

**Studi Komparasi antara Metode *k-Means*, *k-Means++*, dan *k-Medoids* untuk Klusterisasi Tingkat Kesejahteraan Penduduk**

**(Studi Kasus Dinas PPKB dan P3A Kab. Wonogiri)**

**TESIS**

Untuk Memenuhi Sebagian Persyaratan Mencapai Derajat  
Magister S-2

Program Studi Teknik Informatika



**Disusun oleh:**

**PRASDIKA FBS**

**18206050013**

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA  
PROGRAM STUDI TEKNIK INFORMATIKA

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA**

**YOGYAKARTA**

**2020**

## PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Prasdika Fahdiatmoko Bintang Septiandaru

NIM : 18206050013

Jenjang : Magister

Program Studi : Informatika

menyatakan bahwa naskah tesis ini secara keseluruhan adalah hasil penelitian/karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sumbernya.

Yogyakarta, 25 Agustus 2020

Saya yang menyatakan,



Prasdika F.B.S

NIM: 18206050013

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama : Prasdika Fahdiatmoko Bintang Septiandaru

NIM : 18206050013

Jenjang : Magister

Program Studi : Informatika

menyatakan bahwa naskah tesis ini secara keseluruhan benar-benar bebas dari plagiasi. Jika di kemudian hari terbukti melakukan plagiasi, maka saya siap ditindak sesuai ketentuan hukum yang berlaku.

Yogyakarta, 25 Agustus 2020

Saya yang menyatakan,



Prasdika F.B.S.

NIM: 18206050013

STATE ISLAMIC UNIVERSITY  
SUNAN KALIJAGA  
YOGYAKARTA

## PENGESAHAN TUGAS AKHIR

Nomor : B-2246/Un.02/DST/PP.00.9/09/2020

Tugas Akhir dengan judul : STUDI KOMPARASI ANTARA METODE K-MEANS, K-MEANS++, DAN K-MEDOIDS UNTUK KLASTERISASI TINGKAT KESEJAHTERAAN PENDUDUK (STUDI KASUS DINAS PPKB DAN P3A KABUPATEN MONOGIRI)

yang dipersiapkan dan disusun oleh:

Nama : PRASDIKA FAHDIATMOKO BINTANG  
SEPTIANDARU, S. Kom  
Nomor Induk Mahasiswa : 18206050013  
Telah diujikan pada : Senin, 31 Agustus 2020  
Nilai ujian Tugas Akhir : A-

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

### TIM UJIAN TUGAS AKHIR



Ketua Sidang/Penguji I

Maria Ulfah Siregar, S.Kom. MIT., Ph.D.  
SIGNED

Valid ID: 5f6af0c48f835



Penguji II

Dr. Ir. Bambang Sugiantoro, S.Si., M.T.  
SIGNED

Valid ID: 5f71735997a03



Penguji III

Dr. Agung Fatwanto, S.Si., M.Kom.  
SIGNED

Valid ID: 5f6a74e0d4edc



Yogyakarta, 31 Agustus 2020  
UIN Sunan Kalijaga  
Dekan Fakultas Sains dan Teknologi

Dr. Hj. Khurul Wardati, M.Si.  
SIGNED

Valid ID: 5f719f745d96c

## NOTA DINAS PEMBIMBING

Kepada Yth.,  
Dekan Fakultas Sains dan Teknologi  
UIN Sunan Kalijaga Yogyakarta

*Assalamu 'alaikum wr. wb.*

Setelah melakukan bimbingan, arahan, dan koreksi terhadap penulisan tesis yang berjudul:

**Studi Komparasi antara Metode  $k$ -Means,  $k$ -Means++, dan  $k$ -Medoids untuk Klasterisasi Tingkat Kesejahteraan Penduduk (Studi Kasus Dinas PPKB dan P3A Kab. Wonogiri)**

Yang ditulis oleh:

Nama : Prasdika F.B.S  
NIM : 18206050013  
Jenjang : Magister  
Program Studi : Informatika

Saya berpendapat bahwa tesis tersebut sudah dapat diajukan kepada Magister Informatika UIN Sunan Kalijaga untuk diujikan dalam rangka memperoleh gelar Magister Informatika. Wassalamu'alaikum wr. wb.

