

## **FILOSOFI, TEORI DAN IMPLEMENTASI N-SOFT SETS PADA DATA MINING**

### **PHILOSOPHY, THEORY, AND IMPLEMENTATION OF N-SOFT SETS IN DATA MINING**

Fatia Fatimah\*, Selly Anastassia Amellia Kharis

Program Studi Matematika, Fakultas Sains dan Teknologi,  
Universitas Terbuka

\*fatia@ecampus.ut.ac.id

#### **ABSTRAK**

*Pengambilan keputusan merupakan proses yang penting dalam kehidupan individu atau kelompok. N-soft set merupakan inovasi dalam teori pengambilan keputusan khususnya dalam mengatasi masalah ketidakpastian dengan merepresentasikan data dalam berbagai bentuk. Beberapa penelitian telah menggabungkan N-soft set dengan teori-teori lain namun belum ada penelitian yang menggabungkan N-soft set dengan data mining. Oleh karena itu, Penelitian ini membahas filosofi kolaborasi antara N-soft set dan data mining serta pendekatan yang diperlukan untuk mengimplementasikan N-soft set dalam metode data mining. N-soft set dapat diaplikasikan untuk memodelkan ketidakpastian dan ambiguitas dalam data dan data mining dapat mengidentifikasi pola dan tren. Metode pencarian pola dalam data mining seperti Association Rule Mining, Clustering, dan Classification juga dapat menggunakan representasi N-soft set. Hal ini memungkinkan identifikasi pola yang relevan dan hubungan yang ada dalam data dengan mempertimbangkan tingkat keanggotaan objek dalam kelas-kelas yang relevan. Penggabungan N-soft set dengan data mining dapat menghasilkan pemodelan yang lebih representatif*

sesuai kebutuhan pengguna dengan mempertimbangkan lebih banyak faktor dalam pengambilan keputusan. Hal ini disebabkan pola data yang digali telah memiliki tanda atau pengenal sehingga memudahkan untuk mengidentifikasi faktor atau fitur. Penggunaan *N*-soft set pada data mining diharapkan dapat memberikan wawasan baru dalam pengambilan berbagai jenis keputusan dengan mempertimbangkan ketidakpastian, ambiguitas, dan kompleksitas data.

**Kata Kunci:** *N*-soft sets, data mining, pengambilan keputusan

#### ABSTRACT

*Decision making is an essential process in individual or group life. N-soft set is an innovation in decision-making theory, particularly in addressing uncertainty issues by representing data in various forms. Several studies have combined N-soft set with other theories, but no research that combines N-soft set with data mining. Therefore, this study discusses the philosophy of collaboration between the N-soft set and data mining and the approaches needed to implement the N-Soft set into data mining methods. N-soft set can be applied to model uncertainty and ambiguity in data, while data mining can identify patterns and trends. Pattern discovery methods in data mining, such as Association Rule mining, clustering, and classification, can also utilize N-soft set representations. N-soft set data mining collaboration enables the identification of relevant patterns and relationships in data by considering the membership levels of objects in relevant classes. Combining N-soft sets with data mining can produce a more representative model according to user needs by considering more factors in decision-making. Because the pattern of data extracted already has a sign or identifier, making it easier to identify factors or features. Using the N-soft set for data mining is expected to provide new insights into making various decisions by considering the data's uncertainty, ambiguity, and complexity.*

**Keywords:** *N*-soft sets, data mining, decision making

## PENDAHULUAN

Setiap individu ataupun kelompok selalu berhadapan dalam pengambilan keputusan. Keputusan bersifat skala prioritas karena penting dan mendesak, atau analisis keputusan berdampak jangka panjang secara masif, atau keputusan sederhana dalam kehidupan sehari-hari yang perlu segera dan tepat dilakukan. Pengambilan keputusan meski sering dilakukan namun hakikatnya merupakan mata rantai proses meliputi pengenalan terhadap masalah yang sedang dihadapi, mempertimbangkan alternatif yang tersedia, memiliki target atau tujuan yang jelas, mempunyai prediksi atau asumsi hasil yang akan diperoleh, melakukan berbagai kombinasi antara alternatif dengan hasil, dan menetapkan model yang lebih cocok dengan situasi untuk memperoleh keputusan (Render, Stair Jr, Hanna, & Hale, 2018).

Keberlanjutan (*sustainability*) merupakan pendekatan berbasis sistem pada bidang ekonomi, lingkungan, dan sosial untuk kehidupan lebih baik saat ini dan untuk generasi mendatang. Salah satu atribut yang perlu dikembangkan oleh organisasi adalah kemampuan menciptakan inovasi menuju keberlanjutan (de Almeida, Gohr, Morioka, & da Nobrega, 2021). Inovasi tanpa henti dengan tujuan memudahkan pekerjaan, efektif dan efisiensi untuk kehidupan yang lebih baik pada saat sekarang dan untuk generasi puluhan bahkan ratusan tahun yang akan datang. Inovasi berkelanjutan diharapkan dilakukan di semua bidang ilmu sesuai kapasitas masing-masing.

Teori *N-soft set* merupakan salah satu bentuk inovasi pada teori pengambilan keputusan khususnya masalah *uncertainty* yang mampu merepresentasikan berbagai data baik berupa angka, kata, kalimat, simbol, peringkat dan sebagainya (Fatimah, Rosadi, Hakim, & Alcantud, 2018). Karena *N-soft set* merupakan teori baru maka membuka kesempatan untuk peneliti lain dalam mengembangkan inovasi dan pengembangan secara teoritis maupun praktis. Kolaborasi *N-soft set* pada bidang *data mining*

merupakan tantangan yang perlu dilakukan mengingat fungsi masing-masing ilmu tersebut memfasilitasi fleksibilitas berbagai tipe keputusan.

