

## Implementasi DBSCAN dalam *Clustering* Data Minat Mahasiswa Setelah Pandemi Covid19

A Kristianto\*<sup>1</sup>

<sup>1</sup>Institut Informatika Indonesia

E-mail: kristiantoariel@gmail.com<sup>1</sup>

**Abstrak.** Pandemi Covid19 mengubah pola kehidupan manusia, termasuk dunia pendidikan yang juga ikut terkena imbas dari pandemi ini. Proses Kegiatan Belajar Mengajar (KBM) di sekolah maupun pendidikan tinggi harus beradaptasi dengan kondisi saat pandemi ini. Proses KBM terpaksa harus dilaksanakan melalui media daring atau *online*. Hal ini dilakukan demi memutus rantai penyebaran virus Covid19. Penelitian ini bertujuan mencari minat mahasiswa dalam metode perkuliahan setelah pandemi ini usai, apakah dilaksanakan secara daring, luring ataupun gabungan (*hybrid*). Variabel yang digunakan dalam penelitian ini antara lain usia, jenis kelamin, semester, IPK dan segi keilmuan (Eksakta dan Non-Eksakta). Dalam proses mencari variabel penentu tersebut digunakan algoritma DBSCAN untuk mengelompokkan variabel yang berpengaruh dalam minat metode perkuliahan mahasiswa tersebut. Dari 385 responden didapatkan hasil faktor penentu mahasiswa dalam memilih minat metode perkuliahan adalah semester, IPK, dan kelompok ilmu (Eksakta atau Non-Eksakta) sementara usia dan jenis kelamin tidak mempengaruhi pemilihan minat metode perkuliahan mahasiswa tersebut.

**Kata kunci:** metode perkuliahan, daring, luring, DBSCAN

**Abstract.** *The Covid-19 pandemic has changed the pattern of human life, including in the world of education which has also been affected by this pandemic. The process of teaching and learning activities (KBM) in schools and higher education must adapt to the conditions during this pandemic. The KBM process had to be carried out through online media. This was done in order to break the chain of spreading the Covid-19 virus. This study aims to find students' interest in the learning method after this pandemic is over, whether it is carried out online, offline or hybrid. The variables used in this research included age, gender, semester, GPA and scientific aspects (Exact and Non-Exact). In the process of finding the determining variable, the DBSCAN algorithm was used to classify variables that influence the interest in the student's lecture method. From 385 respondents, it was found that the determining factors for students in choosing the interest in the lecture method were semester, GPA, and science group (Exact or Non-Exact) while age and gender did not affect the choice of interest in the student's learning method.*

**Keywords:** *learning method, online, offline, DBSCAN.*

## 1. Pendahuluan

Virus Corona adalah virus yang berasal dari keluarga virus yang menularkan penyakit yang terjadi pada hewan ataupun manusia. Covid19 atau *Corona Virus Disease* adalah varian dari virus Corona yang terjadi pada tahun 2019 [1]. Pasien yang terdeteksi positif terpapar virus Covid19 semakin hari semakin sulit dikendalikan dan pemerintah membutuhkan jalan keluar yang jelas dan sederhana untuk menangani masalah ini. Virus corona adalah virus yang datang dari subfamili Orthocoronavirinae, famili Coronaviridae dan dari ordo Nidovirales. Virus ini tidak hanya menginfeksi hewan tetapi juga manusia, infeksi pada manusia memiliki gejala sebagai infeksi yang mirip dengan SARS dan MERS, hanya saja Covid19 memiliki efek yang lebih kuat. Pemerintah pusat melakukan langkah-langkah pencegahan penyebaran virus berupa Pembatasan Sosial Berskala Besar (PSBB) seperti yang tertuang dalam PP Nomor 21 Tahun 2020. Pemerintah juga mengadakan pengujian dengan tes masal dengan *rapid test* atau *swab test* untuk memeriksa apakah orang tersebut positif atau negatif dengan Covid19 [2][3][4].

