

Komparasi Algoritma Klasifikasi pada Data Mining

Dita Fadma Ristianti¹, Suparman²

^{1,2}Magister Pendidikan Matematika, Universitas Ahmad Dahlan, Indonesia

Abstract. *Big data* adalah salah satu teknik pembelajaran mesin yang paling populer dan menjadi alat penting dalam bisnis. Pengolahan *big data* secara efektif sangat penting bagi perusahaan. Analisis *Big data* dapat menghemat waktu, sehingga menimbulkan istilah data penambangan (*Data Mining*). *Data Mining* adalah teknik analisis data berbasis pada aplikasi statistik; yang bertujuan untuk mengekstrak informasi itu sebelumnya tidak dapat ditentukan, dari jumlah *Big data*. Banyak kasus di dunia nyata, teknik yang terdapat dalam *data mining* yang dapat digunakan untuk mendeteksi masalah yang dihadapi oleh beberapa perusahaan misalnya, clustering, klasifikasi, asosiasi, outlier dan masih banyak lagi. Klasifikasi adalah bentuk dasar dari analisis data. Didalam klasifikasi ada banyak algoritma-algoritma yang dikembangkan misalnya *decision tree*. *Decision tree* dapat mengubah fakta yang sangat besar menjadi pohon keputusan yang mempresentasikan aturan, didalam *decision tree* masih banyak lagi metode-metode yang dikembangkan. Pada jurnal ini akan dibahas komparasi dari pengembangan *decision tree* yaitu *fuzzy decision tree* dan *Rough set decision tree*.

Kata Kunci : *Decision tree*, *fuzzy decision tree*, Klasifikasi, *rough set decision tree*.

1. Pendahuluan

Perkembangan teknologi yang cepat menyebabkan peningkatan global yang signifikan dalam volume data yang memiliki jumlah besar, atau biasa disebut *big data*. *Big data* adalah salah satu teknik pembelajaran mesin yang paling populer. *Big Data* bukan hanya subjek yang menarik bagi para peneliti tetapi juga telah menjadi alat penting dalam bisnis. Pengolahan *big data* secara efektif sangat penting bagi perusahaan [1]. Banyak yang dapat diketahui dari sejumlah informasi yang dapat diperoleh dari *Big big data*. Memperoleh informasi dari *big data* dengan menggunakan metode yang tepat sama dengan mengekstraksi sebanyak mungkin dari tambang yang baru ditemukan. Kebutuhan untuk sampai pada kesimpulan yang akurat secara ilmiah menyoroti perlunya analisis data besar. Analisis data besar dapat mengurangi kehilangan informasi dan menghemat waktu, sehingga menimbulkan istilah data penambangan (*Data Mining*) [2].

Data Mining adalah teknik analisis data berbasis pada aplikasi statistik yang bertujuan untuk mengekstrak informasi itu sebelumnya tidak dapat ditentukan dari jumlah *big data* [3]. *Data Mining* adalah sistem informasi berbasis komputer yang ingin memindai data besar repositori, menghasilkan informasi, dan menemukan pengetahuan [4]. *Data Mining* membantu mengatasi banyak kesulitan dalam mengolah data yang berukuran besar. Peran utama dari data mining adalah menerapkan berbagai prosedur dan algoritma untuk mengambil pola dari data yang sangat besar. Saat ini data dapat diambil dari berbagai jenis volume besar dataset dalam berbagai format seperti file data, video, catatan, teks, gambar, audio, data ilmiah dan jenis baru format data. Data yang dikumpulkan dari berbagai sumber memerlukan analisis data yang tepat untuk proses pengambilan keputusan yang efisien [5].

Banyak kasus di dunia nyata, teknik yang terdapat dalam *data mining* yang dapat digunakan untuk mendeteksi masalah yang dihadapi oleh beberapa perusahaan misalnya, clustering, klasifikasi, asosiasi, outlier dan masih banyak lagi. Klasifikasi adalah bentuk dasar dari analisis data. Klasifikasi juga biasa disebut teknik yang digunakan untuk menentukan anggota kelompok dari data yang telah tersedia. Konsep dasar dari klasifikasi adalah sejumlah data yang mempunyai struktur data yang hampir sama atau serupa akan menghasilkan klasifikasi yang hampir sama atau serupa juga. Klasifikasi adalah metode *data mining* yang banyak diterapkan di berbagai bidang [6].

