

Teknologi *Data Mining* Berbasis Metode *Clustering* Sebagai Ujung Tombak Perkembangan Umkm Di Indonesia Dalam Era Revolusi Industri 4.0

Ernowo Gordon Sukoco Utojo Unus
Universitas Muhammadiyah Malang
ernowogordon18@webmail.umm.ac.id

Abstrak

UMKM di Indonesia membutuhkan sebuah sistem baru yang dapat mengolah data penjualan, pemasaran, produksi mereka dengan cepat dan akurat, lewat teknologi Data Mining berbasis metode clustering menjadikan salah satu hal yang harus ada dalam pengolahan data yang besar di dunia UMKM. Era revolusi 4.0 menuntut kita semua agar lebih berkembang dari segi teknologi. Pengolahan data lewat metode clustering bisa memudahkan para pelaku UMKM dalam mengelola data yang banyak, tanpa adanya hambatan lagi dalam proses pengolahan data. Menginput data produksi atau pemasaran ke dalam server terkadang masih belum akurat tanpa adanya pengolahan data yang baik dan benar. Data Mining berbasis metode clustering menggunakan Teknik bottom-up yang menggabungkan cluster kecil menjadi cluster lebih besar dan top-down yang memecah cluster besar menjadi cluster yang lebih kecil. Teknik tersebut akan membuat pengolahan data menjadi lebih efisien ke depannya.

Kata Kunci: *Data Mining, Clustering, UMKM*

Abstract

MSMEs in Indonesia need a new system that can process their sales, marketing, production data quickly and accurately, through data mining technology based on the clustering method, making it one of the things that must be present in large data processing in the world of MSMEs. The era of revolution 4.0 requires all of us to develop in terms of technology. Data processing through the clustering method can make it easier for MSME players in managing a lot of data, without any obstacles in the data processing process. Entering production or marketing data into the server is sometimes still inaccurate without proper data processing. Data Mining based on the clustering method uses a bottom-up technique which combines small clusters into larger and top-down clusters which break up large clusters into smaller clusters. This technique will make data processing more efficient in the future.

Keywords: *Data mining, clustering, UMKM*

1. Pendahuluan

Seluruh teknologi yang ada di dunia sekarang telah berkembang sangat pesat dan akan terus meningkat seiring dengan berjalannya waktu. Manusia menggunakan teknologi karena memiliki akal. Dengan akalnya manusia ingin keluar dari masalah, ingin hidup lebih baik, lebih aman, dan sebagainya. Perkembangan teknologi terjadi karena seseorang menggunakan akalnya untuk menyelesaikan setiap masalah yang dihadapinya[1]. Banyak ilmuwan, peneliti, pengembang, serta para sarjana muda yang berhasil membuat serta mengembangkan teknologi yang mereka buat saat ini. Kemudian diaplikasikan ke seluruh sektor, baik sektor pendidikan, sosial, ekonomi, dan lain-lain.

Pada saat ini kita telah memasuki Era Revolusi Industri 4.0 dimana terjadi perkembangan yang sangat pesat dalam bidang teknologi. Revolusi Industri 4.0 merupakan kemajuan teknologi baru yang mengintegrasikan dunia fisik, digital, dan biologis, dimana terdapat perubahan cara hidup kerja manusia secara fundamental[2]. Istilah Industri 4.0 sendiri secara resmi lahir di Jerman tepatnya saat diadakan Hannover Fair pada tahun 2011. Negara Jerman memiliki kepentingan yang besar terkait hal ini karena Industri 4.0 menjadi bagian dari kebijakan rencana pembangunannya yang disebut *High-Tech Strategy 2020*.

Sebagian besar pendapat mengenai potensi manfaat Industri 4.0 adalah mengenai perbaikan kecepatan fleksibilitas produksi, peningkatan layanan kepada pelanggan dan

peningkatan pendapatan. Terwujudnya potensi manfaat tersebut akan memberikan dampak positif terhadap perekonomian suatu negara[3]. Jika suatu negara benar-benar memanfaatkan kesempatan ini dengan baik dan benar maka bukan tidak mungkin akan terlahir sebuah negara yang dapat bersaing secara internasional dalam Era Revolusi Industri 4.0.