Kombinasi *N-soft sets* dengan teori-teori lain cukup intensif dilakukan oleh peneliti lain dari berbagai negara. Berikut beberapa penelitian terkait kolaborasi *fuzzy sets* dengan *N-soft sets* (Akram, Adeel, & Alcantud, 2018; Akram, Adeel, & Alcantud, 2019b; Adeel, Akram, Yaqoob, & Chammam, 2020; Zhang, Jia-Hua, & Yan, 2020; Akram, Amjad, & Davvaz, 2021; Fatimah, & Alcantud, 2021; Kamaci & Saqlain, 2021; Mahmood, Ur Rehman, & Ali, 2021; Rehman & Mahmood, 2021), *hesitant fuzzy* dengan *N-soft sets* (Akram, Adeel, & Alcantud, 2019a; Akram & Adeel, 2019; Fatimah, 2020), *vague sets* dengan *N-soft sets* (Chen, Liu, Chen, & Zhang, 2020; Liu, Chen, Chen, & Zhang, 2020), *neutrosophic* dengan *N-soft sets* (Riaz, Naeem, Zareef, & Afzal, 2020; Akram, Shabir, & Ashraf, 2021; Ashraf & Butt, 2021), struktur aljabar pada *N-soft sets* (Riaz, Çağman, Zareef, & Aslam, 2019; Kamaci, 2020; Kamaci & Petchimuthu, 2020; Riaz, Razzaq, Aslam, & Pamucar, 2021), dan kolaborasi *rough sets* dengan *N-soft sets* (Alcantud, Feng, & Yager, 2019; Zhang, Li, & An, 2021). Selanjutnya, *N-soft sets* untuk pengambilan keputusan diantaranya yaitu pendekatan *N-soft sets* jika data tidak lengkap (Fatimah, 2018), *big data* (Fatimah, 2021a, 2021b), reduksi parameter (Akram, Ali, Alcantud, & Fatimah, 2020), contoh kasus pariwisata (Fatimah, & Andriyansah, 2020), dan contoh pada pemilihan umum (Fatimah, Rosadi, Hakim, Alcantud, 2017).

Namun belum ada penelitian yang mengkombinasikan *N-soft sets* dengan *data mining*. Oleh karena itu, artikel ini membahas tentang filosofi yang mendasari prinsip kolaborasi *N-soft sets* dengan *data mining*. Selanjutnya, kami juga mengkaji teori-teori terkait dan pendekatan yang diperlukan untuk mengimplementasikan *N-soft sets* pada metode *data mining* sebagai salah satu tren dalam sains dan teknologi untuk kehidupan berkelanjutan.

## PEMBAHASAN

Filosofi penemuan *N-soft sets* didasarkan pada perlunya untuk merepresentasikan penilaian sejumlah berhingga peringkat dengan pendekatan parameterisasi sesuai kebutuhan atau preferensi setiap manusia sehingga bentuk penilaian tidak selalu harus biner, atau dalam rentang antara 0 dan 1, namun dapat dalam bentuk berhingga angka. Bahkan evaluasi yang diberikan tidak terbatas dalam bentuk angka saja tetapi juga dapat dalam bentuk simbol maupun kalimat.

*N-soft sets* didefinisikan sebagai berikut.

**Definisi 1.** Misalkan  $U$  menyatakan semesta objek dan  $Q$  himpunan atribut,  $A \subseteq Q$ . Misalkan  $G = \{0, 1, \dots, N-1\}$  merupakan himpunan peringkat terurut dengan  $N \in \{2, 3, \dots\}$ . *N-soft set* atas  $U$  dinotasikan dengan  $(F, A, N)$ , didefinisikan sebagai hasil pemetaan  $F: A \rightarrow 2^{U \times G}$  untuk setiap  $q \in A$  terdapat secara tunggal  $(u, g_q) \in U \times G$  sedemikian sehingga  $(u, g_q) \in F(q)$ ,  $u \in U, g_q \in G$  (Fatimah dkk., 2018).

Misalkan  $U = \{u_j \mid j=1, 2, \dots, m\}$  dan  $A = \{q_j \mid j=1, 2, \dots, n\}$  merupakan himpunan berhingga maka tabular *N-soft set* dibuat seperti terlihat pada Tabel 1.

**Tabel 1.** *N-Soft Set*

$(F, A, N)$	$q_1$	$q_2$	...	$q_n$
$u_1$	$r_{11}$	$r_{12}$	...	$r_{1n}$
$u_2$	$r_{21}$	$r_{22}$	...	$r_{2n}$
...	...	...	...	...
$u_m$	$r_{m1}$	$r_{m2}$	...	$r_{mn}$

**Sumber:** Fatimah dkk., (2018)

### 1. Pengambilan Keputusan Menggunakan *N-Soft Set*

Pada sesi ini akan dibahas pengambilan keputusan menggunakan *N-soft sets*. Algoritma keputusan *N-soft sets* memfasilitasi setiap kebutuhan individu termasuk organisasi atau komunitas yang tidak dapat diselesaikan melalui konsep himpunan Cantor dan juga tidak membatasi pengambilan keputusan pada interval  $[0,1]$  seperti pada *fuzzy sets*.

