

**PENENTUAN ALGORITMA *SIMILARITY* YANG AKURAT  
PADA SISTEM BERBASIS *CASE BASED-REASONING (CBR)*  
UNTUK IDENTIFIKASI JENIS IKAN**

**TESIS**

**Diajukan Sebagai Salah Satu Syarat Untuk Menyelesaikan  
Program Strata Dua (S2) Magister Komputer**



**OLEH:**

**S U H A D I**

**341208736**

**PROGRAM STUDI TEKNIK INFORMATIKA  
PROGRAM PASCA SARJANA (S2) MAGISTER KOMPUTER  
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER ERESHA  
JAKARTA  
2011**

## PERSETUJUAN TESIS

Nama : Suhadi  
NPM : 341208736  
Konsentrasi : *Software Engineering*  
Judul Tesis : Penentuan Algoritma *Similarity* Yang Akurat Pada Sistem Berbasis  
*Case Based-Reasoning (CBR)* Untuk Identifikasi Jenis Ikan

Telah disetujui untuk disidangkan pada Sidang Tesis pada Program Pasca Sarjana (S2) Magister Komputer, Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer Eresha.

Jakarta, 26 Maret 2011

Pembimbing Utama

Pembimbing Pendamping

(Romi Satria Wahono, B.Eng, M.Eng)

(Dr. Rufman Iman Akbar Effendi, SE, M.Kom)

Mengetahui,  
Direktur  
Program Pasca Sarjana

(Dr. Rufman Iman Akbar Effendi, SE, M.Kom)

## PENGESAHAN TESIS

Nama : Suhadi  
NPM : 341208736  
Konsentrasi : *Software Engineering*  
Judul Tesis : Penentuan Algoritma *Similarity* Yang Akurat Pada Sistem Berbasis *Case Based-Reasoning (CBR)* Untuk Identifikasi Jenis Ikan

Telah disidangkan dan dinyatakan Lulus Sidang Tesis pada Program Pasca Sarjana (S2) Magister Komputer, Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer Eresha pada tanggal 26 Maret 2011

Nama Penguji

Tanda Tangan

1. Dr. Rufman Iman Akbar Effendi, SE, M.Kom .....  
(*Ketua*)
2. Ir. Damsiruddin Siregar, M.MT .....  
(*Sekretaris*)
3. Bobby Reza, S.Kom, MM .....  
(*Anggota*)

Mengetahui :  
Direktur  
Program Pasca Sarjana

(Dr. Rufman Iman Akbar Effendi, SE, M.Kom.)

## PERNYATAAN KEASLIAN TESIS

Nama : Suhadi  
NPM : 341208736  
Konsentrasi : *Software Engineering*  
Judul Tesis : Penentuan Algoritma *Similarity* Yang Akurat Pada Sistem Berbasis  
*Case Based-Reasoning (CBR)* Untuk Identifikasi Jenis Ikan

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

Jakarta, 26 Maret 2011

Suhadi

Suhadi, 341208736

**Penentuan Algoritma Similarity Yang Akurat Pada Sistem Berbasis Case Based-Reasoning (CBR) Untuk Identifikasi Jenis Ikan; dibawah bimbingan Romi Satria Wahono, B.Eng, M.Eng dan Dr. Rufman Iman Akbar Effendi, SE, M.Kom.  
78+ xiii hal/ 22 tabel/ 13 grafik/ 14 gambar/ 9 lampiran/10 pustaka (1962-2009)**

## ABSTRAK

*Indonesia sebagai negara kepulauan memiliki potensi perikanan yang sangat besar dan beragam, Indonesia memiliki 17.508 pulau dengan garis pantai sepanjang 81.000 km dan 70% (5,8 juta km<sup>2</sup>) dari luas Indonesia adalah lautan, adapun keragaman sumberdaya laut untuk beragam jenis ikan merupakan ciri tersendiri untuk mengenali dan memahami suatu spesies bagi enumerator/petugas lapangan di kabupaten/kota. Identifikasi jenis ikan yang bersifat computing masih terbatas salah satu sistem yang digunakan dalam penelitian ini adalah sistem Case Based-Reasoning (CBR).*