Pandemi Covid19 juga mengubah cara pelaksanaan pendidikan di Indonesia. Siap atau tidak, cara belajar mengajar ikut berubah. Murid Sekolah Dasar hingga Sekolah Menengah Atas, bahkan mahasiswa harus belajar dari rumah secara daring [5][6][7]. Kebijakan belajar secara daring di rumah jelas memberikan gangguan yang besar, seperti dalam proses KBM, penilaian, peluang mendapatkan pekerjaan setelah lulus, dan penilaian dalam seleksi pekerjaan. Hal ini menjadi kekhawatiran bagi lulusan siswa dan mahasiswa [8]. Kerja sama timbal balik antara masing-masing siswa dan guru, mahasiswa dan dosen menjadi hal yang mutlak dilaksanakan untuk mencapai hasil pembelajaran yang efektif. Kendalanya adalah situasi pembelajaran daring ini mengakibatkan penurunan motivasi belajar siswa [9], dan kehadiran orangtua juga menjadi penting untuk mendapatkan hasil lulusan yang baik [10].

Ada banyak pengolahan data yang dapat dilakukan, *linear regression*, *Minimum Distance*, *Mahalanobis Distance*, *Maximum Likelihood*, *Parallelepiped*, *k-Nearest Neighbor*, *Isodata*, *k-Means*, *Improved Split and Merge Classification (ISMC)*, *Naive Bayesian*, *Adaptive Clustering*, dan DBSCAN[11][12]. *Density-Based Spatial Clustering Algorithm with Noise (DBSCAN)* adalah salah satu algoritma yang sering digunakan dalam proses *clustering*. DBSCAN memiliki keunggulan dalam proses *clustering* yaitu DBSCAN mengelompokkan berdasar kepada kepadatan data atau *density* sehingga unggul dalam pengelompokan data yang besar. Kepadatan data dalam DBSCAN ditandai oleh tiga macam *state*, yaitu *core* (inti), *border* (batas), dan *noise* (gangguan) [13][14][15][16]. Oleh karena itu, penelitian ini bertujuan untuk mencari faktor penentu minat mahasiswa dalam berkuliah setelah pandemi menggunakan algoritma DBSCAN.

## 2. Metode Penelitian

### 2.1. Jenis dan Sumber Data

Data yang digunakan dalam penelitian ini adalah data primer, diperoleh dengan menyebarkan kuesioner kepada beberapa mahasiswa dengan berbagai latar belakang. Kuesioner disebar dengan melalui tautan yang telah disiapkan sebelumnya. Jumlah responden dalam penelitian ini ada lebih kurang 400 responden, yang menghasilkan 385 kuesioner yang valid. Adapun beberapa variabel yang digunakan dalam penelitian ini yaitu, dari segi cara perkuliahan: daring, luring, dan gabungan daring serta luring (*hybrid*). Dari segi latar belakang mahasiswa: usia, jenis kelamin, semester dan IPK. Dari segi keilmuan terdapat: Eksakta dan Non-Eksakta.

### 2.2. Teknik Pengolahan Data

Data dalam penelitian ini akan diolah menggunakan *software* bernama RStudio dan akan menggunakan algoritma DBSCAN. Tahapan dalam pengolahan data dijabarkan sebagai berikut :

1. Membuat kuesioner
2. Mengumpulkan dan melakukan *preprocessing* data

3. Melakukan *clustering* data menggunakan algoritma DBSCAN.
4. Penilaian hasil *cluster* menggunakan uji silhouette.
5. Menganalisis *output* yang dihasilkan dan memberikan simpulan.