Pengambilan keputusan *N-soft sets* diperkenalkan pertama kali oleh Fatimah et al. (2018) untuk menangani masalah-masalah yang sering ditemui dalam kehidupan sehari-hari. Pertama, permasalahan yang melibatkan pemilihan objek dimana setiap kriteria atau atribut sama pentingnya maka algoritma yang digunakan adalah *Extended Choice Values* (ECV). Kedua, jika setiap kriteria atau atribut memiliki bobot kepentingan tertentu bagi pengambil keputusan maka algoritma pengambilan keputusan *N-soft sets* menjadi *Extended Weight Choice Values* (EWCV). Dengan kata lain, setiap nilai prioritas atau kepentingan kriteria menjadi pertimbangan dalam menentukan pilihan. Ketiga, masalah keputusan terkait kriteria atau atribut yang memiliki kedudukan sama namun nilai untuk masing-masing objek berdasarkan kriteria diberi ambang batas (*threshold*). *Threshold* dapat merupakan batas minimum atau batas maksimum sesuai kebutuhan pengambil keputusan atau penentu kebijakan. Untuk kasus tersebut, algoritma *N-soft sets* yang digunakan adalah *T-Extended Choice Values* (T-ECV). Sebagai contoh, dalam suatu pemilihan dosen terbaik di fakultas, tim penilai hanya akan mempertimbangkan kandidat dosen yang mendapatkan peringkat lebih besar atau sama dengan empat dalam skala peringkat dari nol sampai lima untuk kriteria yang telah ditetapkan bersama.

Pada perkembangannya, model pengambilan keputusan *N-soft sets* mengalami penyesuaian dan inovasi sesuai kebutuhan praktisi dan keilmuan. Dalam pengumpulan data primer ataupun mengambil data sekunder dari berbagai sumber informasi terkadang ditemui bahwa data yang tersedia tidak lengkap. Solusi yang dapat digunakan dengan pendekatan *N-soft sets* yaitu

algoritma *data filling approach for incomplete N-soft sets* (Fatimah, 2018). Sehingga data yang kosong pada item kriteria tertentu tidak harus dibuang. Untuk atribut yang mendapatkan nilai negatif dan positif secara bersamaan maka pengambilan keputusan ECV dan EWCV mengalami penyesuaian untuk menentukan urutan prioritas objek yang dievaluasi (Kamacı & Petchimuthu, 2020).

Sementara itu, untuk atribut atau parameter yang tidak terlalu penting atau bukan menjadi prioritas utama pertimbangan keputusan maka dapat menggunakan revisi dari dua algoritma ECV dan EWCV sebelumnya (Akram et al., 2020). Konsep reduksi parameter berdasarkan *rough sets* tidak tepat digunakan pada konteks *N-soft sets* karena dapat mengurangi keunikan himpunan atribut atau parameter yang merupakan ciri khas dari teori *N-soft sets* yaitu kemampuan merepresentasikan semua keunikan atribut. Disisi lain, penggunaan reduksi melalui pendekatan *rough sets* berdampak pada hasil rekomendasi alternatif optimal akan bervariasi sesuai dengan pilihan reduksi. Reduksi parameter juga merupakan solusi untuk data yang besar ataupun sangat bervariasi seperti ditemui pada bidang *data mining*.

## 2. Pengambilan Keputusan Menggunakan *Data Mining*

Pada era digital saat ini, data yang tersebar semakin banyak dan beragam. Organisasi di berbagai sektor menghadapi tantangan dalam mengelola volume data yang semakin besar dan kompleks. Data yang besar (*big data*) telah menjadi tren baru yang memiliki potensi besar untuk digali. Penggalian data (*data mining*) adalah proses ekstraksi informasi yang berharga dan bermanfaat dari berbagai *database* yang terkait dengan menggunakan teknik statistik, matematika, *artificial intelligent* dan *machine learning* (Surbakti, 2021). Tujuan utama dari *data mining* adalah menemukan hubungan atau pola di dalam data yang mungkin memberikan indikator yang berharga (Syahril, Erwansyah, & Yetri, 2020). *Data mining* melibatkan penggunaan berbagai teknik dan algoritma komputasional untuk menganalisis data dan mengidentifikasi pola

yang signifikan. Pola-pola tersebut digunakan untuk membaca data lainnya yang akan menghasilkan kesimpulan atau prediksi baru berdasarkan data historis.

*Data mining* dapat diimplementasikan untuk menggali wawasan yang berharga dan mendukung pengambilan keputusan yang lebih baik dalam berbagai bidang, termasuk bisnis, keuangan, pemasaran, kesehatan, ilmu pengetahuan, dan sebagainya. Keputusan-keputusan yang ada dapat didasarkan pada suatu hal yang ilmiah dan berdasarkan pola di masa lalu. *Data mining* dapat membantu dalam berbagai sektor mulai dari memprediksi kelulusan siswa (Kharis dkk, 2023), meramalkan harga saham (Zili, Hendri, & Kharis., 2022), klasifikasi kanker paru-paru (Bustamam dkk, 2021), klasifikasi kanker otak (Kharis, Hadi, & Hasanah, 2019), dan berbagai bidang lainnya. *Data mining* sering dilakukan dengan menggunakan perangkat lunak dan alat yang spesifik untuk memproses dan menganalisis data yang kompleks. Langkah-langkah umum dalam proses *data mining* terdiri dari beberapa kegiatan, mulai dari pengumpulan data, *preprocessing data*, pemilihan fitur yang relevan, pengembangan model, evaluasi model, dan interpretasi hasil.

*Data mining* dalam pengambilan keputusan memiliki berbagai kemampuan. Salah satu kemampuannya adalah dapat mengidentifikasi pola (Indriyawati & Winarti, 2021). Banyak pola yang ada di kehidupan sehari-hari namun tidak disadari secara langsung oleh manusia. *Data mining* memiliki berbagai teknik untuk melakukan representasi data yang jika dipilih dan digunakan dengan tepat maka dapat menguntungkan organisasi. Penggalian nilai dan informasi ini dapat mendukung pengambilan keputusan yang lebih efektif. Pengambilan keputusan merupakan proses penting dan menentukan untuk keputusan selanjutnya dalam manajemen organisasi. Keputusan yang tepat dapat memberikan keuntungan kompetitif, meningkatkan efisiensi, dan mengoptimalkan hasil. Keputusan yang ada tentu semakin meningkat permasalahannya dan kompleksitasnya seiring berkembangnya organisasi. Keputusan yang dahulu dapat diputuskan tanpa metode ilmiah semakin sulit

untuk dilakukan dan resiko yang ditimbulkan semakin sulit dicegah. Oleh karena itu, *data mining* menjadi alat yang penting dalam proses pengambilan keputusan karena dapat merepresentasikan data dengan baik.