Era Revolusi Industri 4.0 juga ditandai dengan bertumbuhnya tren dibidang Otomasi, Internet of Things (IoT), Big Data, dan Teknologi Cloud Computing. Era Industri 4.0 memberikan peluang bagi pekerjaan *data science*. *Data science* adalah bidang yang mempelajari bagaimana data yang terstruktur maupun tidak terstruktur dapat digunakan melalui teknik analisa data sebagai pertimbangan untuk mengambil keputusan. Banyaknya jumlah data yang dihasilkan oleh manusia dan perangkat terus mendorong perkembangan teknologi yang dapat mawadahi kebutuhan untuk mengatur dan memproses informasi yang terus bertambah.

Data science atau ilmu data semakin marak digunakan di berbagai industri, kini *data science* menjadi alat yang sangat berharga bagi industri untuk mentransformasi data menjadi informasi penting. Menurut Founder PHI-Integration Feris Thia, dalam perusahaan biasanya terjadi kesenjangan antara TI, data, dan bisnis. Sebagai solusi dari kesenjangan tersebut maka *data science* dapat menjadi alternatif untuk mengatasinya. Dalam hal ini orang yang bekerja sebagai *data science* disebut dengan data scientist. Hasil analisis dari data scientist kemudian dimodelkan menggunakan machine learning. Salah satu aplikasi yang saat ini dapat digunakan untuk machine learning adalah microsoft azure visual studio. Aplikasi ini sudah include dengan bahasa pemrograman R dan Python sebagai logic untuk output program sehingga menghasilkan insight atau rekomendasi[4].

Dalam ilmu data science ada yang disebut dengan *Data mining*. Dalam perkembangan *data mining* (DM) yang pesat tidak dapat lepas dari perkembangan teknologi informasi yang memungkinkan data dalam jumlah besar terakumulasi. Sebagai contoh, toko swalayan merekam setiap penjualan barang dengan memakai alat POS point of sales). Database data penjualan tsb. bisa mencapai beberapa GB setiap harinya untuk sebuah jaringan toko swalayan berskala nasional. Perkembangan internet juga punya andil cukup besar dalam akumulasi data[5].

Data mining (DM) adalah salah satu bidang yang berkembang pesat karena besarnya kebutuhan akan nilai tambah dari database skala besar yang makin banyak terakumulasi sejalan dengan pertumbuhan teknologi informasi. Definisi umum dari DM itu sendiri adalah serangkaian proses untuk menggali nilai tambah berupa pengetahuan yang selama ini tidak diketahui secara manual dari suatu kumpulan data[5]. Kita harus bisa memanfaatkan peluang dalam bidang teknologi, terutama pada bidang Big Data yang mengacu pada sistem *Data mining* untuk Masyarakat Ekonomi Asean (MEA) 2020, dimana tantangan Usaha Mikro Kecil dan Menengah (UMKM) harus dapat terintegrasi dengan sistem secara otomisasi.

Saat ini Indonesia menghadapi era Revolusi Industri 4.0, dimana rantai suplai bisnis terhubung secara digital yang berakibat penyederhanaan rantai suplai bagi kaum milenial yang kreatif dan inovatif menjadi peluang dan tantangan dalam menjalankan bisnis di Indonesia. [6]. Dengan adanya sistem *data mining* akan lebih memudahkan usaha mikro kecil dan menengah UMKM dalam aktivitas perdagangan serta Bisnis yang dilakukan oleh para pelaku usaha. *data mining* berbasis metode *clustering* yaitu merupakan pengelompokan data tanpa berdasarkan kelas data tertentu.

Prinsip dari *data mining* dengan menggunakan metode *clustering* yaitu memaksimalkan kesamaan antara anggota satu kelas dan meminimumkan kesamaan antar kelas atau cluster[5]. Pengelompokan ini akan membuat pelaku UMKM akan lebih mudah dalam pengerjaan suatu data yang mereka kelola. baik itu dalam bidang keuangan bidang pemasaran dan juga bidang-bidang lainnya yang berkaitan dengan setiap aktivitas UMKM tersebut. Dengan adanya teknologi *data mining* berbasis metode *clustering* maka perkembangan UMKM yang ada di Indonesia akan sangat berkembang pesat karena pada setiap aspek ada di UMKM itu memerlukan pengelompokan data yang harus dikelola dengan baik dan benar menggunakan sistem yang terintegrasi dan terstruktur. Adanya penelitian ini diharapkan akan memudahkan hal tersebut bagi para pelaku UMKM yang ada di Indonesia.