*CBR merupakan penalaran berbasis kasus adalah salah satu metode penyelesaian masalah berbasis pengetahuan untuk mempelajari dan memecahkan masalah berdasarkan pengalaman masa lalu. CBR adalah suatu model komputasi untuk meniru penalaran manusia untuk memberikan kemudahan dalam mencari kasus berdasarkan kemiripan, kemudian case-based reasoner mencari kasus-kasus yang ada pada basis kasus untuk menemukan kasus yang memiliki kemiripan dengan persoalan yang sedang dihadapi (retrieve). Oleh karena itu, proses CBR sering juga disebut dengan istilah “4 Re” yaitu Retrive, Retain, Revise, Reuse. CBR adalah suatu paradigma pemecahan masalah dimana sebuah permasalahan baru diselesaikan dengan cara membandingkan dengan kasus-kasus pada dimasa lampau dan menggunakannya kembali kasus-kasus yang ada untuk menyelesaikan suatu masalah sekarang.*

*Tujuan dari penelitian ini adalah untuk, melakukan evaluasi dan komparasi Algoritma Weighted Euclidean Distance (WED), Algoritma Hamming and Levenshtein Distances (HLD), Algoritma Cosine Coefficient for Text-Based Cases (CCFTBC) dan Algoritma k-Nearest Neighbor (k-NN) untuk identifikasi jenis ikan. Hasil komparasi algoritma yang akurat yang pertama adalah algoritma k-Nearest Neighbor (k-NN), yang kedua adalah algoritma Cosine Coefficient for Text-Based Cases (CCFTBC), yang ketiga adalah algoritma Hamming and Levenshtein Distances (HLD) dan yang keempat adalah algoritma Weighted Euclidean Distance (WED).*

**Kata Kunci:** *Cased Based-Reasoning, Similarity, Komparasi, Evaluasi, Weighted Euclidean Distance, Hamming and Levenshtein, Cosine Coefficient for Text-Based Cases, k-Nearest Neighbor.*

**Suhadi, 341208736**

***The Accurate Determination of Similarity Algorithm Based System Case-Based Reasoning (CBR) To Identify Types of Fish; under the guidance of Romi Satria Wahono, B. Eng, M. Eng and Dr. Rufman Akbar Effendi, SE, M. Kom.***

**78 + xiii page/ 22 tables/ 13 graphs/ 14 pictures/ 9 attachment/10 bibliography (1962-2009)**

### **ABSTRACT**

*Indonesia as an archipelagic country has the potential for fishery is very large and diverse, the Indonesia has 17,508 islands with a coastline of 81,000 km and 70% (5.8 million km<sup>2</sup>) of Indonesia is a vast ocean, while the diversity of marine resources for a variety of fish species is a distinctive feature of to recognize and understand a species for enumerators/field officers at district/city. Identification of fisheries species that are still limited computing one system used in this study is the Case-Based Reasoning (CBR).*

*CBR is a case-based reasoning is a knowledge-based problem solving methods to study and solve problems based on past experience. CBR is a computational model to mimic human reasoning to facilitate the search for cases based on similarity, then case-based Reasoner looking for cases that exist in the base case to find cases that have similarities with the problem at hand (retrieve). Therefore, the CBR process is often also referred to as "4 Re" ie Retrive, Retain, Revise, Reuse. CBR is a problem solving paradigm in which a new problem solved by comparing with the cases in the past and re-use of existing cases to solve a problem now.*

*The purpose of this research is to, conduct evaluation and comparison Weighted Euclidean Distance (WED) algorithm, Hamming and Levenshtein Distances (HLD) algorithm, Cosine Coefficient Algorithm for Text-Based Cases (CCFTBC) algorithm and k-Nearest Neighbor (k-NN) algorithm for identification of fish species. Accurate comparison results of the first algorithm is the of k-Nearest Neighbor algorithm (k-NN) algorithm, the second is the Cosine Coefficient for Text-Based Cases (CCFTBC) algorithm, the third is the Hamming and Levenshtein Distances (HLD) algorithm and the fourth is Weighted Euclidean Distance (WED) algorithm.*