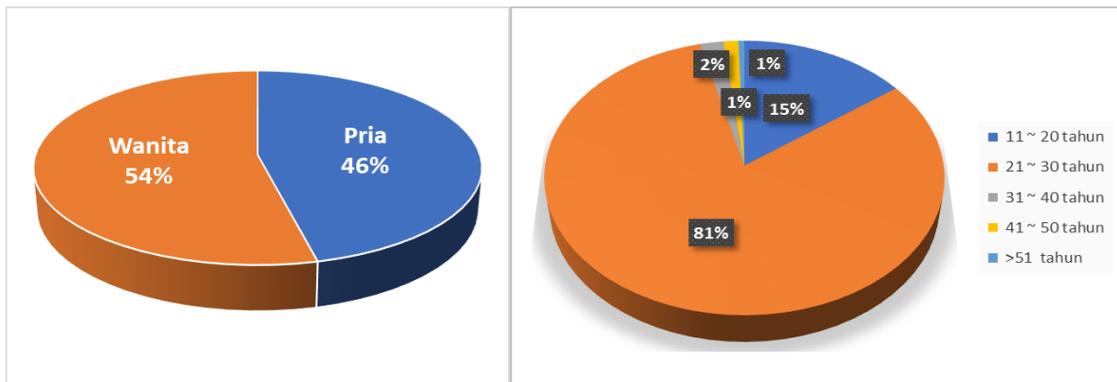
Pembuatan kuesioner menggunakan beberapa variabel. Variabel pertama adalah variabel yang terkait dengan latar belakang mahasiswa, yaitu usia {1}, jenis kelamin {2}, semester {3} dan IPK {4}. Usia akan dikelompokkan menjadi: kelompok 11-20 tahun, kelompok 21-30 tahun, kelompok 31-40 tahun, kelompok 41-50 tahun, kelompok >51 tahun. Jenis kelamin terbagi menjadi dua yaitu pria dan wanita. Variabel semester adalah variabel yang menunjukkan mahasiswa tersebut menginjak semester berapa dalam perkuliahannya dan akan terbagi menjadi: kelompok  $\leq 2$  semester, kelompok  $\leq 4$  semester, kelompok  $\leq 6$  semester, kelompok  $\leq 8$  semester dan kelompok  $> 8$  semester. Variabel IPK akan terbagi menjadi :  $\leq 1$ ,  $\leq 2$ ,  $\leq 3$ , dan  $> 3$ .

Dari segi keilmuan {5} terdapat bidang eksakta dan noneksakta. Eksakta adalah perkuliahan dengan angka dan ilmu pasti, seperti teknik, matematika, fisika, biologi, teknologi informasi, dan sejenisnya. Sementara ilmu noneksakta adalah perkuliahan lebih condong ke berbagai teori, seperti bahasa, ilmu hukum, psikologi, sosial-budaya, komunikasi dan sebagainya.

Sebelum proses *clustering* data menggunakan algoritma DBSCAN, *preprocessing data* adalah proses yang bertujuan untuk menyeragamkan dan merapikan data. Tujuan proses ini adalah untuk membentuk data menjadi suatu bentuk yang homogen, supaya menghindari terjadinya adanya data bias dalam proses *clustering*. Hasil dari *clustering* menggunakan algoritma DBSCAN adalah beberapa *cluster* atau beberapa kelompok yang berisi motivasi mahasiswa memilih cara perkuliahan tersebut berdasarkan variabel yang telah dipersiapkan sebelumnya. Penilaian baik tidaknya *cluster* yang dihasilkan menggunakan *silhouette coefficient*, yang merupakan sebuah metode yang akan memberitahu keefektifitas proses *clustering* dari algoritma tersebut. Proses metode ini adalah mengukur jarak sebaran titik dalam sebuah *cluster*, yaitu jarak setiap titik dan titik lainnya di dalam *cluster* tersebut. *Cluster* yang baik muncul ketika nilai *silhouette coefficient* mendekati 1. Nilai ini menandakan jarak antartitik pada *cluster* tersebut kecil atau *cluster* tersebut rapat. Jika nilai *silhouette coefficient* cenderung mendekati negatif 1 (-1) maka *cluster* tersebut kurang baik. Hal itu menandakan bahwa jarak antartitik dalam *cluster* tersebut semakin jauh.

### 3. Hasil dan Pembahasan

Hasil dari kuesioner yang disebarakan memberikan 385 kuesioner valid. Dari hasil tersebut dapat diketahui beberapa karakteristik seperti yang terlihat pada Gambar 1. Terdapat 178 responden pria (46%) dan 207 responden wanita (54%). Berdasarkan usia /umur responden, 15% atau 57 responden berumur antara 11-20 tahun, 81% atau 313 responden berumur antara 21-30 tahun, 2% atau 8 responden berumur antara 31-40 tahun, 1% atau 5 responden berumur antara 41-50 tahun, dan 1% atau 2 responden berumur antara >51 tahun.



Gambar 1. Persentase jenis kelamin dan umur responden

Hasil dari *clustering* menggunakan algoritma DBSCAN akan memberikan beberapa *cluster* atau kelompok. Kelompok ini tercipta dari kaitan variabel kepada metode perkuliahan yang dipilih, sehingga dapat diketahui apa saja penentu ketika seseorang mahasiswa memilih metode perkuliahan tersebut. Selanjutnya pengukuran efektifitas dan performa algoritma DBSCAN dilakukan menggunakan *Silhouette Coefficient*. Hasil *Silhouette Coefficient* akan mendekati nilai 1, jika *cluster* yang terbentuk cukup *dense* dan tidak menyebar acak. Hal ini juga menunjukkan proses *clustering* berhasil dengan baik. Sebaliknya, jika nilai *Silhouette Coefficient* mendekati -1, hal ini berarti *cluster* yang dihasilkan renggang dan cenderung menjauh, dan menunjukkan proses *clustering* yang kurang baik.