Seiring dengan perkembangan ilmu pengetahuan dan teknologi informasi, kehadiran cabang ilmu baru di bidang komputer data mining telah menarik banyak perhatian dalam dunia sistem informasi. Didalam klasifikasi ada banyak algoritma-algoritma yang dikembangkan misalnya *decision tree*. *Decision tree* adalah metode metode untuk mendekati fungsi target bernilai diskrit, di mana fungsi yang dipelajari diwakili oleh pohon keputusan. Pohon keputusan mengklasifikasikan instance dengan menyortirnya ke bawah pohon dari akar ke beberapa simpul daun, yang menyediakan klasifikasi instance [7]. Dalam penelitian ini akan dibahas pengembangan dari algoritma *decision tree*, karena dalam beberapa kasus pada data real sering mengandung banyak data dengan kelas tidak seimbang. Masalah data kelas tidak seimbang sering dikarenakan salah satu kelas jumlahnya lebih sedikit dari kelas yang lain didalam dataset. Oleh karena itu diusulkan pengembangan algoritma baru yaitu *Fuzzy decision tree* dan *Rough set*.

Fuzzy decision tree (FDT) adalah menggabungkan pohon keputusan simbolik dengan perkiraan alasan yang ditawarkan oleh representasi fuzzy. Tujuannya adalah untuk mengeksplorasi keunggulan komplementer dari keduanya: popularitas dalam aplikasi untuk belajar dan kelengkapan pengetahuan yang tinggi dari pohon keputusan, kemampuan untuk menangani ketidakaktifan dan informasi yang tidak pasti dari representasi fuzzy [8]. Sedangkan, *Rough Set* (RS) menyederhanakan pencarian atribut dominan dalam sistem informasi yang bertujuan untuk meningkatkan efisiensi, kesederhanaan dan kemampuan generalisasi dari kedua algoritma dasar. *Rough set* berbasis *Decision Tree* (RDT) yang menggabungkan alat *Rough set* dengan kemampuan *decision tree* klasik, diusulkan untuk mengatasi masalah *overhead* komputasi [9].

Maka untuk mengatasi permasalahan tersebut, penulis ingin membandingkan beberapa algoritma dari klasifikasi yaitu *Fuzzy Decision Tree*, dan *Rough set* agar mengetahui algoritma yang lebih akurat dalam menyelesaikan permasalahan dari data kelas tidak seimbang untuk mencegah hasil yang seragam. Penelitian ini disusun sebagai berikut. Pada bagian 2, menjelaskan metode-metode yang digunakan. Pada bagian 3, menjelaskan pembahasan. Pada bagian 4, Kesimpulan dan ucapan terimakasih.

2. Metode

2.1 Pengumpulan literatur

Studi pustaka dilakukan dengan diawali mengumpulkan literatur-literatur akademik baik buku teks maupun jurnal-jurnal internasional yang terbaru. Penelitian ini membahas pokok-pokok teori tentang *big data*, klasifikasi, *decision tree*, *fuzzy decision tree* dan *rough set*. Selain itu, perkembangan terbaru tentang beberapa metode di atas baik tentang kualitas metode atau implementasinya dalam berbagai bidang, di bahas berdasarkan jurnal-jurnal yang terkait

2.2 Pengkajian literatur

Penulis membaca dan memahami teori-teori tentang tentang *big data*, klasifikasi, *decision tree*, *fuzzy decision tree*, *rough set* dan materi lain berdasar literatur-literatur yang telah dikumpulkan. Penulis mencatat hal-hal pokok seperti gagasan utama teori, bagaimana prosedur penerapannya, kelebihan dan kekurangannya serta perkembangan teori tersebut hingga masa sekarang. Kekurangan dan kelebihan metode ini menjadi pusat perhatian pada penelitian ini. Hasil-hasil penelitian dari jurnal-jurnal yang dikumpulkan digunakan sebagai pendukung kelebihan atau solusi untuk menutupi kekurangan yang ada.

2.3 Penyusunan hasil kajian

Hasil kajian disusun sebagai berikut:

- **Pendahuluan**
Meliputi latar belakang masalah pada *big data*, *data mining*, klasifikasi, *decision tree*, *fuzzy decision tree* dan *rough set decision tree*. Hasil-hasil penelitian terbaru diungkapkan sekilas guna mendukung bahasan pokok penelitian. Ide dasar solusi dari masalah diungkapkan pada akhir bagian pendahuluan.
- **Metode**
Meliputi langkah-langkah kajian pustaka.
- **Hasil dan pembahasan**
Menguraikan gagasan utama teori *big data*, *data mining*, klasifikasi, *decision tree*, *fuzzy decision tree* dan *rough set decision tree*. Kelebihan dan kekurangan metode di bahas dan kemudian solusi yang ditawarkan berdasarkan hasil-hasil penelitian yang sudah ada sebelumnya.
- **Kesimpulan**
Berisi kesimpulan dari hasil kajian