**Keyword:** *Cased Based-Reasoning, Similarity, Comparison, Evaluation, Weighted Euclidean Distance, Hamming and Levenshtein, Cosine Coefficient for Text-Based Cases, k-Nearest Neighbor.*

## KATA PENGANTAR

Tesis dengan judul “**Penentuan Algoritma *Similarity* Yang Akurat Pada Sistem *Berbasis Case Based-Reasoning (CBR)* Untuk Identifikasi Jenis Ikan**”, penulis dapat menyelesaikan sesuai rencana, tidak lain adalah dukungan dari semua pihak yang telah membantu selesainya penulisan tesis ini. Tesis ini merupakan salah satu persyaratan yang diajukan dalam rangka menempuh ujian akhir untuk memperoleh gelar Magister Komputer (M.Kom) pada Program Pasca Sarjana (S2), Program Studi Teknik Informatika di Sekolah Tinggi Manajemen Informatika dan Komputer Eresha

Dalam kesempatan ini penulis menghaturkan penghargaan dan ucapan terima kasih yang sebesar-besarnya kepada:

1. Istri tercinta dan anak-anakku tersayang (Avif dan Faizh), yang telah memberikan motivasi dan dukungan moril ayahanda dalam menyelesaikan pendidikan program Pascasarjana dengan baik.
2. Bapak Ir. Damsiruddin Siregar, M.MT, sebagai Ketua STMIK Eresha Jakarta, yang telah memberikan kesempatan belajar bagi penulis.
3. Bapak Dr. Rufman Iman Akbar Effendi, SE, M.Kom, sebagai Direktur Program Pasca Sarjana STMIK Eresha Jakarta, sekaligus Pembimbing Pendamping, yang memberikan semangat untuk menyelesaikan Tesis ini.
4. Bapak Romi Satria Wahono, B.Eng, M.Eng, sebagai Pembimbing Utama, yang telah membagi ilmu pengetahuan dan pengalamannya dibidang teknologi dan informasi.
5. Rekan-rekan Angkatan 34, yang kompak selalu, telah banyak memberikan motivasi, semangat belajar serta dukungan moril dalam menyelesaikan Tesis ini.
6. Seluruh Staf Eresha, yang memberikan dukungan moril dan bantuan dalam proses pembelajaran akademik.

Akhir kata, dengan segala keterbatasan yang ada pada penulis tentunya masih banyak kekurangan dan masih jauh dari kesempurnaan, hanya Allah SWT yang memiliki segala kesempurnaanNya, oleh sebab itu kritik dan saran masukan yang bersifat membangun akan sangat membantu penulis untuk memperbaiki pada penelitian selanjutnya.

Penulis

## DAFTAR ISI

	Halaman
Lembar Persetujuan .....	i
Lembar Pengesahan .....	ii
Lembar Pernyataan Keaslian Tesis .....	iii
Abstrak (Indonesia dan Inggris) .....	iv
Kata Pengantar .....	vi
Daftar Isi .....	vii
Daftar Tabel .....	ix
Daftar Grafik .....	x
Daftar Gambar .....	xi
Daftar Lampiran .....	xii
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Permasalahan Penelitian .....	3
1.2.1 Identifikasi Masalah .....	3
1.2.2 Ruang Lingkup Masalah .....	4
1.2.3 Rumusan Masalah .....	4
1.3 Tujuan dan Manfaat Penelitian .....	4
1.3.1 Tujuan .....	4
1.3.2 Manfaat Penelitian .....	4
BAB II LANDASAN TEORI DAN KERANGKA PEMIKIRAN .....	6
2.1 Tinjauan Studi .....	6
2.2 Landasan Teori .....	7
2.2.1 Ikan ( <i>Fishes</i> ) .....	7
2.2.2 Morfologi Ikan ( <i>Fish Morfology</i> ) .....	8
2.2.3 Metode Identifikasi Jenis Ikan .....	11
2.3 <i>Case Based-Reasoning (CBR)</i> .....	16
2.3.1 Pengertian CBR .....	17
2.3.2 Algoritma Similarity .....	22
2.3.3 Algoritma <i>Weighted Euclidean Distance (WED)</i> .....	22