Yogyakarta, Agustus 2020

Pembimbing,



Maria Ulfah Siregar, S.Kom. MIT., Ph.D.

NIP. 19780106 200212 2 001

## ABSTRAK

Seiring berkembang pesatnya pertumbuhan data di dunia nyata dan dengan data yang hanya memiliki sedikit atau bahkan tanpa latar belakang, klusterisasi memiliki peran penting dalam eksplorasi data untuk dapat menentukan pola dan manfaat dari sejumlah besar dataset. Penelitian ini melakukan klusterisasi dengan beberapa algoritma dari *unsupervised learning* yaitu *k-Means*, *k-Means++* dan *k-Medoids*. Data yang digunakan adalah data hasil survei kependudukan dari dinas PPKB dan P3A Kab. Wonogiri Tahun 2018. Dalam penelitian ini dilakukan tahap ekstraksi fitur dengan metode *principal component analysis (PCA)* dan pra-pemrosesan data menggunakan *standard scaler* dan *min-max scaler*. Dengan penerapan ketiga algoritma klusterisasi tersebut didapatkan hasil bahwa algoritma *k-Means++* memiliki nilai rata-rata inerti terkecil dengan 15.954711235353362 pada data hasil standarisasi dengan *standard scaler* dan 1.13967411934335812 pada data hasil standarisasi dengan *min-max scaler*, sedangkan *k-Medoids* menghasilkan rata-rata waktu pemrosesan tercepat dengan 0.0145999434 detik pada data hasil standarisasi dengan *standard scaler* dan 0.0142000198 detik pada data hasil standarisasi dengan *min-max scaler*.

Kata kunci: *Clustering*, *Data Mining*, *k-Means*, *k-Means++*, *k-Medoids*

## **ABSTRACT**

*Along the rapid growth of data growth in the real world and with the data that have little or no background, clustering has an important role in the exploration of the data to determine patterns and benefit from a large number of datasets. This study did clustering with some of unsupervised learning algorithms are k-Means, k-Means++ and k-Medoids. The data used is data from official population survey results PPKB and P3A Wonogiri 2018. In this study, a feature extraction stage was carried out using the principal component analysis (PCA) method and data pre-processing using a standard scaler and a min-max scaler. With the implementation of the three algorithms clustering it showed that the algorithm k-means ++ has an average value of inertia smallest with 15.954711235353362 on data from standardization to the standard scaler and 1.13967411934335812 on outcome data standardization with the min-max scaler, while k-Medoids produce an average time 0.0145999434 seconds with the fastest processing on the data standardization results with standard scaler and 0.0142000198 seconds on the result data standardization with the min-max scaler.*

*Keywords: Clustering, Data Mining, k-Means, k-Means++, k-Medoids*

## KATA PENGANTAR

Dengan memanjatkan Puji dan Syukur Kehadirat Allah SWT atas segala rahmat dan karunia-Nya, akhirnya penulis dapat menyelesaikan penyusunan tesis yang berjudul Studi Komparasi antara Metode *k-Means*, *k-Means++*, dan *k-Medoids* untuk Klasterisasi Tingkat Kesejahteraan Penduduk (Studi Kasus Dinas PPKB dan P3A Kab. Wonogiri).

Tesis ditulis dalam rangka memenuhi persyaratan untuk memperoleh gelar Magister di UIN Sunan Kalijaga Yogyakarta. Penulis menyadari bahwa tesis dapat diselesaikan berkat dukungan dan bantuan dari berbagai pihak, oleh karena itu penulis berterima kasih kepada semua pihak yang secara langsung maupun tidak langsung memberikan kontribusi dalam menyelesaikan Tesis ini. Selanjutnya ucapan terima kasih penulis sampaikan kepada:

1. Bapak Prof. Dr. Phil. Al Makin, S.Ag., M.A. sebagai Rektor UIN Sunan Kalijaga Yogyakarta.
2. Bapak Dr. Bambang Sugiantoro, S.Si., M.T. sebagai Kepala Program Studi Magister Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta, yang telah memberikan izin dan kemudahan sehingga penulis dapat menyelesaikan studi di Program Studi Program Magister Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Ibu Maria Ulfah Siregar, S.Kom. MIT., Ph.D. sebagai pembimbing, yang dengan sabar membimbing penulis dan meluangkan waktunya untuk memberikan bimbingan, petunjuk, dan memberikan motivasi



dan semangat kepada penulis untuk menyelesaikan penulisan Tesis ini.