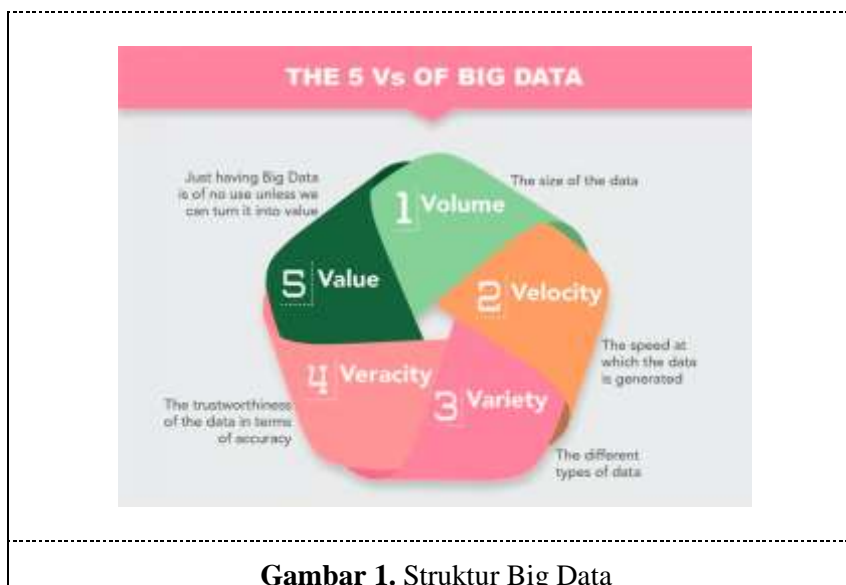
3. Hasil dan Pembahasan

Pada penulisan penelitian ini, penulis menggunakan buku, prosiding, dan jurnal sebagai referensi guna memaparkan model algoritma Fuzzy *Decision tree*, dan *Rough Set*.

3.1. Big Data

Big Data pertama kali diperkenalkan ke dunia komputasi oleh Roger Magoulas untuk mendefinisikan sejumlah besar data yang tidak dapat dikelola dan diproses oleh teknik manajemen data tradisional karena kompleksitas dan ukuran data yang sangat besar. *Big Data* didefinisikan berdasarkan ukurannya, terdiri dari kumpulan set data yang besar, kompleks dan independen, masing-masing dengan potensi untuk berinteraksi. Selain itu, aspek penting dari *Big Data* adalah fakta bahwa *Big Data* tidak dapat ditangani dengan teknik manajemen data standar karena ketidakkonsistenan dan kemungkinan kombinasi yang tidak dapat diprediksi [10].

Big Data dikaitkan dengan kumpulan data besar dan ukurannya di atas fleksibilitas perangkat lunak basis data umum untuk menangkap, menyimpan, menangani, dan mengevaluasi [11, 12]. Analisis *Big Data* sangat penting bagi analis, peneliti dan pelaku bisnis untuk membuat keputusan yang lebih baik yang sebelumnya tidak tercapai. Struktur big data yang berisi lima dimensi yaitu *volume*, *velocity*, *variety*, *veracity*, dan *value* seperti pada **Gambar 1**.