2.3.4	Algoritma <i>Hamming and Levenshtein (HLD)</i>	23
2.3.5	Algoritma <i>Cosine Coefficient for Text-Based Cases (CCFTBC)</i>	23
2.3.6	Algoritma <i>k-Nearest Neighbor (k-NN)</i>	24
2.4	Kerangka Pemikiran	26
BAB III	METODE PENELITIAN	29
3.1	Analisa Kebutuhan	29
3.1.1	Penentuan Masalah	29
3.1.2	Metode Pengumpulan Data	29
3.1.3	Sumber Data	30
3.1.4	Alat Bantu Penelitian	30
3.2	Perancangan Penelitian	30
3.2.1	Penentuan Penelitian	30
3.2.2	Penerapan Algoritma <i>Similarity</i>	38
3.3	Teknik Analisis	45
3.3.1	Teknik Analisis Pengujian Algoritma WED	45
3.3.1	Teknik Analisis Pengujian Algoritma HLD	49
3.3.1	Teknik Analisis Pengujian Algoritma CCFTBC	52
3.3.1	Teknik Analisis Pengujian Algoritma k-NN	55
BAB IV	HASIL DAN PEMBAHASAN	59
4.1	Hasil	59
4.1.1	Penentuan Hasil Sampel	59
4.2	Pembahasan	62
4.2.1	Metode Evaluasi	62
4.2.2	Metode Komparasi	71
4.2.3	Analisis Hasil	72
4.3	Implikasi Penelitian	73
BAB V	KESIMPULAN DAN SARAN	76
5.1	Kesimpulan	76
5.2	Saran	76

DAFTAR PUSTAKA .....	77
DAFTAR RIWAYAT HIDUP .....	78
LAMPIRAN-LAMPIRAN .....	79

## DAFTAR TABEL

	Halaman
Tabel 2.1 Contoh hasil penelusuran famili <i>Scombridae</i> dengan pohon keputusan .....	14
Tabel 3.1 Penentuan jumlah sampel dari populasi tertentu dengan taraf kesalahan 1%, 5% dan 10% .....	31
Tabel 3.2 Teknik perhitungan populasi .....	36
Tabel 3.3 Pengelompokan bagian jenis ikan .....	37
Tabel 3.4 Hasil perhitungan menurut bagian kasus jenis ikan .....	38
Tabel 3.5 Faktor Pada Bagian Kasus Jenis Ikan .....	39
Tabel 3.6 Faktor A1 atau Jenis Kepala .....	39
Tabel 3.7 Faktor A2 atau Bagian Sirip .....	39
Tabel 3.8 Faktor A3 atau Bagian Sisik .....	40
Tabel 3.9 Faktor A4 atau Bagian Badan .....	40
Tabel 3.10 Faktor A5 atau Bagian Ekor .....	40
Tabel 3.11 Faktor A6 atau Bagian Warna .....	41
Tabel 3.12 Jumlah Nilai Kemiripan .....	42
Tabel 3.13 Basis kasus yang tersimpan pada jenis ikan .....	43
Tabel 3.14 Kriteria Kemiripan .....	43
Tabel 3.15 Hasil kemiripan untuk semua jenis ikan .....	44
Tabel 3.16 Perbandingan penghitungan nilai awal dengan hasil perhitungan algoritam WED .....	47
Tabel 3.17 Perbandingan penghitungan nilai awal dengan hasil perhitungan algoritam HLD .....	50
Tabel 3.18 Perbandingan penghitungan nilai awal dengan hasil perhitungan algoritam CCFTBC .....	54
Tabel 3.19 Perbandingan penghitungan nilai awal dengan hasil perhitungan algoritam k-NN .....	57
Tabel 4.1 Hasil teknik pengambilan sampel .....	59
Tabel 4.2 Hasil penghitungan jenis ikan berdasarkan kriteria kemiripan .....	63
Tabel 4.3 Hasil kemiripan dan kriteria untuk semua jenis ikan .....	70

