

BAB I

PENDAHULUAN

1.1 Latar Belakang

Keberadaan lahan merupakan aspek penting dalam kehidupan manusia dan makhluk hidup lainnya, akan tetapi persoalan kerusakan hutan dan lahan terus terjadi dan mengalami peningkatan sehingga mengakibatkan lahan menjadi kritis (Ramayanti, Yuwono, & Awaluddin, 2015). Akibat pemanfaatan sumber daya alam yang berlebih dan tidak dibarengi dengan usaha konservasi tanah dan air, juga dapat menimbulkan ketidakseimbangan lingkungan yaitu terus meningkatnya luas lahan kritis (Nugroho, 2000). Dalam Keputusan Menteri Kehutanan Nomor 52/KptsII/2001 tentang pedoman penyelenggaraan pengelolaan Daerah Aliran Sungai dijelaskan bahwa lahan kritis merupakan lahan yang telah mengalami kerusakan sehingga lahan tersebut tidak dapat berfungsi secara baik sesuai dengan peruntukannya sebagai media produksi maupun sebagai media tata air (Prabandaru, Nugraha, & Sukmono, 2016). Penentuan lahan kritis dilakukan dengan pembagian 3 kawasan yaitu kawasan budidaya pertanian, kawasan hutan lindung dan kawasan lindung di luar kawasan hutan. Menentukan tingkat kekritisian lahan mengacu pada Peraturan Direktur Jenderal Bina Pengelolaan Daerah Aliran Sungai dan Perhutanan Sosial tentang petunjuk teknis penyusunan data spasial lahan kritis dengan peraturan nomor P.4/V-SET/2013, yaitu pemberian skor serta pembobotan tiap parameter pada masing-masing Kawasan (Rosyada, Prasetyo, & Hani'ah, 2015).

Laju kerusakan hutan di Indonesia tercatat telah mencapai 59,63 juta hektar. Yaitu Kawasan hutan konservasi mencapai 4,7 hektar, hutan lindung mencapai 10,5 juta hektar dan hutan alam produksi mencapai 44,4 juta hektar (Zuhud, 2009). Kementerian Kehutanan mencanangkan kegiatan Rehabilitasi Hutan dan Lahan (RHL) untuk mengatasi lahan kritis serta memulihkan, mempertahankan dan meningkatkan fungsi hutan dan lahan sehingga daya dukung, produktivitas dan peranannya sebagai penyangga kehidupan tetap terjaga (Indrihastuti, Murti Laksono, & Tjahjono, 2016). Oleh karena itu, peta dan data lahan kritis adalah suatu hal yang sangat penting untuk penyusunan program RHL agar efektif dan program berjalan dengan optimal, serta untuk harmonisasi program antar sektor yang terkait dengan rehabilitasi hutan dan lahan (Dephut, 2014).

Kabupaten Grobogan merupakan salah satu kabupaten di Provinsi Jawa Tengah dengan luas 1976 km². Kabupaten Grobogan tercatat memiliki lahan kritis seluas 428.687 ha. Sebesar 238.170 ha adalah lahan kritis di luar kawasan hutan dan 190.517 ha merupakan lahan kritis di dalam kawasan hutan (Ariyanto, Sunarminto, & Shiddieq, 2011). Berdasarkan laporan Badan Pusat Statistik tahun 2016 Kabupaten Grobogan menempati urutan pertama di Kabupaten Grobogan yang memiliki luas lahan kritis sebesar 203.131,10 ha (BPS, 2016). Parameter yang menjadi faktor kekritisannya lahan, yaitu tingkat bahaya erosi, produktivitas lahan, erosi lahan, manajemen lahan, singkapan batuan, kemiringan lereng dan curah hujan (Hendro HS, Nahdi, Budiastuti, & Purnomo, 2012). Pengklasifikasian lahan kritis akan berfungsi untuk mengelompokkan berbagai macam tingkatan lahan kritis berdasarkan rentang nilai hasil berbagai parameter (Prasetya & Gunawan, 2006).

Decision Tree merupakan metode klasifikasi yang telah banyak digunakan, karena merupakan salah satu metode yang populer (Handarko, 2015). *Decision tree* banyak digunakan karena kemampuan mereka dalam menghasilkan konsep yang jelas (Bhavitha & Madhuri, 2012). Kefleksibelan membuat metode ini atraktif, khususnya karena memberikan keuntungan berupa visualisasi saran (dalam bentuk *Decision tree*) yang membuat prosedur prediksinya dapat diamati (Gorunescu, 2011). Visualisasi saran tersebut dikenal sederhana dan menarik bahkan untuk praktisi awam yang tidak memiliki latar belakang statistik (Sohn & Kim, 2012). Metode ini juga mampu melakukan klasifikasi sekaligus menunjukkan hubungan antar atribut (Ariadi & Arieshanti, 2015). Induksi di *Decision Tree* telah dikembangkan secara mandiri di statistik dan di komunitas *machine learning* (Rahayu, Satria, & Supriyanto, 2015). Ada beberapa pengembangan algoritma yang telah dikembangkan berdasarkan *Decision Tree* seperti ID3 (*Iterative Dichotomiser 3*), C4.5, CART (*Classification and Regression Tree*), CHAID (*Chi-squared Automatic Interaction Detector*) (Gupta, Rawat, Jain, Arora, & Dhimi, 2017).