Gambar 1. Struktur Big Data

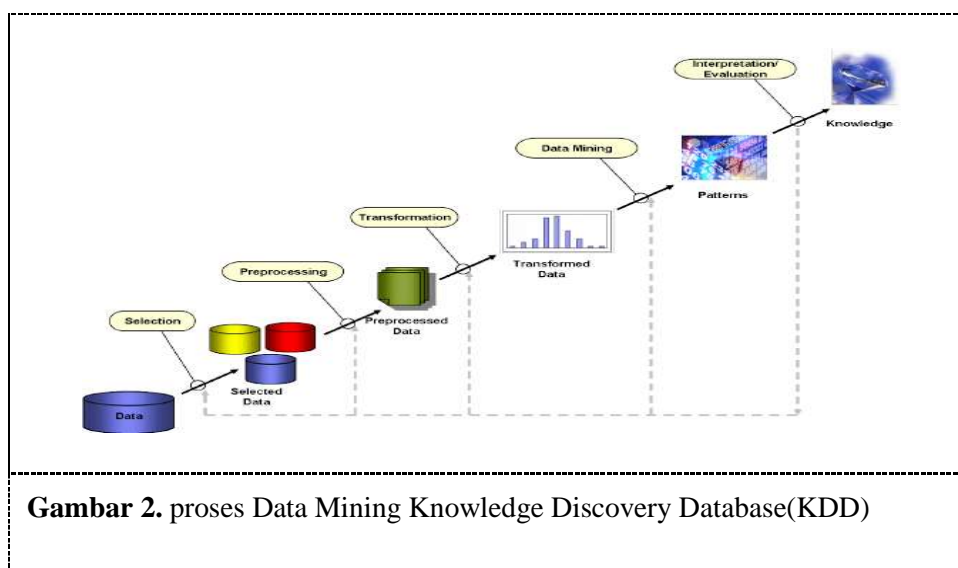
Volume mengacu pada ukuran data yang terutama menunjukkan bagaimana menangani data base skalabilitas besar dan data base dimensi tinggi dan kebutuhan pemrosesan. *Velocity* mendefinisikan kedatangan aliran data yang berkesinambungan dari informasi yang berguna ini diperoleh. Selain itu, big data telah meningkatkan through-put yang ditingkatkan, konektivitas dan kecepatan komputasi perangkat digital yang telah mempercepat pengambilan, proses, dan produksi data. *Veracity* menentukan kualitas informasi dari berbagai tempat. *Variety* menjelaskan cara mengirim berbagai jenis data, misalnya data sumber mencakup tidak hanya data relasional tradisional terstruktur tetapi juga mencakup kuasi-terstruktur, semi-terstruktur dan data tidak terstruktur seperti teks, data sensor, audio, video, grafik dan banyak lagi lebih banyak tipe. *Value* sangat penting untuk mendapatkan ekonomi nilai data berbeda yang bervariasi secara signifikan [11].

Big data pada saat ini menjadi salah satu teknik pembelajaran mesin yang paling populer. Big Data bukan hanya subjek yang menarik bagi para peneliti tetapi juga telah menjadi alat penting dalam bisnis. Pengolahan big data secara efektif sangat penting bagi perusahaan [1]. Analisis Big data dapat menghemat waktu, sehingga menimbulkan istilah data penambangan (Data Mining).

3.2. Data Mining

Suatu proses di mana data mentah sedang dipersiapkan dan disusun sedemikian rupa sehingga informasi yang berharga dapat diekstraksi darinya disebut Analisis data. Proses pengorganisasian dan berpikir tentang data adalah cara untuk melakukannya menerima apa yang data lakukan dan tidak mengandung. Analisis Data adalah proses inspeksi, pembersihan, transformasi, dan pemodelan data. Tujuan dari analisis data adalah untuk menyortir informasi yang bermanfaat, memberikan kesimpulan, dan membantu dalam pengambilan keputusan. Analisis data terdiri dari beberapa langkah dan pendekatan, termasuk beragam teknik di bawah serangkaian nama, dalam domain bisnis, ilmu pengetahuan, dan ilmu sosial yang berbeda [13]. *Data Mining* adalah penemuan informasi yang tidak diketahui dari basis data [14]. *Data Mining* adalah teknik analisis data berbasis pada aplikasi statistik; yang bertujuan untuk mengekstrak informasi itu sebelumnya tidak dapat ditentukan dari jumlah *big data* [3]. *Data Mining* adalah sistem informasi berbasis komputer yang ingin memindai data besar repositori, menghasilkan informasi, dan menemukan pengetahuan [4].

Data Mining membantu mengatasi banyak kesulitan dalam mengolah data yang berukuran besar. Peran utama dari *data mining* adalah menerapkan berbagai prosedur dan algoritma untuk mengambil pola dari data yang sangat besar . Saat ini data dapat diambil dari berbagai jenis volume besar dataset dalam berbagai format seperti file datar, video, catatan, teks, gambar, audio, data ilmiah dan jenis baru format data. Data yang dikumpulkan dari berbagai sumber memerlukan analisis data yang tepat untuk proses pengambilan keputusan yang efisien [5]. Pada proses *Data Mining* yang biasa disebut *Knowledge Discovery Database (KDD)* terdapat beberapa proses seperti pada Gambar 2.



