

## **BAB II**

### **TINJAUAN PUSTAKA**

Berbagai topik yang berkaitan dengan Penelitian kali ini telah banyak dilakukan untuk menyelesaikan berbagai masalah di segala bidang ilmu, sehingga dapat diartikan bahwa Klasifikasi berguna untuk mendapatkan Hasil yang terbaik. Menurut penelitian yang dilakukan (Amin et al., 2015) membahas mengenai sistem pendukung keputusan untuk persetujuan pemberian dana pinjaman menggunakan data nasabah yang diambil dari Bank Pasar Yogyakarta menggunakan metode C4.5 menghasilkan persetujuan debitur baru dilakukan dengan menggunakan C4.5 serta memiliki tingkat akurasi tinggi. Pada penelitian (Duan et al., 2009) membahas pengklasifikasian jenis pelanggan batubara menggunakan data internal yang diambil dari Shanxi Coal Transportation and Marketing Corporation menggunakan metode C4.5 menghasilkan pohon keputusan berdasarkan algoritma C4.5 untuk analisis pelanggan logistik batubara, mengadopsi Pessimistic Error Pruning(PEP).

Pernah dilakukan oleh (Azizah et al., 2018) uji komparasi Naïve Bayes dengan C4.5 pada kasus penggunaan media belajar online menggunakan dataset bebas dari The Open University England menghasilkan akurasi Naïve Bayes lebih baik walaupun perbedaan hanya tipis namun hasil recall yang lebih tinggi di dapatkan pada C4.5. Dalam percobaan (A.Anushya, 2020) mengenai klasifikasi kualitas buah pisang menggunakan data citra pisang yang diambil sendiri dengan ketentuan khusus menggunakan metode C4.5 dan Jaringan Syaraf Tiruan menghasilkan nilai akurasi yang lebih baik ketika menggunakan Jaringan Syaraf Tiruan. Hal lain pernah dilakukan oleh (Perwej et al., 2018) membuat klasifikasi klinis penderita penyakit jantung menggunakan data pasien yang diambil dari sumber yang tidak disebutkan menggunakan metode kombinasi ROCK dengan K-Means dan kombinasi antara ROCK dengan C4.5 menghasilkan akurasi yang tinggi pada klasifikasi penyakit jantung apabila ROCK di kombinasikan dengan C4.5.

Sebelumnya (Fadlil et al., 2019) pernah membuat sistem identifikasi jamur menggunakan data yang diambil sendiri menggunakan metode Jaringan Syaraf

Tiruan menghasilkan akurasi sistem yang baik. Seperti halnya (Gustina et al., 2017) yang membahas mengenai sistem identifikasi jamur menggunakan data primer yang diambil sendiri menggunakan metode klasifikasi jarak menghasilkan akurasi 82%.

Pada kasus yang diteliti oleh (Gustian & Hundayani, 2018) membahas tentang klasifikasi tingkat drop out mahasiswa menggunakan data yang diambil dari Universitas XYZ di Sukabumi menggunakan metode AHP dan C4.5 menghasilkan akurasi sebesar 97.67% pada C4.5. Pada uji coba yang dilakukan oleh (Zhen et al., 2011) pernah dibuat klasifikasi bidikan video menggunakan data primer yang diambil dari berbagai sumber menggunakan metode C4.5 menghasilkan C4.5 mampu mengklasifikasikan adegan yang tayang pada suatu cuplikan klip video. Sedangkan apa yang diteliti oleh (Yu et al., 2010) mengenai deteksi intisel dari citra mikroskopis kanker serviks menggunakan data citra sel dari serviks uteri menggunakan metode C4.5 menghasilkan akurasi klasifikasi yang menjanjikan, sebesar 97,8%. Ada lagi dalam kasus yang diteliti oleh (Li et al., 2016) yang membuat sebuah model prediksi kualitas tanah dan menggabungkan komposisi tanah di Lishu menggunakan data nutrisi tanah dari Distrik Lishu di Provinsi Jilin menggunakan metode C4.5 menghasilkan usulan agar algoritma C4.5 digunakan untuk membangun model data mining untuk memprediksi kualitas tanah. Hal lain pernah dilakukan pengujian oleh (Krasser et al., 2007) untuk membangun sistem identifikasi spam gambar berdasarkan header dan properti file menggunakan data spam email yang diambil dari sumber internal menggunakan metode C4.5 menghasilkan usulan bahwa klasifikasi ini dapat secara efektif berfungsi sebagai klasifikasi spam tingkat pertama dalam framework klasifikasi berlapis.