**Tabel 1.** Nilai *Silhouette Coefficient* DBSCAN

<i>Cluster</i>	<i>Silhouette Value</i>
{1}&{2}	0.92
{1}	0.9
{2}	0.91
{3}	1
{4}	1
{5}	0.99

Tabel 1 di atas merupakan nilai *Silhouette Coefficient* dari *clustering* dengan algoritma DBSCAN. Hasil pada Tabel 1 menunjukkan bahwa DBSCAN memiliki performa yang baik. Hal ini dapat dilihat dari nilai *Silhouette Coefficient* yang mendekati 1. *Cluster* responden memilih metode perkuliahan luring (tatap muka) adalah mahasiswa dengan semester  $\leq 2$ ,  $\leq 4$ ,  $\leq 6$  dan kelompok ilmu Eksakta. *Cluster* responden memilih menggunakan metode perkuliahan daring (*online*) adalah mahasiswa dengan semester  $> 8$  dan kelompok ilmu Non-Eksakta. *Cluster* responden memilih metode perkuliahan *hybrid* adalah mahasiswa dengan semester  $\leq 4$  dan  $\leq 6$  serta IPK  $> 3$ . Dari kelompok yang terbentuk, dapat terlihat faktor yang mempengaruhi minat metode perkuliahan mahasiswa adalah semester yang sedang ditempuh mahasiswa tersebut, IPK yang dimiliki dan kelompok ilmu (eksakta atau non-eksakta).

#### 4. Kesimpulan

Dari hasil penelitian ini dapat disimpulkan bahwa algoritma DBSCAN menunjukkan performa yang baik dalam proses *clustering* dengan data yang besar. Ditemukan faktor penentu pemilihan minat metode perkuliahan mahasiswa adalah semester, IPK, dan kelompok ilmu (Eksakta atau Non-Eksakta). Hal ini menunjukkan ketika mahasiswa memilih berkuliah daring (*online*) atau luring (tatap muka) adalah berdasar pada variabel tersebut. Mahasiswa dengan semester di atas 8 lebih memilih pembelajaran secara *online*. Mahasiswa dari kelompok ilmu non-eksakta juga memilih untuk berkuliah secara *online* dan

mahasiswa dengan IPK >3 juga memilih kuliah secara *online*. Sementara itu, mahasiswa dengan semester kurang dari 6 lebih memilih perkuliahan secara *offline*. Mahasiswa dari kelompok ilmu eksakta juga memilih untuk berkuliahan secara *offline*. Variabel usia dan jenis kelamin tidak mempengaruhi pemilihan minat metode perkuliahan mahasiswa tersebut, baik daring ataupun luring.