Gambar 2. proses Data Mining Knowledge Discovery Database(KDD)

Pada proses *Knowledge Discovery Database (KDD)* terdapat beberapa fase yaitu sebagai berikut:

- a. *Seleksi Data (Selection)*
Selection (seleksi/pemilihan) data dari kumpulan data operasional perlu dilakukan sebelum melalui tahap penggalian informasi dalam *Knowledge Discovery Database (KDD)* dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional
- b. *Pemilihan Data (Preprocessing/Cleaning)*
Proses *Preprocessing* yaitu membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Dilakukan juga proses enrichment, yaitu proses “memperkaya” data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi eksternal.
- c. *Transformasi (Transformation)*
Pada fase Transformasi yang dilakukan adalah mentransformasi bentuk data yang belum memiliki entitas yang jelas kedalam bentuk data yang valid atau siap untuk dilakukan proses *Data Mining*
- d. *Data Mining*
Pada fase *Data Mining* yang dilakukan adalah menerapkan algoritma atau metode pencarian pengetahuan.
- e. *Interpretasi/Evaluasi (Interpretation/Evaluation)*
Pada fase terakhir ini yang dilakukan adalah proses pembentukan keluaran yang mudah dimengerti yang bersumber pada proses *Data Mining* Pola informasi

Banyak kasus di dunia nyata, teknik yang terdapat dalam *data mining* yang dapat digunakan untuk mendeteksi masalah yang dihadapi oleh beberapa perusahaan misalnya, clustering, klasifikasi, asosiasi, outlier dan masih banyak lagi. Klasifikasi adalah bentuk dasar dari analisis data. Klasifikasi juga biasa disebut teknik yang digunakan untuk menentukan anggota kelompok dari data yang telah tersedia. Konsep dasar dari klasifikasi adalah sejumlah data yang mempunyai struktur data yang hampir sama atau serupa akan menghasilkan klasifikasi yang hampir sama atau serupa juga. Klasifikasi adalah metode *data mining* yang banyak diterapkan di berbagai bidang [6].

Seiring dengan perkembangan ilmu pengetahuan dan teknologi informasi, kehadiran cabang ilmu baru di bidang komputer data mining telah menarik banyak perhatian dalam dunia sistem informasi. Didalam klasifikasi ada banyak algoritma-algoritma yang dikembangkan misalnya *decision tree*.

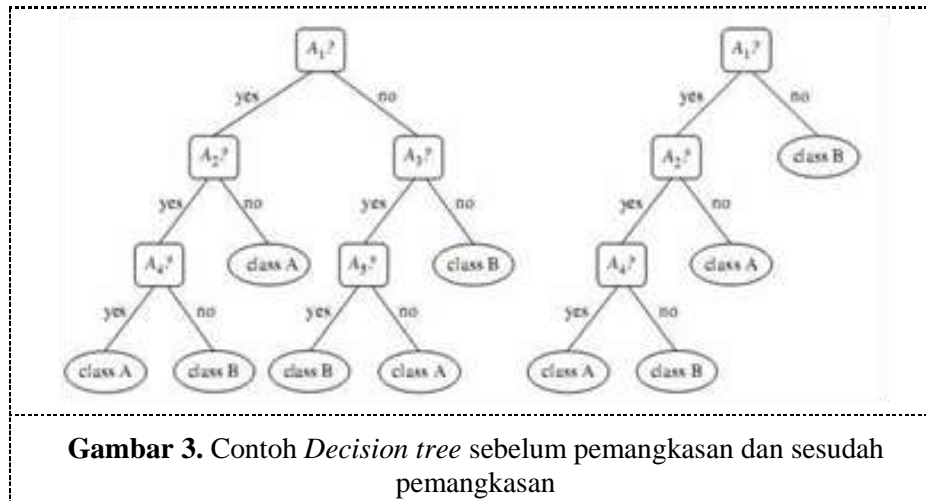
3.3. *Decision Tree*

Decision tree adalah pendekatan yang sangat populer dan praktis untuk klasifikasi. *Decision tree* menggunakan pohon keputusan sebagai model prediktif yang memetakan pengamatan tentang item ke kesimpulan tentang nilai target item. Ini adalah salah satu pendekatan pemodelan prediktif yang digunakan dalam statistik, *big data*, dan *medicine learning*. Konsep yang ditawarkan dalam *decision tree* yaitu merubah data menjadi pohon keputusan aturan-aturan keputusan yang selanjutnya dapat disederhanakan dengan menghilangkan cabang-cabang atau aturan-aturan yang tidak perlu.

Decision tree memiliki struktur pohon seperti bagan alur, di mana setiap simpul internal mewakili tes pada atribut, masing-masing cabang mewakili hasil tes, label kelas diwakili oleh setiap simpul daun (atau simpul terminal). Diberikan tuple X , nilai atribut tuple diuji terhadap *decision tree*. Path dilacak dari root ke node leaf yang menampung prediksi kelas untuk tuple [13].