*Data mining* memiliki kemampuan untuk melakukan prediksi dan peramalan yang berguna dalam pengambilan keputusan (Haba, 2021). Hal ini yang memungkinkan organisasi untuk membuat keputusan lebih tepat dalam merencanakan strategi jangka panjang dan mengantisipasi perubahan pasar. Organisasi dapat mengantisipasi perubahan dan mengambil langkah-langkah yang tepat untuk menghadapinya dengan memahami tren pasar, perilaku pelanggan, dan faktor-faktor lain yang mempengaruhi bisnis. Misalnya, organisasi dapat memprediksi penjualan di tahun berikutnya, mengidentifikasi warna kesukaan pelanggan, preferensi produk, atau hal yang sedang menjadi perbincangan yang dapat digunakan untuk mengembangkan strategi pemasaran yang efektif. Peramalan dan prediksi yang dilakukan dengan *data mining* dapat mengoptimalkan rantai pasokan, manajemen persediaan, dan keputusan produksi yang dilakukan oleh organisasi. Dengan demikian, *data mining* memberikan wawasan yang berharga bagi organisasi untuk membuat keputusan yang lebih tepat, merencanakan strategi yang efektif, dan beradaptasi dengan perubahan yang terjadi di pasar.

*Data mining* dalam pengambilan keputusan memiliki peran dalam segmentasi pelanggan. Segmentasi pasar dilakukan *data mining* dengan mengidentifikasi kelompok pelanggan yang memiliki karakteristik serupa dan mengembangkan strategi pemasaran yang lebih efektif. Organisasi dapat mengidentifikasi pola, preferensi, dan perilaku yang berbeda di antara kelompok pelanggan sehingga dapat memetakan segmentasi pelanggan. Sebagai contoh, organisasi dapat mengidentifikasi kelompok pelanggan yang memiliki selera dan preferensi produk yang serupa, tingkat loyalitas yang tinggi, atau kecenderungan untuk merespon kampanye pemasaran tertentu. Segmentasi pelanggan yang telah diidentifikasi organisasi melalui *data mining* dapat dikembangkan

menjadi strategi pemasaran yang lebih efektif karena sudah jelas hal yang dituju oleh organisasi. Organisasi dapat membuat promosi atau penawaran yang disesuaikan dengan segmentasi pelanggan. Selain itu, segmentasi pelanggan dapat membantu organisasi dalam menetapkan harga, distribusi, dan strategi penjualan. *Data mining* membantu dalam pembuatan keputusan khususnya dalam segmentasi pelanggan yang memungkinkan organisasi untuk mengenali karakteristik pelanggannya dan mengelompokkannya. Akhirnya, organisasi memahami hal yang menjadi preferensi dan perilaku pelanggannya sehingga dapat memberikan pelayanan yang sesuai.

Salah satu manfaat *data mining* dalam pengambilan keputusan adalah organisasi dapat mencegah terjadinya penipuan dengan melakukan deteksi dini menggunakan *data mining* (Wali dkk, 2023). Melalui penerapan teknik dan algoritma deteksi penipuan yang canggih, *data mining* memungkinkan organisasi untuk secara proaktif mengurangi risiko kerugian dan melindungi diri dari aktivitas ilegal yang merugikan. *Data mining* dalam konteks ini dapat menganalisis pola-pola yang tersembunyi dalam data transaksi, mencari tanda-tanda yang mencurigakan seperti anomali, pola perilaku yang tidak biasa, atau indikator lain yang menunjukkan adanya penipuan. Misalnya, *data mining* dapat mengidentifikasi pola transaksi yang tidak wajar, seperti jumlah transaksi yang tinggi dalam waktu singkat, transaksi dengan jumlah yang tidak proporsional, atau pola pembelian yang tidak konsisten dengan perilaku pelanggan sebelumnya.

Berdasarkan identifikasi pola transaksi yang tidak wajar, organisasi dapat proaktif untuk melakukan tindakan yang akan merugikan organisasi atau pelanggan organisasi. Misalnya, organisasi dapat mengirimkan peringatan kepada pelanggan mereka terkait dengan adanya transaksi yang mencurigakan, membatasi akses, hingga memblokir akun yang terlibat. *Data mining* dapat mengungkap pola-pola penipuan yang sulit terdeteksi secara manual dengan menganalisis data transaksi secara menyeluruh. Hal ini memungkinkan organisasi dapat

mengambil tindakan cepat dan tepat guna untuk melindungi diri dari berbagai jenis kerugian yang akan terjadi dan menghancurkan reputasi organisasi yang disebabkan oleh penipuan. *Data mining* membuat organisasi lebih responsif dalam mengenali ancaman penipuan dan dapat membuat keputusan untuk meminimalkan dampak negatif yang mungkin ditimbulkan dan mencegah aktivitas ilegal yang terjadi dalam organisasi ataupun kepada pelanggan organisasi.