Percobaan lainnya oleh (Giaccom Ribeiro & Garcia Fonseca, 2013) dimana mereka membangun sistem klasifikasi tutupan lahan perkotaan Sao Paulo menggunakan data dari sensor WorldView-2 yang diambil dari hasil observasi yang mereka lakukan menggunakan metode C4.5 menghasilkan C4.5 terintegrasi dengan alat analisis gambar lainnya, menghadirkan potensi besar dalam analisis gambar resolusi tinggi untuk diterapkan di perkotaan. Juga pernah diteliti oleh (Pertiwi et al., 2018) tentang klasifikasi tingkat putus sekolah tiap provinsi di

Indonesia menggunakan dataset yang diambil dari laman data.go.id menggunakan metode C4.5 menghasilkan akurasi yang baik sebesar 71.2%.

C4.5 juga dapat diteliti seperti yang dilakukan oleh (Qian & Qiu, 2014) untuk menangani pemilihan fitur untuk prediksi harga listrik menggunakan dataset yang diambil dari New South Wales electricity dan the Spanish electricity menggunakan metode C4.5 menghasilkan pembuktian bahwa menggunakan C4.5 lebih baik daripada yang menggunakan ID3. Dan juga seperti yang dilakukan percobaan oleh (Sucipto et al., 2017) yang mengimplementasikan metode klasifikasi multikelas pada penyakit ikan menggunakan data penyakit ikan yang diambil dari sumber yang tidak disebutkan menggunakan metode C4.5 menghasilkan klasifikasi yang baik pada tiap kelas penyakit ikan dengan tingkat akurasi tertinggi sebesar 88.4%.

Penelitian terkait Klasifikasi telah banyak dilakukan oleh peneliti sebelumnya yang dirangkum dalam Tabel 2.1 berikut.

**Tabel 2.1** Tinjauan Pustaka

No	Publikasi (Penulis, Tahun)	Domain Penelitian	Data	Kontribusi
1	Rafik Khairul Amin, et al, 2015	Implementasi Decision Tree Menggunakan Algoritma C4.5 dalam Pengambilan Keputusan Aplikasi Pinjaman oleh Debitur (Studi Kasus: Bank Pasar Daerah Istimewa Yogyakarta)	Data Nasabah	Penerapan penetapan persetujuan debitur baru dilakukan dengan menggunakan pohon keputusan yang menghasilkan tingkat akurasi tertinggi
2	Fu Duan, et al, 2009	Algoritma C4.5 Untuk Mengklasifikasikan Pelanggan Batubara	Database Shanxi Coal	Hasil klasifikasi tersebut layak dan memiliki nilai praktis yang penting
3	Erwina Nurul Azizah, et al, 2018	Komparasi Kinerja antara pengklasifikasi C4.5 dan Naive Bayes dalam memprediksi	Dataset aktifitas web	Mendapatkan tingkat akurasi yang mirip, namun recall C4.5 lebih baik

		kinerja akademik siswa di Lingkungan Belajar Virtual		
4	A Anusya, 2020	Mengklasifikasikan kualitas pisang berdasarkan tekstur	Primer	Dalam penelitian ini gambar pisang ditangkap dan diklasifikasikan ke dalam kelompok yang berbeda dengan Decision Tree dan Neural Networks menggunakan fitur GLCM.
5	Yusuf Perwej, et al, 2018	Mengklasifikasikan tingkat dan jenis penyakit jantung	data primer	Akurasi K-means berdasarkan ROCK 76%, K-Mean berdasarkan ROCK dengan ID3 87%, dan K-Mean berdasarkan ROCK dengan algoritma klasifikasi C4.5 95%.
6	Abdul Fadlil, et al, 2019	Identifikasi Gambar Jamur Menggunakan Ekstraksi Fitur Statistik Orde 1 dengan Teknik Klasifikasi Jaringan Saraf Tiruan	Data Tes	Keakuratan sistem untuk mengidentifikasi dan mengenali citra jamur sangat dipengaruhi oleh parameter jaringan internal.
7	Dudih Gustian, et al, 2017	Kombinasi Metode AHP Dengan C4.5 : Klasifikasi Level Drop Out Mahasiswa	Data Tes	Menghasilkan luaran yang sama
8	Sapriani Gustina, et al, 2017	Membuat Sistem Identifikasi Jamur Menggunakan Metode Ekstraksi Ciri Statistik Orde 1 dan Klasifikasi Jarak	Observasi	Hasil eksperimen dari pengujian sistem menunjukkan tingkat akurasi yang terbaik 82% diperoleh dari metode klasifikasi jarak Manhattan.
9	Yao Zhen, et al, 2011	Mengklasifikasikan teknik pengambilan video dalam beberapa acara	Data Video	C4.5 mampu mengklasifikasikan program komentar