Adapun salah satu penelitian sebelumnya yang berkaitan dengan klasifikasi lahan kritis seperti penelitian yang dilakukan oleh Nur Khotimah pada tahun 2018 yang membandingkan metode *Decision Tree* Algoritma C4.5, *Naïve Bayes*, dan *K-Nearest Neighbour (k-NN)* untuk klasifikasi Lahan Kritis di Kabupaten Pemalang. Menggunakan *dataset* parameter kekritisian Kabupaten Pemalang tahun 2013, dengan jumlah 391 data dan memiliki 3 atribut dan satu label. Penelitian tersebut menghasilkan bahwa *Decision Tree* algoritma C4.5 memiliki akurasi tertinggi

sebesar 77.75%, disusul dengan *Naïve Bayes* sebesar 77.49%, dan terakhir k-NN memiliki akurasi sebesar 73.91% (Khotimah & Istiawan, 2018).

Pada penelitian terkait komparasi ID3, CART dan C4.5 yang dilakukan oleh Mohankumar, Amuthakkani, dan Jeyamala pada tahun 2016 dengan menggunakan dataset kelayakan pinjaman nasabah bank diperoleh hasil bahwa C4.5 memiliki hasil *precision*, *recall*, dan akurasi lebih baik daripada ID3 dan CART. Lim dan shan shih juga telah membandingkan tingkat akurasi, kompleksitas dan waktu training dari ID3, CART dan C4.5 dan hasilnya menunjukkan bahwa C4.5 mempunyai tingkat akurasi yang bagus dan mudah untuk diinterpretasikan (Lim, Loh, Shih, & Cohen, 2000).

Split atribut merupakan proses utama dalam pembentukan pohon keputusan (*Decision Tree*) di C4.5 (Zhang & Jiang, 2012). Ada dua metode pemilihan split atribut paling populer yaitu *information gain* dan *gain ratio* (Han, Kamber, & Pei, 2012). Metode *information gain* yang diperkenalkan Quinlan cocok untuk *dataset* dengan variabel diskret akan tetapi metode yang diperkenalkan tidak cocok untuk *dataset* dengan *missing value*, sedangkan metode *gain ratio* dapat bekerja untuk *dataset* dengan variabel kontinu dan *missing value*. Penelitian yang menggunakan *gain ratio* untuk split atribut Chaurasia pada tahun 2018 menggunakan *dataset Breast Cancer Wisconsin* dengan jumlah data 683, 10 atribut dan 1 label menghasilkan nilai akurasi sebesar 93.41%, nilai *sensifity* 93.4% dan nilai *specificity* sebesar 90.37% (Chaurasia, Pal, & Tiwari, 2018). Penelitian sebelumnya yang menggunakan *information gain* untuk split atribut yang dilakukan oleh Wisti Dwi Septiani pada tahun 2017 menggunakan data penyakit hepatitis yang didapat

dari *Machine Learning Repository UCI* (Universitas California Invene) dengan 155 data, 19 atribut dan 1 label menghasilkan nilai akurasi sebesar 77.29% (Septiani, 2017). Berdasarkan uraian diatas penulis akan menggunakan *Information Gain* dan *Gain Ratio* untuk split atribut pada Algoritma C4.5 dengan studi kasus klasifikasi lahan kritis di Kabupten Grobogan.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang di atas maka rumusan masalah yang diangkat dalam penelitian ini adalah:

1. Bagaimana hasil klasifikasi Algoritma C4.5 jika menggunakan metode pemilihan split atribut *Information Gain* untuk Klasifikasi Lahan Kritis di Kabupaten Grobogan?
2. Bagaimana hasil klasifikasi Algoritma C4.5 jika menggunakan metode pemilihan split atribut *Gain Ratio* untuk Klasifikasi Lahan Kritis di Kabupaten Grobogan?
3. Bagaimana perbandingan metode pemilihan split atribut *Information Gain* dan *Gain Ratio* pada klasifikasi Lahan Kristis di Kabupaten Grobogan?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan pada penelitaian yang hendak dicapai adalah:

1. Mengetahui hasil klasifikasi Algoritma C4.5 jika menggunakan metode pemilihan split atribut *Information Gain* untuk Klasifikasi Lahan Kritis di Kabupaten Grobogan.

2. Mengetahui hasil klasifikasi Algoritma C4.5 jika menggunakan metode pemilihan split atribut *Gain Ratio* untuk Klasifikasi Lahan Kritis di Kabupaten Grobogan.
3. Pemilihan metode split atribut terbaik antara *Information Gain* dan *Gain Ratio* pada *decision tree* Algoritma C4.5 untuk klasifikasi Lahan Kritis di Kabupaten Grobogan.

1.4 Manfaat Penelitian

Penelitian penerapan *gain ratio* dan *information gain* pada proses pemilihan split atribut pada *decision tree* algoritma C4.5 untuk klasifikasi lahan kritis di Kabupaten Grobogan diharapkan dapat memberikan manfaat sebagai berikut:

1. Manfaat praktis dari hasil penelitian ini adalah dapat dijadikan bahan pertimbangan dalam pengambilan keputusan masyarakat mengenai peristiwa tertentu.
2. Manfaat teoritis dari hasil penelitian ini yaitu diharapkan dapat memberikan kontribusi untuk pengembangan teori yang berkaitan dengan penerapan proses pemilihan split atribut di Algoritma C4.5

1.5 Batasan Masalah

Agar pembahasan tidak terlalu meluas, penulis perlu memberikan batasan permasalahan, adapun batasan masalahnya adalah metode pemilihan split atribut yang digunakan adalah *Information Gain* dan *Gain Ratio*, model validasi yang digunakan adalah *k-fold cross validation* dengan nilai k yaitu 10, data yang digunakan dalam penelitian kali ini berupa data lahan kritis tahun 2016 yang merupakan publikasi dan laporan terakhir BPDAS Pemali Jratun.