*Data mining* mempunyai peran penting dalam mengungkap faktor atau fitur penting yang berpengaruh terhadap suatu variabel terikat. Melalui seleksi fitur dapat diketahui fitur-fitur yang berpengaruh atau tidak terhadap suatu variabel terikat. Jika tidak berpengaruh maka dalam prosesnya nanti fitur tersebut dapat dieliminasi sehingga tidak mengganggu akurasi dari hasil. Misalnya, melalui seleksi fitur, organisasi dapat mengidentifikasi variabel-variabel seperti gen-gen yang berpengaruh dalam penyakit kanker paru-paru (Rustam & Kharis, 2020). Setelah faktor-faktor yang relevan diidentifikasi, organisasi dapat menggunakan hasil analisis *data mining* untuk melakukan pengoptimalan proses. Dengan memahami hubungan antara faktor-faktor ini, organisasi dapat mengambil langkah-langkah yang diperlukan untuk meningkatkan kinerja dan efisiensi proses mereka. Misalnya, mereka dapat mengidentifikasi mutasi gen mana yang menyebabkan timbulnya suatu penyakit, mengidentifikasi faktor-faktor penyebab permasalahan atau pemborosan atau mengidentifikasi praktik terbaik yang dapat diterapkan untuk meningkatkan kinerja.

### **3. Pengambilan Keputusan Menggunakan *N-Soft Set* dan *Data Mining***

Pengambilan keputusan yang efektif merupakan hal yang penting dalam menghadapi permasalahan yang semakin kompleks dan dinamis. Sebuah organisasi membutuhkan pendekatan yang kuat untuk mengatasi ketidakpastian, ambiguitas, dan kompleksitas data. Salah satu metode dalam pengambilan keputusan dapat dilakukan dengan menggunakan pendekatan gabungan *N-soft*

*set* dan *data mining*. Dengan menerapkan *N-soft set*, organisasi dapat memodelkan ketidakpastian dan ambiguitas dalam data, sementara *data mining* memungkinkan identifikasi pola dan tren yang signifikan. Kombinasi kedua pendekatan ini memberikan wawasan yang lebih mendalam tentang data, memungkinkan pengambilan keputusan yang lebih tepat, dan membantu organisasi menghadapi tantangan yang kompleks.

*N-soft set* adalah kerangka kerja matematis yang memungkinkan penanganan ketidakpastian dan ambiguitas dalam data. Dalam konteks pengambilan keputusan, *N-soft set* memungkinkan untuk menyertakan tingkat keyakinan dan ambiguitas dalam proses analisis. Sedangkan *data mining* adalah teknik analisis yang digunakan untuk menggali informasi yang berharga dari data yang kompleks dan besar. Dengan menggabungkan konsep *N-soft set* dengan *data mining*, organisasi dapat menghasilkan pemodelan yang lebih realistis dan mempertimbangkan lebih banyak faktor dalam pengambilan keputusan. Hal ini disebabkan pola data yang digali telah memiliki tanda atau pengenal sehingga memudahkan untuk mengidentifikasi faktor atau fitur dari set data.

*N-soft set* dan *data mining* adalah dua konsep yang berbeda tetapi keduanya memiliki hubungan. Hubungan pertama antara *N-soft set* dan *data mining* ada pada representasi data. *Data mining* melibatkan pemrosesan dan analisis data untuk menemukan pola, tren, dan informasi berharga. *N-soft set* dapat digunakan sebagai salah satu metode representasi data dimana setiap objek dapat memiliki derajat keanggotaan dalam kelas tertentu. Hal ini memungkinkan penggabungan informasi tidak pasti dalam proses analisis data. Dengan memberikan derajat keanggotaan pada setiap objek terhadap kelas tertentu, *N-soft set* memungkinkan representasi yang lebih fleksibel dan akurat dari keadaan nyata. Hal ini memungkinkan penggabungan informasi yang tidak pasti, yang dapat menghasilkan pemodelan yang lebih kaya dan wawasan yang lebih mendalam dalam analisis data. Akibatnya, organisasi dapat mengoptimalkan proses

pengambilan keputusan mereka dengan menggunakan *N-soft set* dalam *data mining*. Organisasi dapat menggabungkan teknik-teknik *data mining* yang kuat dengan representasi data yang lebih inklusif dan adaptif menggunakan *N-soft set*. Hasilnya adalah pengambilan keputusan yang lebih baik dimana organisasi dapat mempertimbangkan tingkat ketidakpastian dan ambiguitas dalam data mereka.

Pengambilan keputusan dengan *data mining* juga dapat menggabungkan *N-soft set* pada tahap *preprocessing data*. Sebelum menerapkan algoritma *data mining*, seringkali perlu dilakukan tahap *preprocessing data* untuk membuat data mentah menjadi data berkualitas. Proses *preprocessing data* terdiri dari *data cleaning*, *data integration*, *data reduction*, dan *data transformation* (Syahfitri, 2017). *N-soft set* dapat membantu dalam *preprocessing data* dengan memperhitungkan ketidakpastian dan ambiguitas dalam pengklasifikasian objek ke dalam kelas-kelas tertentu. Dalam konteks ini, *N-soft set* memungkinkan perhitungan derajat keanggotaan setiap objek dalam kelas tertentu dengan mempertimbangkan kemungkinan variasi atau ketidakpastian dalam pengklasifikasian. Hal ini memungkinkan adanya fleksibilitas dalam pengelompokan objek ke dalam kelas-kelas yang relevan. Dengan memanfaatkan konsep *N-soft set* pada tahap *preprocessing data*, organisasi dapat meningkatkan akurasi dan kualitas hasil *data mining*.