4. Seluruh dosen dan staf administrasi pada Program Studi Informatika Program Magister Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta, yang secara langsung atau tidak langsung telah memberi bantuan kepada penulis dalam menyelesaikan penulisan tesis.
5. Orang Tua tercinta yang telah mendidik dengan penuh rasa kasih sayang dan senantiasa memberi do'a, semangat dan dorongan kepada penulis.
6. Istri tercinta yang selalu membirikan do'a dan dukungan kepada penulis.
7. Ara dan Hasan yang menjadi semangat penulis untuk menyelesaikan penelitian.
8. Seluruh rekan-rekan sejawat di Magister Informatika yang telah saling mendukung untuk melalui perjuangan bersama-sama.

Akhirnya penulis berharap semoga Tesis ini dapat memberikan manfaat. Permintaan maaf penulis ucapkan jika seandainya dalam penulisan Tesis ini terdapat kekurangan dan kekeliruan, penulis juga menerima kritik dan saran yang bersifat membangun demi menyempurnakan penulisan tesis ini.

## DAFTAR ISI

Halaman Judul .....	i
Pernyataan Keaslian .....	ii
Pernyataan Bebas Plagiasi .....	iii
Halaman Pengesahan .....	iv
Nota Dinas Pembimbing .....	v
Abstrak .....	vi
<i>Abstract</i> .....	vii
Kata Pengantar .....	viii
Daftar Isi .....	x
Daftar Tabel .....	xiii
Daftar Gambar .....	xiv
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	4
1.6. Sistematika Penulisan.....	4
<b>BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....</b>	<b>5</b>
2.1. Tinjauan Pustaka.....	5
2.2. Landasan Teori .....	7
2.2.1. <i>Data Mining</i> .....	8
2.2.2. Teknik <i>Data mining</i> .....	10
2.2.3. Machine Learning.....	12
2.2.4. Unsupervised Learning.....	14

2.2.5.	Klasterisasi.....	15
2.2.6.	Distance Matrics .....	16
2.2.7.	Ekstrasi fitur.....	19
2.2.8.	Data Preprocessing .....	21
2.2.8.1	<i>Standard Scaler</i> .....	24
2.2.8.2	<i>Min Max Scaler</i> .....	25
2.2.9.	<i>k-Means</i> .....	25
2.2.10.	<i>k-Means++</i> .....	29
2.2.11.	<i>k-Medoids</i> .....	30
2.2.12.	Jumlah Klaster Optimal .....	32
2.2.12.1.	Elbow Methode.....	32
2.2.12.2.	Silhouette Score.....	34
<b>BAB III</b>	<b>METODE PENELITIAN .....</b>	<b>35</b>
3.1.	Jenis dan Desain Penelitian .....	35
3.2.	Data dan Sumber Data.....	37
3.3.	Ekstraksi Fitur .....	38
3.4.	Data Preprocessing .....	39
3.4.1.	<i>Standard Scaler</i> .....	40
3.4.2.	<i>Min Max Scaler</i> .....	41
3.5.	Algoritma <i>k-Means</i> .....	41
3.6.	Algoritma <i>k-Medoids</i> .....	42
3.7.	Elbow Methode .....	43
<b>BAB IV</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>47</b>
4.1.	Ekstrasi fitur.....	47
4.2.	Data Preprocessing.....	49
4.2.1.	<i>Standard Scaler</i> .....	50
4.2.2.	<i>Min Max Scaler</i> .....	51
4.3.	Elbow Methode.....	53

4.4. Silhouette Score .....	53
4.5. Hasil dan Analisis Pengujian .....	54
4.5.1. Pengujian Terhadap Data <i>Standard Scaler</i> .....	54
4.5.2. Pengujian Terhadap Data <i>Min Max Scaler</i> .....	60
4.6. Evaluasi Metode Klasterisasi .....	65
4.6.1. Perbandingan Hasil Percobaan .....	65
4.6.2. Evaluasi Hasil Pengujian .....	66
4.6.3. Perbandingan Hasil Klasterisasi .....	67
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>68</b>
5.1. Kesimpulan .....	68
5.2. Saran .....	69