				sesuai dengan topik acara olahraga.
10	Yu Peng, et al, 2010	Mengklasifikasikan sel kanker serviks menggunakan C4.5	Data Penelitian Internal	Mempu mengklasifikasikan citra sel dengan baik dengan tingkat akurasi hasil klasifikasi sebesar 97%
11	Li Dongming, et al, 2016	Penerapan Decision Tree C4.5 Algoritma untuk Model Peramalan Grade Kualitas Tanah	Dataset sampel tanah kota Lishu	Memberikan dasar teori yang dapat diandalkan untuk melakukan pemupukan presisi.
12	Sven Krasser, et al, 2007	Mengidentifikasi Spam Gambar berdasarkan Header dan Properti File menggunakan C4.5 Decision Trees dan Support Vector Machine Learning	Data Sampel Internal	SVM Mendapatkan hasil akurasi klasifikasi yang lebih tinggi dibandingkan dengan C4.5
13	Barbara M G Ribeiro, et al, 2013	Klasifikasi Lahan Perkotaan Menggunakan C4.5	Hasil Observasi	Klasifikasi C4.5 mampu melakukan klasifikasi dengan cepat dengan data citra yang banyak dan besar.
14	Annas Gading Pertiwi, et al, 2017	Klasifikasi Provinsi Berdasarkan Tingkat Dropout menggunakan Algoritma C4.5	Dataset : data.go.id	Mengklasifikasikan provinsi berdasarkan angka putus sekolah dilakukan dengan algoritma C4.5 dengan akurasi yang cukup tinggi, yaitu 71,2%.
15	Hehui Qian, et al, 2014	Seleksi Fitur Menggunakan Algoritma C4.5 Untuk Prediksi Harga Listrik	Data Tes	Kinerja metode menggunakan C4.5 lebih baik daripada yang menggunakan ID3.
16	Sucipto, et al, 2016	Metode Klasifikasi Multi-kelas pada Algoritma C4.5 untuk Penyakit Ikan	Data Tes	Digunakan untuk menilai penyakit ikan di mana baik kelas terendah dan akurasi kelas tertinggi

				ditemukan masing-masing sebesar 55,3% dan 88,4%.
17	(Andriyani <i>et al.</i> , 2019)	Memodelkan Klasifikasi rekam medis (Gejala, pemeriksaan fisik dan penyakit penyerta) pada hasil diagnose pasien stroke	Data Rekam medis	Menghasilkan rule base diagnose pasien stroke pada pasien stroke berdasarkan data pada rekam medis yang diperoleh pada pemeriksaan pasien di unit gawat darurat
18	(Pratomo and Andriyani, 2021)	Melakukan klasifikasi citra jamur menggunakan ekstraksi ciri feature Orde 1 dan histogram menggunakan Algoritma C4.5	Citra Jamur hasil Observasi	Nilai akurasi klasifikasi jamur feature Orde 1 sebesar 84%, sedangkan untuk feature Histogram menghasilkan nilai akurasi 92%

Data pada rekam medis dapat digunakan untuk membantu dokter dalam mendiagnosa pasien stroke di unit gawat darurat (Andriyani *et al.*, 2019) data tersebut diolah dengan menggunakan algoritma C4.5, proses pembersihan data dengan bantuan para medis dan dokter spesialis stroke dalam membaca data rekam medis sebelum diolah ketahap klasifikasi.

Pada penelitian (Pratomo and Andriyani, 2021) melakukan klasifikasi citra jamur menggunakan ekstraksi ciri feature Orde 1 dan histogram menggunakan Algoritma C4.5 dengan data primer dan melakukan analisis hasil penelitian dibandingkan dengan penelitian sebelumnya yang dilakukan oleh (Fadlil *et al.*, 2019) dengan Ekstraksi ciri yang sama namun dengan klasifikasi yang berbeda yaitu Jaringan Syaraf Tiruan.