Pencarian pola merupakan salah satu tujuan utama dalam *data mining*, dimana data yang ada akan dianalisis untuk menemukan pola yang relevan atau pengetahuan yang tersembunyi. Dalam konteks ini, metode-metode seperti *Association Rule Mining*, *Clustering*, atau *Classification* dapat diterapkan pada *dataset* yang direpresentasikan menggunakan *N-soft set*. Pendekatan ini memungkinkan penemuan pola dengan mempertimbangkan tingkat keanggotaan objek dalam kelas-kelas tertentu. Dengan menerapkan metode *Association Rule Mining*, *data mining* dapat menemukan hubungan atau keterkaitan antara item-item dalam *dataset* yang direpresentasikan menggunakan

*N-soft set*. Melalui analisis ini, pola asosiasi yang signifikan dapat diidentifikasi, yang mengungkapkan hubungan antara item-item tertentu yang muncul bersama secara berulang. Pendekatan ini memungkinkan organisasi untuk mendapatkan wawasan tentang hubungan dan korelasi yang ada dalam data mereka dengan mempertimbangkan tingkat keanggotaan objek dalam kelas-kelas yang relevan.

Penelitian sebelumnya telah menggabungkan konsep *soft set* dengan *association rule mining*. Penelitian Herawan dan Deris menggunakan *mining regular association rules* dan *maximal association rules* dengan menggunakan teori *soft set* pada *dataset* transaksional (2011). Penelitian dilakukan dengan mengubah transformasi *dataset* transaksional menjadi sistem informasi bernilai *Boolean*. Dalam konteks ini, *dataset* transaksional direpresentasikan sebagai *soft set*. Dengan menggunakan konsep *co-occurrence* parameter dalam sebuah transaksi, peneliti telah mendefinisikan gagasan *regular* dan *maximal association rules* antara dua set parameter, juga *support*, *confidence*, *maximal support*, dan *maximal confidences* yang didefinisikan dengan benar menggunakan teori *soft set*. Hasil penelitian menunjukkan bahwa *soft regular* dan *soft maximal association rules* memberikan aturan yang identik dibandingkan dengan *regular* dan *maximal association rules*. Pada penelitian tersebut meyakini bahwa implementasi teori *soft set* pada *data mining* berbagai tingkatan *association rules* dan sistem pendukung keputusan akan menjadi lebih mudah.

Selain itu, *data mining* dengan menggunakan metode *clustering* dapat mengelompokkan objek-objek dalam *dataset* berdasarkan tingkat keanggotaan mereka dalam kelas-kelas yang diwakili oleh *N-soft set*. Proses ini memungkinkan identifikasi berbagai kelompok yang memiliki karakteristik atau pola yang mirip. Dengan demikian, organisasi dapat mengidentifikasi segmen pelanggan, kategori produk, atau kelompok data lainnya yang dapat memberikan wawasan yang berharga dalam pengambilan keputusan. Penelitian sebelumnya telah menggunakan pendekatan

*soft set* untuk klasterisasi *dataset* penilaian siswa (Saedudin, Kasim, Mahdin, Hasibuan, 2017). Pada penelitian tersebut menggunakan *Maximum Degree of Domination* dalam teori *soft set* (MDDS). MDDS digunakan untuk memilih fitur terbaik dalam pengelompokan data pendidikan. Pemilihan fitur memperhatikan derajat keanggotaan untuk menemukan fitur yang bermakna dari *dataset* sehingga fitur yang ada memiliki korelasi tinggi atau saling ketergantungan tinggi satu sama lain. Hasil yang diperoleh pada penelitian tersebut berpotensi memberikan rekomendasi dalam memberikan nilai akhir suatu mata kuliah dengan lebih cepat dan akurat.

Penerapan metode *classification* dalam *data mining* juga dapat memanfaatkan representasi *N-soft set* dalam menentukan kelas-kelas objek. Pengklasifikasian objek ke dalam kelas tertentu akan mempertimbangkan tingkat keanggotaan objek dalam kelas-kelas yang diwakili oleh *N-soft set*. Dengan cara ini, *data mining* dapat membangun model klasifikasi yang lebih fleksibel dan dapat menangani ketidakpastian atau ambiguitas dalam pengklasifikasian objek. Penelitian sebelumnya telah menggunakan *fuzzy soft set* (FSS) dalam pengklasifikasian berdasarkan pengukuran kesamaan jarak *Hamming*, yang disebut HDFSSC (Yanto, Saedudin, Lashari, & Haviluddin, 2018). Model yang diusulkan melalui empat fase, yaitu akuisisi data, fuzzifikasi fitur, fase pelatihan, dan fase pengujian. Hasil eksperimen menunjukkan bahwa HFSSC memiliki akurasi yang tinggi bahkan dengan kompleksitas yang lebih rendah pada tahap pengujian.

Secara keseluruhan, metode-metode pencarian pola dengan memanfaatkan representasi *N-soft set* dalam *data mining* dapat menghasilkan penemuan yang lebih informatif dan berharga. Dalam proses ini, tingkat keanggotaan objek dalam kelas-kelas tertentu memberikan informasi tambahan yang memperkaya analisis pola yang dilakukan. Penggunaan *N-soft set* dalam memberikan tanda atau ciri kepada objek sehingga dapat diketahui fitur tersebut berpengaruh atau tidak pada variabel terikat. Dengan demikian, penggunaan *N-soft set* dalam pencarian pola dalam *data mining* dapat meningkatkan kualitas hasil analisis dan mendukung pengambilan keputusan yang lebih baik.

## KESIMPULAN

Pengambilan keputusan menggunakan *data mining* memiliki peran penting dalam berbagai sektor. *Data mining* memungkinkan organisasi untuk menggali wawasan yang berharga dari data yang kompleks dan besar, dan mendukung pengambilan keputusan yang lebih baik. Pengambilan keputusan menggunakan kombinasi *N-soft set* dan *data mining* menjadikan pendekatan yang lebih kuat dalam menghadapi ketidakpastian, ambiguitas, dan kompleksitas data. *N-soft set* mengelompokkan objek ke dalam kelas dengan mempertimbangkan tingkat ketidakpastian dan ambiguitas sedangkan *data mining* memungkinkan identifikasi pola dan tren yang signifikan. Dengan menggabungkan kedua pendekatan ini, organisasi dapat mempertimbangkan lebih banyak faktor dalam pengambilan keputusan dan menghasilkan wawasan yang lebih mendalam tentang data yang dianalisis. Untuk pengembangan lebih lanjut, penulis akan meneliti tentang algoritma pengambilan keputusan *N-soft set association rule* serta implementasinya pada data real. Kolaborasi *N-soft set* pada setiap tahapan *data mining* merupakan potensi penelitian lanjutan sebagai upaya terus berinovasi sebagai ilmuwan untuk berkontribusi dalam mengembangkan kemampuan kolaboratif berkelanjutan (*Collaborative Capability for Sustainability*).