## DAFTAR GRAFIK

Halaman

Grafik 3.1 Perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma WED .....	48
Grafik 3.2 Perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma HLD .....	51
Grafik 3.3 Perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma CCFTBC .....	55
Grafik 3.4 Perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma k-NN .....	58
Grafik 4.1 Hasil Analisis perbedaan $\Sigma$ populasi awal dan $\Sigma$ populasi akhir .....	62
Grafik 4.2 Hasil Persentasi kemiripan berdasarkan jumlah sampel .....	63
Grafik 4.3 Hasil perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma WED .....	64
Grafik 4.4 Hasil analisis nilai desimal penghitungan dengan menggunakan algoritma WED .....	65
Grafik 4.5 Hasil perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma HLD .....	66
Grafik 4.6 Hasil analisis nilai desimal penghitungan dengan menggunakan algoritma HLD .....	66
Grafik 4.7 Hasil perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma CCFTBC .....	67
Grafik 4.8 Hasil analisis nilai desimal penghitungan dengan menggunakan algoritma CCFTBC .....	68
Grafik 4.9 Hasil perbandingan penghitungan nilai awal (cluster) dengan hasil perhitungan algoritma k-NN .....	69
Grafik 4.10 Hasil analisis nilai desimal penghitungan dengan menggunakan algoritma k-NN .....	69
Grafik 4.11 Hasil analisis persentasi masing-masing algoritma berdasarkan kriteria kemiripan.....	71
Grafik 4.12 Hasil analisis akurasi perbandingan algoritma WED, HLD, CCFTBC dan k-NN pada sistem CBR .....	72
Grafik 4.13 Hasil analisis akurasi perbandingan algoritma WED, HLD, CCFTBC dan k-NN pada sistem CBR .....	73

## DAFTAR GAMBAR

	Halaman
Gambar 2.1 <i>Morfology</i> jenis ikan secara umum .....	8
Gambar 2.2 Kepala Jenis Ikan .....	9
Gambar 2.3 Sirip Jenis Ikan .....	9
Gambar 2.4 Sisik Jenis Ikan .....	10
Gambar 2.5 Ekor Jenis Ikan .....	10
Gambar 2.6 Badan Jenis Ikan .....	11
Gambar 2.7 Warna Jenis Ikan.....	11
Gambar 2.8 Contoh model identifikasi jenis ikan Tuna dengan famili <i>combridae</i> .....	12
Gambar 2.9 Pohon Keputusan .....	13
Gambar 2.10 <i>Flowchart</i> proses identifikasi jenis ikan .....	15
Gambar 2.11 Model Siklus <i>Case-Based Reasoning (CBR)</i> .....	17
Gambar 2.12 Contoh pencarian metode k-NN .....	25
Gambar 2.13 Kerangka Pemikiran .....	27
Gambar 3.1 Alur proses evaluasi dan komparasi pada algoritma sistem CBR ...	32
Gambar 4.1 Hasil teknik pengambilan sampel dengan kesalahan 5% .....	32

## DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Tabel Hasil Perhitungan Menurut Bagian Kasus Jenis Ikan .....79
Lampiran 2	Tabel Hasil Penghitungan Jumlah Nilai Kemiripan Pada Bagian Kasus ..... 87
Lampiran 3	Tabel Hasil Penghitungan Basis Kasus Yang Tersimpan Pada Jenis Ikan ..... 93
Lampiran 4	Tabel Hasil Penghitungan Kemiripan Untuk Jenis Ikan ..... 101
Lampiran 5	Tabel Hasil Perhitungan Algoritma <i>Weighted Euclidean Distance (WED)</i> ..... 109
Lampiran 6	Tabel Hasil Perhitungan Algoritma <i>Hamming And Levenshtein Distances (HLD)</i> ..... 115
Lampiran 7	Tabel Hasil Perhitungan Algoritma <i>Cosine Coefficient For Text-Based Cases (CCFTBC)</i> ..... 121
Lampiran 8	Tabel Hasil Perhitungan Algoritma <i>k-Nearest Neighbor (k-NN)</i> ..... 127
Lampiran 9	Tabel Hasil Perhitungan Hasil Kemiripan dan Kriteria Untuk Semua Jenis Ikan ..... 133