Pohon ini dibangun pada fase pertama dengan secara rekursif memisahkan set pelatihan berdasarkan kriteria optimal lokal sampai semua atau sebagian besar catatan milik masing-masing partisi bertuliskan label kelas yang sama. Pohon dapat menyesuaikan data. Fase pemangkasan menangani masalah over fitting data di pohon keputusan. Tahap pemangkasan menggeneralisasi pohon dengan menghilangkan noise dan outlier [14].

Keakuratan Klasifikasi meningkat pada fase pemangkasan. Fase pemangkasan hanya mengakses pohon yang sudah dewasa. Fase pertumbuhan membutuhkan beberapa lintasan melewati data pelatihan [14]. Pohon keputusan dapat dibangun relatif cepat dibandingkan dengan metode klasifikasi lainnya. Contoh dari *decision tree* ditunjukkan pada Gambar 3.



Gambar 3. Contoh *Decision tree* sebelum pemangkasan dan sesudah pemangkasan

3.4. Fuzzy Decision Tree

Algoritma *decision tree* adalah salah satu metodologi paling populer untuk metode klasifikasi. Pengetahuan yang dihasilkan, keputusan simbolis pohon bersama dengan mekanisme inferensi sederhana, telah dipuji karena kelengkapannya. *Decision tree* yang paling dapat dipahami telah dirancang untuk simbolik yang sempurna [15]. *Decision tree* bekerja dengan baik di domain rendah, tetapi tidak dapat memodelkan ketidakjelasan. Untuk mengatasi masalah batas tajam, Fuzzy *decision tree* (FDT) telah diusulkan yang menggabungkan teori fuzzy dengan *Decision tree*[16].

Fuzzy *decision tree* (FDT) adalah penggabungan pohon keputusan simbolik dengan perkiraan alasan yang ditawarkan oleh representasi fuzzy. Tujuannya adalah untuk mengeksplorasi keunggulan komplementer dari keduanya popularitas dalam aplikasi dan kelengkapan pengetahuan yang tinggi dari pohon keputusan, kemampuan untuk menangani ketidakaktifan dan informasi yang tidak pasti dari representasi fuzzy [8]. Keuntungan nyata dari pohon keputusan fuzzy adalah bahwa mereka menggunakan rutinitas yang sama dengan pohon keputusan simbolis (tetapi dengan representasi fuzzy). Hal ini memungkinkan pemanfaatan struktur pohon yang dapat dipahami yang sama untuk pemahaman dan verifikasi pengetahuan. Ini juga memungkinkan pemrosesan yang lebih kuat dengan terus menerus [15].

3.5. Rough Set

Rough set decision tree adalah penggabungan keunggulan algoritma induksi *Rough set* dan *decision tree*. *Rough set* berbasis *Decision Tree* (RDT) yang menggabungkan alat *Rough set* dengan kemampuan *decision tree* klasik, diusulkan untuk mengatasi masalah *overhead* komputasi. Teori *Rough Set* (RS) menyederhanakan pencarian atribut dominan dalam sistem informasi yang bertujuan untuk meningkatkan efisiensi, kesederhanaan dan kemampuan generalisasi dari kedua algoritma dasar [9].

Rough set [17] mampu mengatasi kerugian dari pengklasifikasi. Banyaknya fitur tambahan algoritma klasifikasi, diekstraksi untuk penemuan aturan membuat pengklasifikasi bekerja lebih banyak rumit yang terkadang tidak dapat ditangani. Akibatnya mereka mencoba memilih fitur mengurangi waktu perhitungan. Tetapi metode ini juga menurunkan kinerja. Karena fitur pilihan, tingkat akurasi berkurang karena setiap fitur penting untuk klasifikasi yang halus. *Rough set decision tree* adalah model baru untuk mengatasi masalah ini. Teori *rough set* adalah komponen dari solusi hybrid dari metode

pembelajaran mesin dan data mining. Itu hubungan indiscernibility yang menyebabkan aturan keputusan minimal dari contoh pelatihan adalah notasi penting dalam model *rough set decision tree*. Identifikasi *Rough set* fitur minimal dilakukan dengan aturan IF– THEN Pendekatan baru untuk menghitung atribut dominan (perkiraan mengurangi) dan menggunakannya untuk membangun pohon keputusan yang disebut *Rough set decision tree* (RDT). Aturan keputusan yang dihasilkan dari RDT disimpan dalam RDT Rules Database (RDTRD) [9].