Artikel ini memiliki banyak keterkaitan dengan penelitian sebelumnya yaitu "REVOLUSI INDUSTRI 4.0 MERUPAKAN PELUANG DAN TANTANGAN BISNIS BAGI GENERASI MILENIAL DI INDONESIA" Cisilia Sundari, Fakultas Ekonomi Universitas Tidar. Yang membahas tentang peluang bagi UMKM dan Usaha Kreatif sangat terbuka luas, dan bahkan mendapatkan perhatian khusus dari pemerintah. Dukungan kepada UMKM dan Usaha Kreatif berupa: pendanaan, pelatihan dan pendampingan- pendampingan bagi pemula, dengan tujuan agar

UMKM dan Usaha Kreatif tersebut dapat tumbuh dan berkembang sehingga secara makro akan berdampak pada semakin kokohnya perekonomian Negara. Selain untuk pertumbuhan ekonomi bagi negara berfungsi juga sebagai sarana pemerataan kemakmuran bagi masyarakat[6].

Melalui artikel ini penulis akan lebih khusus membahas Bagaimana cara UMKM memanfaatkan data sains sebagai dasar untuk mengembangkan bisnis mereka di era revolusi industri 4.0. Dengan mengacu pada teknologi *Data Mining* maka dengan penulis berkeinginan agar para UMKM bisa lebih mudah untuk menjalankan bisnis mereka sesuai dengan perkembangan teknologi yang ada di dunia saat ini. Semakin maju teknologi yang digunakan dan semakin terstruktur teknologi yang dipakai maka penulis beranggapan bisnis yang dilakukan oleh para pelaku UMKM akan lebih maju dan lebih untung dari segi sistem dan perencanaan.

Pada artikel ini penulis akan mengkaji lebih dalam bagaimana Teknologi *Data Mining* Berbasis Metode *Clustering* Sebagai Ujung Tombak Perkembangan Umkm Di Indonesia Dalam Era Revolusi, dalam hal ini metode penelitian yang digunakan adalah Metode Deskriptif karena metode ini digunakan untuk mencari unsur-unsur, ciri-ciri, sifat-sifat atau sesuai atau cocok dengan harapan atau teori yang sudah baku [7]. Dengan menggunakan metode penelitian tersebut diharapkan dapat mengumpulkan data dari berbagai sumber dan dapat dijadikan sebuah artikel ilmiah yang sesuai dengan dasar-dasar kaidah yang ada.

2. Metode Penelitian

Penelitian ini menggunakan metode kualitatif hal ini dilakukan dengan tujuan memahami peristiwa, kegiatan, perilaku dan pelaku peristiwa dalam kondisi tertentu, serta dalam situasi ilmiah (natural). Pemilihan metode kualitatif dalam penelitian ini untuk memberikan rangkaian dan gambaran tentang kondisi yang dihadapi oleh para pelaku UMKM di Indonesia dalam menghadapi Revolusi Industri 4.0 tentang peluang dan tantangannya. Metode pengumpulan data menggunakan desk study yaitu cara pengumpulan data dan informasi melalui pemeriksaan dan analisis data dan informasi yang menggunakan data sekunder. Data sekunder diperoleh melalui penelusuran literatur. Yang menjadi sumber data dalam penelitian ini adalah literatur, artikel, jurnal, penelitian ilmiah, serta laman internet yang berkenaan dengan penelitian yang dilakukan.

Objek penelitian merupakan masalah yang diteliti dalam penelitian. Penelitian ini mengambil objek kegiatan bisnis yang dilakukan kaum milenial yang berlokasi di Indonesia. Subjek penelitian adalah pelaku utama yang akan diteliti. Yang menjadi subjek dalam penelitian ini adalah para pelaku UMKM di Indonesia yang melakukan bisnis dengan memanfaatkan kondisi Revolusi Industri 4.0 yang terjadi. Kemudian subject tersebut kita teliti serta mencari solusi. Solusi yang ingin diberikan yaitu bagaimana para pelaku UMKM bisa menggunakan Teknologi *Data Mining* Berbasis Metode *Clustering* sebagai ujung tombak dalam menjalankan bisnis mereka. Penelitian ini menggunakan pendekatan 3 tahapan sebagai berikut:

