>Linear Regression</i>
ML	<i>Machine Learning</i>
MSE	<i>Mean Square Error</i>
MAE	<i>Mean Absolute Error</i>
m	<i>meter</i>
NN	<i>Neural Network</i>
RSS	<i>Residual Some of Squares</i>
RMSE	<i>Root Mean Square Error</i>
RF	<i>Random Forest</i>
Salt	<i>Salinity</i>
SVM	<i>Support Vector Machine</i>
Temp	<i>Temperature</i>
USL	<i>Unsupervised Learning</i>
ppt	<i>part per thousand</i>
psu	<i>practical salinity units</i>

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1. SOP akuisisi CTD.....	51
Lampiran 2. Prosedur konversi data CTD.....	58
Lampiran 3. Alur QC CTD .....	59
Lampiran 4. Prediktor X2 dan X3.....	60
Lampiran 5. Daftar Survei Balai Teksurla.....	62



SEKOLAH PASCASARJANA

# **MACHINE LEARNING DATA OSEANOGRAFI INDONESIA UNTUK PREDIKSI POTENSI SUMBER DAYA LAUT**

## **ABSTRAK**

Data dan informasi kelautan sangat penting bagi kelangsungan hidup manusia, oleh karena itu data dan informasi tersebut menarik bagi investor karena adanya potensi nilai ekonomi. Data dan informasi kelautan yang ada belum dikelola dan belum difungsikan untuk memprediksi sumber daya laut dengan menggunakan teknik *Machine Learning* (ML), sehingga penguasaan data dan informasi kelautan menjadi permasalahan dalam konservasi sumber daya laut. Konservasi sumber daya laut secara berkelanjutan yang benar, dilaksanakan dengan cara menguasai data dan informasi kelautan. Saat ini penelitian ilmiah kelautan telah memasuki era baru *Artificial Intelligence* (AI), seperti penggunaan ML dalam penelitian ilmiah kelautan cukup menjanjikan, khususnya untuk memantau keanekaragaman hayati laut, memodelkan data CTD, dan memprediksi sumber daya laut. ML untuk membangun model prediktif saat ini telah banyak digunakan oleh para peneliti, namun spesifikasi penelitian dengan teknik ML untuk memprediksi potensi sumber daya laut merupakan bidang penelitian yang relatif baru. Tujuan penelitian ini, membuat model data oseanografi Indonesia untuk memprediksi potensi sumber daya laut dengan teknik ML menggunakan algoritma *Decision Tree* (DT), *Linear Regression* (LR), dan *Random Forest* (RF), ketiganya untuk dibandingkan. Metodologinya, data mentah yang telah dikumpulkan harus dikonversi supaya menjadi data oseanografi *Conductivity, Temperature, Depth* (CTD) yang lengkap dengan informasi. Data itu diklasifikasikan kedalam data latih dan data uji dengan *software orange3*. Data latih digunakan untuk pemodelan dengan algoritma, sehingga menghasilkan model data. Model tersebut diuji dengan data uji hingga menghasilkan model dengan skor terbaik dan selanjutnya difungsikan untuk prediksi data target. Hasil prediksi ketiga model menunjukkan nilai salinitas 34 *parts per thousand* (ppt), dan nilai aktualnya pun sebesar 34 ppt. Artinya, pada area perairan laut dengan nilai salinitas 34 ppt, maka salinitas akan mempengaruhi kelarutan oksigen (O<sub>2</sub>) serta berperan besar dalam keberlangsungan dan pertumbuhan tingkat kesuburan sumber daya hayati yang didukung oleh kondisi suhu permukaan laut 29°C. Sehingga nilai salinitas 34 ppt yang diprediksi dengan teknik ML, diasumsikan memiliki korelasi yang kuat dengan adanya potensi sumber daya laut. Penelitian ini berpotensi memberikan kontribusi data dan informasi sebagai referensi penelitian lanjutan yang inovatif bagi para peneliti dan bahan keputusan investasi bagi para investor serta pedoman bagi para nelayan untuk menghasilkan tangkapan ikan yang produktif.

**Keywords:** Data Oseanografi Indonesia, *Machine Learning*, Prediksi, Potensi, Sumber Daya Laut

# **MACHINE LEARNING INDONESIAN OCEANOGRAPHIC DATA FOR PREDICTING MARINE RESOURCE POTENTIAL**

## **ABSTRACT**

Marine data and information are very important for human survival, therefore this data and information is attractive to investors because of the potential economic value. Existing marine data and information have not been managed and have not been used to predict marine resources using Machine Learning (ML) techniques, so that mastery of marine data and information has become a problem in marine resource conservation. Correct, sustainable conservation of marine resources is carried out by controlling marine data and information. Currently, marine scientific research has entered a new era of Artificial Intelligence (AI), as the use of ML in marine scientific research is quite promising, especially for monitoring marine biodiversity, modeling CTD data, and predicting marine resources. ML to build predictive models is currently widely used by researchers, however research specifications using ML techniques to predict marine resource potential is a relatively new research field. The aim of this research is to create a model of Indonesian oceanographic data to predict marine resource potential using ML techniques using the Decision Tree (DT), Linear Regression (LR) and Random Forest (RF) algorithms, all three to be compared. The methodology is that the raw data that has been collected must be converted into Conductivity, Temperature, Depth (CTD) oceanographic data that is complete with information. The data is classified into training data and test data with Orange3 software. Training data is used for modeling with algorithms, thereby producing a data model. The model is tested with test data to produce a model with the best score and is then used to predict target data. The prediction results of the three models show a salinity value of 34 parts per thousand (ppt), and the actual value is 34 ppt. This means that in marine waters with a salinity value of 34 ppt, salinity will affect the solubility of oxygen (O<sub>2</sub>) and play a major role in the sustainability and growth of the fertility level of biological resources which is supported by sea surface temperature conditions of 29°C. So the salinity value of 34 ppt predicted using ML techniques is assumed to have a strong correlation with the presence of marine resource potential. This research has the potential to contribute data and information as a reference for innovative further research for researchers and investment decision material for investors as well as guidelines for fishermen to produce productive fish catches.

**Keywords:** Machine Learning, Marine Resources, Indonesian Oceanographic Data, Prediction, Potential