Dari hasil literatur yang didapat, komparasi dari kedua metode tersebut akan di jelaskan pada Tabel.1:

Table 1. Tinjauan Study Terdahulu.

Penulis/Judul/Tahun Terbit	Deskripsi	Algoritma
<i>R. Revathy & R. Lawrance / Rough Set Based Decision Tree Model for Classification / 2003</i>	Model Rough set berbasis Decision Tree (RDT) menggabungkan alat RS dengan kemampuan DT klasik, diusulkan untuk mengatasi masalah overhead komputasi. Eksperimen membandingkan kinerja RDT dengan pendekatan RS dan algoritma ID3. Kinerja pendekatan RDT dibandingkan RS diamati lebih baik dalam akurasi dan kompleksitas aturan sementara RDT dan ID3 sebanding.	Data Mining, Data preprocessing, Relief filter, Random Forest, C4.5 Classification
<i>Sonajharia Minz & Rajni Jain/ Rough Set based Decision Tree Model for Classification / 2015</i>	<i>Rough set decision tree</i> adalah penggabungan keunggulan algoritma induksi <i>Rough set</i> dan <i>decision tree</i> . RDT dapat berfungsi sebagai model untuk klasifikasi karena menghasilkan aturan dan penghapusan yang lebih sederhana atribut yang tidak relevan pada tahap sebelum induksi pohon. Ini memfasilitasi lebih sedikit memori persyaratan untuk langkah-langkah selanjutnya saat menjalankan model dan untuk mengklasifikasikan data uji serta contoh aktual. Untuk set data nyata, kadang-kadang jumlah pengurangan (mungkin ratusan) ada dan kadang-kadang tidak ada reduksi. Ini menyediakan potensial untuk penyempurnaan model RDT lebih lanjut. Ketersediaan banyak tawaran pengurangan ruang lingkup untuk menghasilkan pohon menghindari evaluasi atribut yang sulit atau mustahil untuk diukur. Ini juga menawarkan opsi menggunakan pohon keputusan berbiaya rendah	Rough set, supervised learning, decision tree, feature selection, classification, data mining.
<i>Yi lai Chen, Tao Wang, Ben-sheng Wang & Zhou-jun Li / A Survey of Fuzzy Decision Tree Classifier / 2019</i>	Survei metode Pohon Keputusan Fuzzy (FDT) penunjukan dan berbagai masalah yang ada. keuntungan dari pengklasifikasi FDT daripada pengklasifikasi pohon keputusan tradisional, membahas FDT termasuk kriteria pemilihan atribut, inferensi untuk penetapan keputusan dan kriteria berhenti	Fuzzy decision tree, Classifier, Attribute selection, Decision assignment . Stopping criteria
<i>Liam Evans, Niels Lohse, & Mark</i>	Algoritma fuzzy-decision-tree diterapkan untuk memberikan pendekatan yang lebih objektif	Experience-based decision support

<i>Summers / A fuzzy-decision-tree approach for manufacturing technology selection exploiting experience-based information / 2013</i>	mengingat bukti kasus implementasi teknologi manufaktur sebelumnya. pendekatan berbasis pohon keputusan fuzzy (FDT) untuk membenaran teknologi manufaktur menggunakan pendekatan berbasis pengalaman. Model ini mempertimbangkan berbagai elemen kuantitatif dan kualitatif yang mewakili data pengambilan keputusan historis dari kasus-kasus sebelumnya. Metodologi telah dikembangkan untuk mempertimbangkan pengalaman dan intuisi yang terbentuk melalui sejarah kasus. Dibandingkan dengan model subjektif seperti AHP yang membutuhkan penilaian yang tepat, pendekatan ini didasarkan pada informasi kinerja teknologi yang sebelumnya dianggap lebih obyektif format.	Manufacturing technology selection Fuzzy decision tree Data mining
---	--	--