1. Tahap pertama adalah tahap Orientasi atau deskripsi dengan grand tour question. Pada tahap ini peneliti mendeskripsikan apa yang dilihat, didengar, dirasakan dan ditanyakan. Peneliti biasanya baru mengenal sepintas informasi yang diperolehnya. Dalam tahap deskripsi data yang diperoleh cukup banyak, bervariasi dan belum tersusun secara jelas.
2. Tahap kedua adalah tahap reduksi/fokus, peneliti mereduksi segala informasi yang telah diperoleh pada tahap pertama untuk memfokuskan pada masalah tertentu. Memilih data mana yang menarik, penting dan berguna, serta baru. Kemudian data tersebut dikelompokkan dalam dikategori yang ditetapkan sebagai fokus penelitian.
3. Tahap ketiga adalah tahap selection. Pada tahap ini peneliti menguraikan fokus yang telah ditetapkan menjadi lebih rinci.

Teknik pengumpulan data dalam penelitian kualitatif digunakan teknik wawancara secara mendalam in dept interview, observasi berperan serta, dan dokumentasi [7].

3. Hasil Penelitian dan Pembahasan

3.1 Data Mining

Perkembangan data mining (DM) yang pesat tidak dapat lepas dari perkembangan teknologi informasi yang memungkinkan data dalam jumlah besar terakumulasi. Sebagai contoh, toko swalayan merekam setiap penjualan barang dengan memakai alat POS (point of sales). Database data penjualan tsb. bisa mencapai beberapa GB setiap harinya untuk sebuah jaringan toko swalayan berskala nasional. Perkembangan internet juga punya andil cukup besar dalam akumulasi data. Tetapi pertumbuhan yang pesat dari akumulasi data itu telah menciptakan kondisi yang sering disebut sebagai "rich of data but poor of information" karena data yang

terkumpul itu tidak dapat digunakan untuk aplikasi yang berguna. Tidak jarang kumpulan data itu dibiarkan begitu saja seakan-akan “kuburan data” (data tombs)[5].

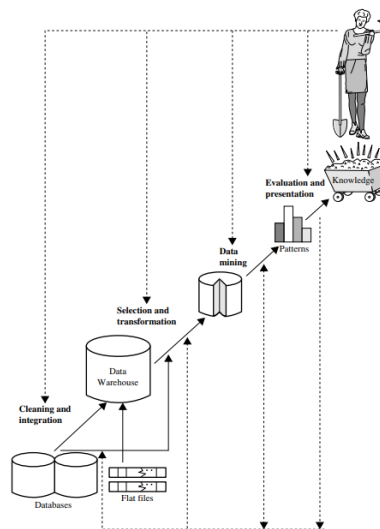
DM adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Patut diingat bahwa kata mining sendiri berarti usaha untuk mendapatkan sedikit barang berharga dari sejumlah besar material dasar. Karena itu DM sebenarnya memiliki akar yang panjang dari bidang ilmu seperti kecerdasan buatan (artificial intelligent), machine learning, statistik dan database. Beberapa teknik yang sering disebut-sebut dalam literatur DM antara lain: clustering, classification, association rule mining, neural network, genetic algorithm dan lain-lain.

Yang membedakan persepsi terhadap DM adalah perkembangan teknik-teknik DM untuk aplikasi pada database skala besar. Sebelum populernya DM, teknik-teknik tersebut hanya dapat dipakai untuk data skala kecil saja.

3.2 Proses Data Mining

Tahap-Tahap Data Mining Karena DM adalah suatu rangkaian proses, DM dapat dibagi menjadi beberapa tahap digambarkan melalui gambar berikut Gambar 1 [8].

1. Pembersihan data (untuk membuang data yang tidak konsisten dan noise)
2. Integrasi data (penggabungan data dari beberapa sumber)
3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)
4. Aplikasi teknik DM
5. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai)
5. Presentasi pengetahuan (dengan teknik visualisasi)



Gambar 1. Ilustrasi proses Data Mining [8].

3.3 Teknik Data Mining (Clustering)

Banyak Teknik di dalam Data Mining, salah satu tekniknya adalah Teknik Clustering. Clustering atau klusterisasi adalah metode pengelompokan data. Menurut Tan, 2006 clustering adalah sebuah proses untuk mengelompokkan data ke dalam beberapa cluster atau kelompok sehingga data dalam satu cluster memiliki tingkat kemiripan yang maksimum dan data antar cluster memiliki kemiripan yang minimum[9].

Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut dengan cluster. Objek yang di dalam cluster memiliki kemiripan karakteristik antar satu sama lainnya dan berbeda dengan cluster yang lain. Partisi tidak dilakukan secara manual melainkan dengan suatu algoritma clustering. Oleh karena itu, clustering sangat berguna dan bisa menemukan group atau kelompok yang tidak dikenal dalam data. Clustering banyak digunakan dalam berbagai aplikasi seperti misalnya pada business intelligence, pengenalan pola citra, web search, bidang ilmu biologi, dan untuk keamanan (security). Di dalam business intelligence, clustering bisa mengatur banyak customer ke dalam banyaknya kelompok. Contohnya mengelompokkan customer ke dalam beberapa cluster dengan kesamaan karakteristik yang kuat. Clustering juga dikenal sebagai data segmentasi karena clustering

mempartisi banyak data set ke dalam banyak group berdasarkan kesamaannya. Selain itu clustering juga bisa sebagai outlier detection.

Clustering melakukan pengelompokan data tanpa berdasarkan kelas data tertentu. Bahkan clustering dapat dipakai untuk memberikan label pada kelas data yang belum diketahui itu. Karena itu clustering sering digolongkan sebagai metode unsupervised learning. Prinsip dari clustering adalah memaksimalkan kesamaan antar anggota satu kelas dan meminimumkan kesamaan antar kelas/cluster. Clustering dapat dilakukan pada data yang memiliki beberapa atribut yang dipetakan sebagai ruang multidimensi. Banyak algoritma clustering memerlukan fungsi jarak untuk mengukur kemiripan antar data, diperlukan juga metode untuk normalisasi bermacam atribut yang dimiliki data.

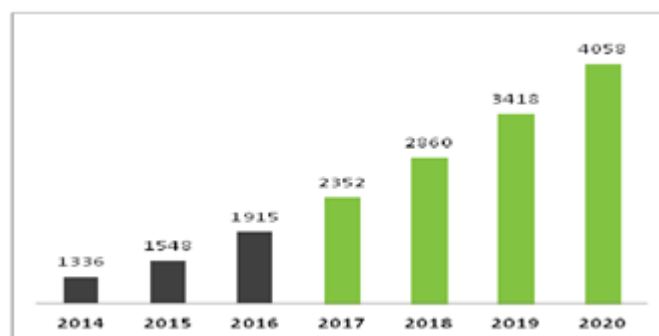
Beberapa kategori algoritma clustering yang banyak dikenal adalah metode partisi dimana pemakai harus menentukan jumlah k partisi yang diinginkan lalu setiap data dites untuk dimasukkan pada salah satu partisi, metode lain yang telah lama dikenal adalah metode hierarki yang terbagi dua lagi : bottom-up yang menggabungkan cluster kecil menjadi cluster lebih besar dan top-down yang memecah cluster besar menjadi cluster yang lebih kecil. Kelemahan metode ini adalah bila salah satu penggabungan/pemecahan dilakukan pada tempat yang salah, tidak dapat didapatkan cluster yang optimal[5].

3.4 Hasil

Dalam Penelitian ini harus memiliki data yang lengkap terlebih dahulu. Kita butuh akses yang cepat, terukur, dan terbuka terhadap data yang tersedia dan fleksibilitas untuk mengekstrak dan memperbarui subset data dalam lingkungan yang akan digunakan untuk data mining. Contoh lingkungan yang baik adalah lingkungan yang mendukung BI dan mudah dalam membuat pasar data dari gudang pusat. Disamping mengakses dan mengirim data, arsitektur IT juga harus bisa mempunyai kapasitas yang cukup untuk melakukan proses data mining, atau mudah untuk menambah kapasitas. Hal ini berarti menambah sistem ekstra terdedikasi untuk mining, atau menjalankannya sebagai tambahan dari proses yang sudah ada dalam salah satu sistem.

Data tentu saja harus tersedia. Jumlah data mentah biasanya bukan masalah, tetapi jumlah data yang bersih, bermanfaat, relevan, dan terintegrasi mungkin kurang dari yang dipikirkan. Tidak ada aturan tetap tentang jumlah data yang dibutuhkan untuk memulai mining. Sebagai aturan dasar, beberapa ribu records, dan sepuluh atau lebih atribut, adalah awal yang baik. Angka tersebut bergantung dari teknik data mining yang digunakan. Alat yang digunakan untuk data mining harus dapat mendukung akses data, praproses, mining, visualisasi, penyimpanan, dan pemeliharaan hasil. Hal ini dapat didukung dari paket tunggal, atau mungkin membutuhkan beberapa alat. Integrasi antar alat adalah penting.