## DAFTAR PUSTAKA

- Adeel, A., Akram, M., Yaqoob, N., & Chammam, W. (2020). Detection and severity of tumor cells by graded decision-making methods under fuzzy  $N$ -soft model. *Journal of Intelligent & Fuzzy Systems*, 39(1), 1303-1318.
- Akram, M., Ali, G., Alcantud, J. C., & Fatimah, F. (2021). Parameter reductions in  $N$ -soft sets and their applications in decision-making. *Expert Systems*, 38(1). <https://doi.org/10.1111/exsy.12601>.
- Akram, M., & Adeel, A. (2019). TOPSIS approach for MAGDM based on interval-valued hesitant fuzzy  $N$ -soft environment. *International Journal of Fuzzy Systems*, 21(3), 993-1009.
- Akram, M., Adeel, A., & Alcantud, J. C. R. (2018). Fuzzy  $N$ -soft sets: A novel model with applications. *Journal of Intelligent & Fuzzy Systems*, 35(4), 4757-4771.
- Akram, M., Adeel, A., & Alcantud, J. C. R. (2019a). Group decision-making methods based on hesitant  $N$ -soft sets. *Expert Systems with Applications*, 115, 95-105.
- Akram, M., Adeel, A., & Alcantud, J. C. R. (2019b). Hesitant fuzzy  $N$ -soft sets: A new model with applications in decision-making. *Journal of Intelligent & Fuzzy Systems*, 36(6), 6113-6127.
- Akram, M., Amjad, U., & Davvaz, B. (2021). Decision-making analysis based on bipolar fuzzy  $N$ -soft information. *Computational and Applied Mathematics*, 40(6), 1-39.

- Akram, M., Shabir, M., & Ashraf, A. (2021). Complex neutrosophic N-soft sets: A new model with applications. *Neutrosophic Sets and Systems*, 42, 278-301.
- Alcantud, J. C. R., Feng, F., & Yager, R. R. (2019). An N-soft set approach to rough sets. *IEEE Transactions on Fuzzy Systems*, 28(11), 2996-3007.
- Ashraf, A., & Butt, M. A. (2021). Extension of TOPSIS method under single-valued neutrosophic N-soft environment. *Neutrosophic sets and systems*, 41, 286-303
- Bustamam, A., Rustama, Z., Selly, A. A., Wibawa, N. A., Sarwinda, D., & Husna, N. (2021). Lung cancer classification based on support vector machine-recursive feature elimination and artificial bee colony. *Annals of Mathematical Modeling*, 3(1), 40-52.
- de Almeida, J. M. G., Gohr, C. F., Morioka, S. N., & da Nobrega, B. M. (2021). Towards an integrative framework of collaborative capabilities for sustainability: A systematic review and research agenda. *Journal of Cleaner Production*, 279, 123789.
- Chen, Y., Liu, J., Chen, Z., & Zhang, Y. (2020). Group decision-making method based on generalized Vague N-soft sets. In *2020 Chinese Control And Decision Conference (CCDC)*, 4010-4015. IEEE.
- Fatimah, F. (2020). Pengambilan keputusan multi hesitant N-soft sets. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 4(6), 1110-1116.
- Fatimah, F., (2018). Pengambilan keputusan incomplete N-Soft Sets pada data untuk mengukur indikator sustainable development goals. *Peran Matematika, Sains, dan Teknologi dalam Mencapai Tujuan Pembangunan Berkelanjutan*, 209-228. Tangerang Selatan: Universitas Terbuka.

- Fatimah, F. (2021a). *N*-soft sets: Tantangan dalam riset big data. *Science and Technology for Society 5.0* (pp.1-23), Tangerang Selatan, Universitas Terbuka.
- Fatimah, F. (2021b). *N*-soft sets: Literature review and research potential. *Proceeding of the 1st International Seminar of Science and Technology for Society Development ISST 2021*, 27-39.
- Fatimah, F., & Andriyansah (2020). Analisis fasilitas pariwisata menggunakan prosedur pengambilan keputusan *N*-soft set. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(1), 135-141.
- Fatimah, F., & Alcantud, J.C.R. (2021). The multi-fuzzy *N*-soft set and its applications to decision-making. *Neural Computing & Application*, 33(17), 11437-11446.
- Fatimah, F., Rosadi, D., Hakim, RB. F., Alcantud, J. C. R. (2017). A social choice approach to graded soft sets, *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1-6. doi: 10.1109/FUZZ-IEEE.2017.8015428.
- Fatimah, F., Rosadi, D., Hakim, R.B.F., & Alcantud, J. C. R. (2018). *N*-soft sets and their decision making algorithms. *Soft Computing*, 2, 3829-3842. <https://doi.org/10.1007/s00500-017-2838-6>.
- Haba, A. R. K. (2021). Penerapan metode linier regresi dalam prediksi produksi pia. *Jurnal Sistem dan Teknik Komputer*, 6(2), 178-180.
- Herawan, T., & Deris, M. M. (2011). A soft set approach for association rules mining. *Knowledge-Based Systems*, 24(1), 186-195.