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Indonesia sebagai negara kepulauan memiliki potensi perikanan yang sangat besar dan beragam, Indonesia memiliki 17.508 pulau dengan garis pantai sepanjang 81.000 km dan 70% (5,8 juta km<sup>2</sup>) dari luas Indonesia adalah lautan, adapun keragaman sumberdaya laut untuk jenis ikan diketahui terdapat 8.500 jenis ikan. Komisi Nasional Pengkajian Sumberdaya Perikanan Laut melaporkan bahwa potensi lestari sumberdaya perikanan laut Indonesia adalah 6,4 juta ton per tahun dengan porsi terbesar dari jenis ikan pelagis kecil yaitu sebesar 3,2 juta ton (52,54%), jenis ikan demersal 1,8 juta ton (28,96%) dan perikanan pelagis besar 0,97 juta ton (15,81%). Potensi perikanan tangkap diperkirakan mencapai 6,26 juta ton per tahun dengan jumlah tangkapan yang diperbolehkan sebesar 5,007 juta ton atau 80% dari *Maximum Sustainable Yield (MSY)* [1]. Potensi sumberdaya perikanan yang besar tersebut dapat dimanfaatkan untuk meningkatkan kesejahteraan masyarakat tetapi potensi tersebut belum dioptimalkan secara maksimal.

Dalam Undang-Undang Republik Indonesia Nomor 9 Tahun 1985 tentang Perikanan disebutkan bahwa yang dimaksud dengan "**ikan**" adalah:

- a. *Pisces* (ikan bersirip)
- b. *Crustacea* (udang, rajungan, kepiting dan sejenisnya)
- c. *Mollusca* (kerang, tiram, cumi-cumi, gurita, siput dan sejenisnya)
- d. *Coelenterata* (ubur-ubur dan sejenisnya)
- e. *Echinodermata* (teripang, bulu babi dan sejenisnya)
- f. *Amphibi* (kodok dan sejenisnya)
- g. *Reptilia* (buaya, penyu, kura-kura dan sejenisnya)
- h. *Mammalia* (paus, lumba-lumba, pesut, duyung dan sejenisnya)
- i. *Algae* (rumput laut dan tumbuh-tumbuhan lain yang hidup di dalam air)
- j. Biota air lainnya yang ada kaitannya dengan jenis-jenis tersebut di atas

Ikan sebagai salah satu organisme yang menjadi kajian ekologi, sehingga harus dijaga kelestariannya, identifikasi terhadap organisme tersebut dengan menempatkan atau memberikan identitas suatu individu melalui prosedur deduktif ke dalam suatu taksonomi dengan menggunakan kunci determinasi. Kunci determinasi adalah kunci jawaban yang digunakan untuk menetapkan identitas suatu individu, kegiatan identifikasi bertujuan untuk mencari dan mengenal ciri-ciri taksonomi yang sangat bervariasi dan memasukkannya ke dalam suatu taksonomi. Selain itu untuk mengetahui identitas atau nama suatu individu atau spesies dengan cara mengamati beberapa karakter atau ciri morfologi spesies tersebut dengan membandingkan ciri-ciri yang ada sesuai dengan kunci determinasi [2].

Ikan merupakan produksi hasil tangkapan dari Nelayan/Perusahaan Perikanan (PP)/Rumah Tangga Perikanan (RTP) yang hasilnya untuk dijual. Ikan dijual oleh Nelayan/PP/RTP dengan cara melelang ke pembeli/tengkulak di Tempat Pendaratan Ikan (TPI) yang berlokasi di Pelabuhan Perikanan setempat, bersamaan dengan lelang tersebut petugas statistik perikanan tangkap (enumerator) akan mencatat semua jenis ikan yang akan dijual/lelang sehingga akan diketahui jumlah produksi per jenis ikan, jenis ikan yang ditangkap dan dari daerah mana hasil penangkapan ikan tersebut. Pencatatan yang dilakukan enumerator akan diidentifikasi dan diklasifikasikan berdasarkan spesiesnya untuk mengidentifikasi jenis ikan enumerator harus mempunyai pengetahuan tentang jenis-jenis ikan.