Kita juga harus memperhatikan skalabilitas alat yang akan digunakan. Kita akan selalu ingin menambah data ekstra, mengeksplorasi banyak sejarah, atau menerima hasil lebih cepat. Waktu pemrosesan seharusnya tidak bergerak lebih jauh dari linieritas jumlah data, baik jumlah atribut maupun jumlah records.



Gambar 2. Data perkembangan UMKM

Berdasarkan Gambar 2 diatas, perkembangan UMKM dari tahun ke tahun makin meningkat seiring perkembangan zaman serta teknologi. Menurut riset Deloitte tahun 2016, dari 57,9 juta UMKM di Indonesia (berdasarkan data Kementerian Koperasi & UMKM, tahun 2015) ternyata hanya 9 persen dari pelaku yang serius menggunakan internet untuk menjual produknya dengan jejaring sosial yang terintegrasi maupun menggunakan platform e-commerce

(Kompas.com). Para UMKM perlu memerlukan sistem yang bisa menunjang keberlangsungan bisnis mereka dengan baik.

4. Kesimpulan

Di Era Revolusi Industri 4.0 perubahan secara signifikan terjadi di segala bidang, termasuk juga bidang ekonomi. Dalam bidang ekonomi dengan adanya perkembangan teknologi digital yang di manfaatkan oleh kaum milenial dalam meraih sukses dibidang usaha sangat terbuka luas. Kaum milenial dengan berbekal ketekunan, kreatif dan inovatif serta memanfaatkan fasilitas teknologi digital dapat membangun UMKM dan Usaha Kreatif yang memiliki peluang besar untuk dapat meraih kesuksesan. Di Era ini bukan jamannya lagi perusahaan kecil menjadi mangsa bagi perusahaan besar, bahkan bisa terjadi usaha pemula dengan berbekal ketekunan, kreatif dan inovatif dapat bersaing dengan perusahaan yang besar.

Referensi

- [1] M. Ngafifi, "Kemajuan Teknologi Dan Pola Hidup Manusia Dalam Perspektif Sosial Budaya," *J. Pembang. Pendidik. Fondasi dan Apl.*, vol. 2, no. 1, pp. 33–47, 2014, doi: 10.21831/jppfa.v2i1.2616.
- [2] H. Hamdan, "Industri 4.0: Pengaruh Revolusi Industri Pada Kewirausahaan Demi Kemandirian Ekonomi," *J. Nusant. Apl. Manaj. Bisnis*, vol. 3, no. 2, p. 1, 2018, doi: 10.29407/nusamba.v3i2.12142.
- [3] H. Prasetyo and W. Sutopo, "Industri 4.0: Telaah Klasifikasi Aspek Dan Arah Perkembangan Riset," *J@ti Undip J. Tek. Ind.*, vol. 13, no. 1, p. 17, 2018, doi: 10.14710/jati.13.1.17-26.
- [4] J. Barnes, "Chapter 13 An Introduction To Cross- Curricular Learning," no. June, 2015.
- [5] I. Pramudiono, "Pengantar Data Mining : Menambang Permata Pengetahuan di Gunung Data," *Kuliah Umum Ilmu Komputer.com*, pp. 1–4, 2003.
- [6] C. Sundari, "Revolusi Industri 4.0 Merupakan Peluang Dan Tantangan Bisnis Bagi Generasi Milenial Di Indonesia," *Pros. Semin. Nas. DAN CALL Pap.*, no. Fintech dan E-Commerce untuk Mendorong Pertumbuhan UMKM dan Industri Kreatif, pp. 555–563, 2019.
- [7] Ms. Prof. Dr. Suryana, "Metodologi Penelitian : Metodologi Penelitian Model Praktis Penelitian Kuantitatif dan Kualitatif," *Univ. Pendidik. Indones.*, pp. 1–243, 2012, doi: 10.1007/s13398-014-0173-7.2.
- [8] S. Agarwal, *Data mining: Data mining concepts and techniques*. 2014.
- [9] R. L. Hale, "Cluster analysis in school psychology: An example," *J. Sch. Psychol.*, vol. 19, no. 1, pp. 51–56, 1981, doi: 10.1016/0022-4405(81)90007-8.