## Referensi

- [1] N. H. Nasution *et al.*, “Gambaran Pengetahuan Masyarakat Tentang Pencegahan Covid-19 Kecamatan Padangsidempuan Batunadua, Kota Padangsidempuan,” *Jurnal Biomedika dan Kesehatan*, vol. 4, no. 2, pp. 47–49, 2021.
- [2] I. Wahidah, R. Athallah, N. F. S. Hartono, M. C. A. Rafqie, and M. A. Septiadi, “Pandemik COVID-19: Analisis Perencanaan Pemerintah dan Masyarakat dalam Berbagai Upaya Pencegahan,” *Jurnal Manajemen dan Organisasi*, vol. 11, no. 3, pp. 179–188, 2020, doi: 10.29244/jmo.v11i3.31695.
- [3] Yelvi Levani, Aldo Dwi Prastya, and Siska Mawaddatunnadila, “Coronavirus Disease 2019 (COVID-19): Patogenesis, Manifestasi Klinis dan Pilihan Terapi,” *Jurnal Kedokteran dan Kesehatan*, vol. 17, no. 1, pp. 44–57, 2021, [Online]. Available: <https://jurnal.umj.ac.id/index.php/JKK/article/view/6340>.
- [4] I. P. M. H. Purba, “Implementasi Undang-Undang Nomor 6 Tahun 2018 Tentang Keekarantinaan Kesehatan di Jawa Timur Menghadapi Pandemi COVID-19,” *Journal of Chemical Information and Modeling*, vol. 4, no. 1, pp. 1–11, 2021.
- [5] M. Siahaan, “Dampak Pandemi Covid-19 Terhadap Dunia Pendidikan,” *Jurnal Kajian Ilmiah*, vol. 1, no. 1, pp. 73–80, 2020, doi: 10.31599/jki.v1i1.265.
- [6] P. R. Sari, N. B. Tussyantari, and M. Suswandari, “DAMPAK PEMBELAJARAN DARING BAGI SISWA SEKOLAH DASAR SELAMA SELAMA COVID-19 Universitas Veteran Bangun Nusantara , Sukoharjo,” *Prima Magistra: Jurnal Ilmiah Kependidikan*, vol. 2, no. 1, pp. 9–15, 2021.
- [7] U. Hanifah Salsabila, L. Irna Sari, K. Haibati Lathif, A. Puji Lestari, and A. Ayuning, “Peran Teknologi Dalam Pembelajaran Di Masa Pandemi Covid-19,” *Al-Mutharahah: Jurnal Penelitian dan Kajian Sosial Keagamaan*, vol. 17, no. 2, pp. 188–198, 2020, doi: 10.46781/al-mutharahah.v17i2.138.
- [8] R. H. S. Aji, “Dampak Covid-19 pada Pendidikan di Indonesia: Sekolah, Keterampilan, dan Proses Pembelajaran,” *SALAM: Jurnal Sosial dan Budaya Syar-i*, vol. 7, no. 5, pp. 395–402, 2020, doi: 10.15408/sjsbs.v7i5.15314.
- [9] A. Cahyani, I. D. Listiana, and S. P. D. Larasati, “Motivasi Belajar Siswa SMA pada Pembelajaran Daring di Masa Pandemi Covid-19,” *IQ (Ilmu Al-qur’an): Jurnal Pendidikan Islam*, vol. 3, no. 01, pp. 123–140, 2020, doi: 10.37542/iq.v3i01.57.
- [10] W. A. F. Dewi, “Dampak COVID-19 terhadap Implementasi Pembelajaran Daring di Sekolah Dasar,” *Edukatif: Jurnal Ilmu Pendidikan*, vol. 2, no. 1, pp. 55–61, 2020, doi: 10.31004/edukatif.v2i1.89.
- [11] R. K. Dinata, S. Safwandi, N. Hasdyna, and N. Azizah, “Analisis K-Means Clustering pada Data Sepeda Motor,” *INFORMAL: Informatics Journal*, vol. 5, no. 1, pp. 10–17, 2020, doi: 10.19184/isj.v5i1.17071.
- [12] Y. D. Darmi and A. Setiawan, “Penerapan Metode Clustering K-Means Dalam Pengelompokan Penjualan Produk,” *Jurnal Media Infotama*, vol. 12, no. 2, pp. 148–157, 2017, doi: 10.37676/jmi.v12i2.418.
- [13] A. Kristianto, E. Sedyono, and K. D. Hartomo, “Implementation dbscan algorithm to clustering

- satellite surface temperature data in indonesia,” *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 6, no. 2, pp. 109–118, 2020, doi: 10.26594/register.v6i2.1913.
- [14] A. Kristianto, “Analisa Performa K-Means dan DBSCAN dalam Clustering Minat Penggunaan Transportasi Umum,” *Elkom : Jurnal Elektronika dan Komputer*, vol. 14, no. 2, pp. 368–372, 2021, doi: 10.51903/elkom.v14i2.551.
- [15] S. Budiman, D. Safitri, and D. Ispriyanti, “Perbandingan Metode K-Means Dan Metode Dbscan Pada Pengelompokan Rumah Kost Mahasiswa Di Kelurahan Tembalang Semarang,” *JURNAL GAUSSIAN*, vol. 5, no. 4, pp. 757–762, 2016.
- [16] N. M. A. S. Devi, I. K. G. D. Putra, and I. M. Sukarsa, “Implementasi Metode Clustering DBSCAN pada Proses Pengambilan Keputusan,” *Jurnal Ilmiah Lontar Komputer*, vol. 6, no. 3, pp. 655–661, 2015.