Enumerator adalah petugas penyuluh lapangan yang melaksanakan pengumpulan data khusus data statistik perikanan tangkap yang berada wilayah Kabupaten/Kota, yang berasal dari instansi dinas yang menangani Kelautan dan Perikanan di Kabupaten/Kota. Pengumpulan data statistik perikanan tangkap oleh enumerator adalah data jenis ikan. Dilapangan enumerator mengalami kesulitan dalam pengumpulan data jenis ikan, kesulitannya adalah mengidentifikasi jenis ikan [3], dengan pengetahuan tentang identifikasi jenis ikan sangat kurang, dikarenakan enumerator dilapangan bervariasi disiplin ilmunya serta tingkat pendidikan yang berbeda, sehingga pemahaman pengetahuan tentang identifikasi jenis ikan masih kurang, enumerator inilah dituntut untuk menguasai teknik identifikasi jenis ikan tersebut dengan cermat dan akurat.

Identifikasi jenis ikan juga sangat sulit dilakukan terutama di perairan Zona Ekonomi Eksklusif Indonesia (ZEEI), kesulitan menentukan daerah potensial sebagai daerah penangkapan ikan (*fishing ground*) termasuk didalamnya jenis ikan yang ditangkap sehingga teknologi yang digunakan sekarang ini adalah menggunakan penginderaan jarak jauh, pada umumnya nelayan di Indonesia masih menggunakan cara-cara konvensional, yaitu hanya dengan memanfaatkan panca indera yang dimiliki oleh nelayan. Keterbatasan panca indera nelayan dalam menduga *fishing ground* dan jenis ikan yang ditangkap tidak hanya menyebabkan inefisiensi penggunaan bahan bakar sebanyak 60%-70%, tetapi juga menyebabkan terkonsentrasinya kapal-kapal penangkap ikan di lokasi tertentu. Sebagai akibatnya pada daerah tertentu terjadi pengeksploitasian secara berlebihan (*over fishing*). Jika hal ini dibiarkan terus menerus dalam jangka waktu tertentu kelestarian sumberdaya perikanan akan terganggu, sebaliknya pada daerah yang memiliki potensi ikan yang cukup besar justru tidak dimanfaatkan secara optimal [4].

Pada penelitian ini akan dilakukan evaluasi dan komparasi algoritma similarity pada sistem CBR yang meliputi Algoritma Weighted Euclidean Distance (WED), Algorithm Hamming and Levenshtein Distances (HLD), Algorithm *Cosine Coefficient for Text-Based Cases (CCFTBC)* dan Algorithm *k-Nearest Neighbor (k-NN)*, untuk mengidentifikasi jenis ikan enumerator membandingkan dan memperkirakan jenis-jenis ikan pada buku panduan klasifikasi jenis ikan hal ini sulit dilakukan jika ada spesies baru yang belum ada dibuku dan jika identifikasi jenis ikan ini dilakukan oleh orang awam/orang yang belum mengetahui jenis ikan secara spesifik, maka identifikasi jenis ikan kurang akurat yang berakibat pada pengelompokan jenis ikan juga kurang akurat mempunyai dampak perbedaan persepsi dalam hal produksi per jenis ikan (ikan lainnya akan lebih besar produksinya).

## **1.2 Permasalahan Penelitian**

### **1.2.1 Identifikasi Masalah**

Identifikasi masalah yang mendasari penelitian ini adalah sebagai berikut:

- a. Bagaimana mengatasi kesulitan enumerator dalam melakukan identifikasi jenis ikan dilapangan karena sulit, tidak akurat dan memerlukan waktu yang lama?

- b. Bagaimana mengakurasi identifikasi jenis ikan dilapangan secara akurat dalam *Case Based-Reasoning* untuk penyelesaian masalah?

### **1.2.2 Ruang Lingkup Masalah**

Enumerator dilapangan kesulitan dan memerlukan waktu yang lama dalam melakukan identifikasi jenis ikan, sehingga dalam penelitian ini akan dilakukan evaluasi dan komparasi algoritma *similarity* pada sistem CBR, algoritma yang akan di evaluasi dan dikomparasi adalah Algoritma WED, Algoritma HLD, Algoritma CCFTBC dan Algoritma k-NN.

### **1.2.3 Rumusan Masalah**

Rumusan masalah berdasarkan hasil dari identifikasi masalah dan ruang lingkup masalah yaitu proses identifikasi jenis ikan diperlukan waktu tidak akurat dan diperlukan analisis algoritma *similarity* yang akurat untuk identifikasi jenis ikan.