- Indriyawati, H., & Winarti, T. (2021). Pemodelan data mining pola kelayakan kemampuan lulusan dengan kebutuhan stakeholder menggunakan algoritma apriori. *JITSI: Jurnal Ilmiah Teknologi Sistem Informasi*, 2(3), 78-84
- Kamacı, H. (2020). Introduction to N-soft algebraic structures. *Turkish Journal of Mathematics*, 44(6), 2356-2379.
- Kamacı, H., & Petchimuthu, S. (2020). Bipolar N-soft set theory with applications. *Soft Computing*, 24(22), 16727-16743.
- Kamacı, H., & Saqlain, M. (2021).  $n$ -ary fuzzy hypersoft expert sets. *Neutrosophic Sets and Systems*, 43, 180-211.
- Kharis, S. A. A., Hadi, I., & Hasanah, K. A. (2019, December). Multiclass classification of brain cancer with multiple multiclass Artificial bee colony feature selection and support vector machine. In *Journal of Physics: Conference Series* (Vol. 1417, No. 1, p. 012015). IOP Publishing.
- Kharis, S. A. A., Hertono, G. F., Wahyuningrum, E., Yumiati, Y., Irawan, S. R., Danial, T. A., & Saputra, D. S. (2023). Design of student success prediction application in online learning using Fuzzy-KNN. *Barekeng: Jurnal Ilmu Matematika dan Terapan*, 17(2), 0971-0980.
- Liu, J., Chen, Y., Chen, Z., & Zhang, Y. (2020). Multi-attribute decision making method based on neutrosophic vague N-soft sets. *Symmetry*, 12(5), 853.
- Mahmood, T., Ur Rehman, U., & Ali, Z. (2021). A novel complex fuzzy N-soft sets and their decision-making algorithm. *Complex & Intelligent Systems*, 1-26.

- Rehman, U. U., & Mahmood, T. (2021). Picture fuzzy N-soft sets and their applications in decision-making problems. *Fuzzy Information and Engineering*, 1-33.
- Render, B., Stair Jr, R. M., Hanna, M. E., Hale. T. S. (2018). *Quantitative analysis for management*, 13th Edition. Pearson Education India.
- Riaz, M., Çağman, N., Zareef, I., & Aslam, M. (2019). N-soft topology and its applications to multi-criteria group decision making. *Journal of Intelligent & Fuzzy Systems*, 36(6), 6521-6536.
- Riaz, M., Naeem, K., Zareef, I., & Afzal, D. (2020). Neutrosophic N-soft sets with TOPSIS method for multiple attribute decision making. *Neutrosophic sets and systems*, 32, 146-170.
- Riaz, M., Razaq, A., Aslam, M., & Pamucar, D. (2021). M-parameterized N-soft topology-based TOPSIS approach for multi-attribute decision making. *Symmetry*, 13(5), 748.
- Rustam, Z., & Kharis, S. A. A. (2020). Comparison of support vector machine recursive feature elimination and kernel function as feature selection using support vector machine for lung cancer classification. In *Journal of Physics: Conference Series* (Vol. 1442, No. 1, p. 012027). IOP Publishing.
- Saedudin, R. R., Kasim, S. B., Mahdin, H., & Hasibuan, M. A. (2017). Soft set approach for clustering graduated dataset. In *Recent Advances on Soft Computing and Data Mining: The Second International Conference on Soft Computing and Data Mining (SCDM-2016)*, Bandung, Indonesia, August 18-20, 2016 Proceedings Second (pp. 631-637). Springer International Publishing.

- Surbakti, N. K. (2021). Data mining pengelompokan pasien rawat inap peserta BPJS menggunakan metode clustering (Studi kasus: RSUD. Bangkatan). *Journal of Information and Technology*, 1(2), 47-53.
- Syahfitri, I. N. (2017). Penerapan data mining untuk menentukan besar pinjaman pada koperasi simpan pinjam dengan algoritma C4. 5 (Studi Kasus: Koperasi Wanita XYZ). *Jurnal Informatika dan Sistem Informasi*, 3(2), 18-27.
- Syahril, M., Erwansyah, K., & Yetri, M. (2020). Penerapan data mining untuk menentukan pola penjualan peralatan sekolah pada brand wigglo dengan menggunakan algoritma apriori. *Jurnal Teknologi Sistem Informasi Dan Sistem Komputer TGD*, 3(1), 118-136.
- Muhammad Wali, S. T., Efitra, S., Kom, M., Sudipa, I. G. I., Kom, S., Heryani, A., ... & Sepriano, M. (2023). Penerapan & implementasi Big Data di berbagai sektor (Pembangunan berkelanjutan era industri 4.0 dan society 5.0). PT. Sonpedia Publishing Indonesia.
- Yanto, I. T. R., Saedudin, R. R., Lashari, S. A., & Havaluddin. (2018). A numerical classification technique based on fuzzy soft set using hamming distance. In *Recent Advances on Soft Computing and Data Mining: Proceedings of the Third International Conference on Soft Computing and Data Mining (SCDM 2018)*, Johor, Malaysia, February 06-07, 2018 (pp. 252-260). Springer International Publishing.
- Zhang, H., Jia-Hua, D., & Yan, C. (2020). Multi-attribute group decision-making methods based on Pythagorean fuzzy N-soft sets. *IEEE Access*, 8, 62298-62309.

- Zhang, H., Jia-Hua, D., & Yan, C. (2020). Multi-attribute group decision-making methods based on Pythagorean fuzzy N-soft sets. *IEEE Access*, 8, 62298-62309.
- Zili, A. H. A., Hendri, D., & Kharis, S. A. A. (2022). Peramalan harga saham dengan model hybrid arima-Garch dan metode Walk Forward. *Jurnal Statistika dan Aplikasinya*, 6(2), 341-354.