4. Kesimpulan

Decision tree adalah pendekatan yang sangat populer dan praktis untuk klasifikasi. Konsep yang ditawarkan dalam *decision tree* yaitu merubah data menjadi pohon keputusan aturan-aturan keputusanyang selanjutnya dapat disederhanakan dengan menghilangkan cabang-cabang atau aturan-aturan yang tidak perlu. Pengembangan algoritma *decision tree* baru yaitu *Fuzzy decision tree* dan *Rough set decision tree*. *Fuzzy decision tree* (FDT) adalah menggabungkan pohon keputusan simbolik dengan perkiraan alasan yang ditawarkan oleh representasi fuzzy. Tujuannya adalah untuk mengeksplorasi keunggulan komplementer dari keduanya: popularitas dalam aplikasi untuk belajar dan kelengkapan pengetahuan yang tinggi dari pohon keputusan, kemampuan untuk menangani ketidakaktifan dan informasi yang tidak pasti dari representasi fuzzy. Sedangkan, *Rough set decision tree* adalah menggabungkan keunggulan algoritma induksi *Rough set* dan *decision tree*. *Rough set* berbasis *Decision Tree* (RDT) yang menggabungkan alat *Rough set* dengan kemampuan *decision tree* klasik, diusulkan untuk mengatasi masalah *overhead* komputasi. Teori *Rough Set* (RS) menyederhanakan pencarian atribut dominan dalam sistem informasi yang bertujuan untuk meningkatkan efisiensi, kesederhanaan dan kemampuan generalisasi dari kedua algoritma dasar. Kedua algoritma pengembangan dari *decision tree* sangat membantu dalam menyelesaikan masalah dari algoritma klasifikasi khususnya pada *decision tree*, yang membedakannya yaitu data yang akan diolah yang menentukan algoritma yang cocok digunakan.

5. Ucapan Terimakasih

Terimakasih diucapkan kepada mamah dan ayah yang telah, sedang dan selalu mendoakan serta mendukung penulis untuk berprestasi. Terimakasih juga diucapkan kepada pimpinan Universitas Ahmad Dahlan dan Fakultas Pascasarjana UAD yang selalu memberikan dukungan kepada kami untuk selalu mengembangkan diri. Kepada ketua prodi Magister Pendidikan Matematika yang telah membimbing kami untuk menjadi mahasiswa S2 yang berkualitas dan berkaliber internasional dengan menghasilkan paper yang berkualitas. Kepada bapak ibu dosen prodi Magister Pendidikan Matematika yang selalu membagi ilmu kepada mahasiswa dengan tulus ikhlas. Selain itu terimakasih kepada Magister Pendidikan Matematika Universitas Ahmad Dahlan sebagai penyelenggara STEEEM 2019.

References

- [1] J. Han and M.Kamber 2011 Elevier **8**
- [2] Begum Çıgsar dan Deniz Unual 2019 *Hindawi Scientific Programing* **2019**.
- [3] Natra 2016 *International Journal of Computer Science and Information Security (IJCSIS)* **3**.
- [4] Tariq O. Fadl Elsid dan Mirghani. A. Elthahir 2014 *International Journal of Computer Science*

T. T 16

- [5] K. Sumathi, S. Kannan, dan K. Nagarajan 2016 *International Journal Computer Applications* **141**
- [6] Z. Sun, Q. Song, X. Zhu, H. Sun, B. Xu, and Y. 2015. *Pattern Recognit.* Volume B **4** : 1623–1637.
- [7] Xiaojing Yuan, Xiaohui Yuan, Fan Yang, Jing Peng dan Bill P. Buckles 2003 *FLAIRS* **2003**
- [8] Yi-lai Chen, Tao Wang, Ben-sheng Wang dan Zhou-jun Li. 2009 *Fuzzy Information and Engineering* 2009
- [9] Sonajharia Minz dan Rajni Jain 2003. *IEEE* **15**
- [10] Elena Geanina, Florina Camelia, Anca, dan Manole 2012 *Databese System Jurnal* III
- [11] S.Vijayarani dan S.Sharmila 2016 *Informatics Engineering, an International Journal* 4
- [12] Neelam Singh, Neha Garg, dan Varsha Mittal 2013 *International Journal of Scientific & Engineering Research* 4
- [13] Thankachan S. dan Suchithra 2017 *Journal of Computer Science & Mobile Computing* 6 160-168
- [14] Ullah I. 2010 *International Journal of Computer Science & Information Technology (IJCSIT)* 2
- [15] Himani Sharma dan Sunil Kumar 2013 *International Journal of Science and Research (IJSR)* 2013.
- [16] D.Lanavanya dan K. Usha Rani 2011 *International Journal of Computer Applications* 26.
- [17] Yi-lai Chen, Tao Wang, Ben-sheng Wang, dan Zhou-jun Li 2008 *Fuzzy Information and Engineering Branch of the Operations Research Society of China* 2
- [18] Vinita A. Gupta dan Sunita Sono 2014 *SSRG International Journal of Computer Science and Engineering* 1
- [19] Zdzislaw Pawlak, Jerzy Grzymala-Busse, Roman Slowinski, dan Wojciech Ziarko 1995 *Communications of the ACM* **38**