Sedangkan pertanyaan penelitian (*research question*) yang dipilih adalah algoritma *similarity* apa yang akurat diterapkan pada sistem CBR untuk identifikasi jenis ikan.

## **1.3 Tujuan dan Manfaat Penelitian**

Tujuan dan manfaat penelitian ini dapat memberikan hasil sebagai bahan referensi bagi pengembangan selanjutnya dengan rincian sebagai berikut:

### **1.3.1 Tujuan**

Penelitian ini bertujuan untuk, melakukan evaluasi dan komparasi Algoritma *Similarity* pada sistm CBR yaitu algoritma yang akan di evaluasi dan dikomparasi adalah Algoritma WED, Algoritma HLD, Algoritma CCFTBC dan Algoritma k-NN, sehingga dapat mengevaluasi akurasi proses pengidentifikasian jenis ikan.

### 1.3.2 Manfaat Penelitian

Adapun manfaat dari penulisan ini dapat memberikan manfaat antara lain sebagai berikut:

- a. Manfaat praktis dapat digunakan sebagai acuan untuk identifikasi jenis ikan di lapangan bagi petugas enumerator.
- b. Manfaat teoritis, diharapkan dapat memberikan masukan atas teori identifikasi penalaran berbasis komputer (*Cased Based-Reasoning*) khususnya berhubungan dengan algoritma similarity.
- c. Manfaat kebijakan, untuk memberikan masukan bagi pengambil kebijakan dalam rangka mengembangkan dan memberikan kontribusi untuk perbaikan masalah pengidentifikasian jenis ikan di lapangan

## BAB II

### LANDASAN TEORI DAN KERANGKA PEMIKIRAN

#### 2.1 Tinjauan Studi

Tinjauan pustaka diambil dari beberapa buku, jurnal dan artikel yang berhubungan dengan penelitian ini adalah sebagai berikut:

- a. Penelitian yang dilakukan oleh Yixin Chen dan kawan-kawan (Department of Computer Science University of New Orleans, New Orleans, LA US), tahun 2005, p237-244, dengan judul “*A Content-Based Image Retrieval System for Fish Taxonomy*”. Penelitian ini memanfaatkan gambar konten digital berbasis taksonomi untuk mengidentifikasi dan pembelajaran jenis ikan dengan menggunakan pendekatan algoritma *similarity* pada sistem *Case Based Image Reasoning* (CBIR) untuk mendapatkan identitas jenis ikan yang akurat.
- b. Penelitian yang dilakukan oleh Uma Murthy dan kawan-kawan (Department of Computer Science, Virginia Tech Blacksburg, USA) dan Ricardo Torres, Evandro J. Ramos, and Tiago R. C. Falcão (Institute of Computing University of Campinas, SP, Brazil), tahun 2009, p:435-436, dengan judul “*Species Identification: Fish Images with CBIR and Annotations*”. Penelitian ini memberikan metode untuk membandingkan gambar jenis ikan dengan informasi tekstualnya untuk identifikasi jenis ikan, dengan menggunakan pendekatan *Cased Based-Reasoning* informasi pada *retrival* dapat menjadikan pengetahuan baru.
- c. Penelitian yang dilakukan oleh Daniel Heesch, Alexei Yavlinsky, Stefan Rürger (Department of Computing, South Kensington Campus, Imperial College London, London, England), tahun 2003, p:456-466 dengan judul “*Performance comparison of different similarity models for CBIR with relevance feedback*. Penelitian ini memberikan kesimpulan bahwa perbandingan kinerja gambar konten dengan menggunakan algoritma, *k-Neighbours* dapat digunakan untuk menentukan teknik kesamaan antara dua image gambar yang berbeda, dengan teknik tersebut model vektor melalui umpan balik relevansi gambar akan menghasilkan kesamaan *retrival* pada sistem *Cased Based-Reasoning*.
- d. Penelitian yang dilakukan Marjo Markkula, Marius Tico, Katja Nirkkonen (Department of Information Studies, University of Tampere, Finland), tahun 2001,