## DAFTAR TABEL

Tabel 2.1. Penelitian Terdahulu .....	5
Tabel 3.1. Langkah-langkah preprocessing .....	22
Tabel 4.1. Data PCA .....	47
Tabel 4.2. Data <i>Standard Scaler</i> Pengujian 1 .....	54
Tabel 4.3. Data <i>Standard Scaler</i> Pengujian 2 .....	55
Tabel 4.4. Data <i>Standard Scaler</i> Pengujian 3 .....	57
Tabel 4.5. Data <i>Standard Scaler</i> Pengujian 4 .....	58
Tabel 4.6. Data <i>Standard Scaler</i> Pengujian 5 .....	59
Tabel 4.7. Data <i>Min Max Scaler</i> Pengujian 1 .....	60
Tabel 4.8. Data <i>Min Max Scaler</i> Pengujian 2 .....	61
Tabel 4.9. Data <i>Min Max Scaler</i> Pengujian 3 .....	62
Tabel 4.10. Data <i>Min Max Scaler</i> Pengujian 4 .....	63
Tabel 4.11. Data <i>Min Max Scaler</i> Pengujian 5 .....	64
Tabel 4.12. Rata-rata dengan data <i>Standard Scaler</i> .....	65
Tabel 4.13. Rata-rata dengan data <i>Min Max Scaler</i> .....	65

## DAFTAR GAMBAR

Gambar 2.1. Tahap-tahap <i>Data Mining</i> .....	9
Gambar 2.2. Model Sistem AI .....	14
Gambar 2.3. Elbow Methode .....	34
Gambar 3.1. Alur Penelitian .....	36
Gambar 3.2. Grafik Persebaran Penduduk Kab. Wonogiri .....	38
Gambar 4.1. Grafik data PCA .....	48
Gambar 4.2. Plot Data PCA .....	49
Gambar 4.3. Grafik dengan <i>standard scaler</i> .....	50
Gambar 4.4. Plot data dengan <i>standard scaler</i> .....	51
Gambar 4.5. Grafik dengan <i>min max scaler</i> .....	52
Gambar 4.6. Plot data <i>min max scaler</i> .....	52
Gambar 4.7. Jumlah kluster optimal menggunakan elbow metode ...	53
Gambar 4.8. Jumlah kluster optimal menggunakan silhouette score ...	54
Gambar 4.9. Hasil klasterisasi <i>Standard Scaler</i> Uji-1 .....	55
Gambar 4.10. Hasil klasterisasi <i>Standard Scaler</i> Uji-2 .....	56
Gambar 4.11. Hasil klasterisasi <i>Standard Scaler</i> Uji-3 .....	57
Gambar 4.12. Hasil klasterisasi <i>Standard Scaler</i> Uji-4 .....	58
Gambar 4.13. Hasil klasterisasi <i>Standard Scaler</i> Uji-5 .....	59
Gambar 4.14. Hasil klasterisasi <i>Min Max Scaler</i> Uji-1 .....	60
Gambar 4.15. Hasil klasterisasi <i>Min Max Scaler</i> Uji-2 .....	61
Gambar 4.16. Hasil klasterisasi <i>Min Max Scaler</i> Uji-3 .....	62
Gambar 4.17. Hasil klasterisasi <i>Min Max Scaler</i> Uji-4 .....	63
Gambar 4.18. Hasil klasterisasi <i>Min Max Scaler</i> Uji-5 .....	64

Gambar 4.19. Komparasi hasil klasterisasi *k-Means*, *k-Means ++* dan *k-Medoids*..... 68



# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Dinas Kependudukan dan Keluarga Berencana dan Pemberdayaan Perempuan dan Perlindungan Anak (DPPKBP3A) Kabupaten Wonogiri merupakan salah satu instansi yang melakukan penyuluhan terhadap terciptanya keluarga sejahtera. Sebagai upaya untuk meningkatkan kesejahteraan masyarakat maka dilakukan pendataan tentang keluarga berencana dan sejahtera dalam beberapa periode untuk mengetahui data pertumbuhan tingkat kesejahteraan. Disebutkan dalam Undang-Undang Republik Indonesia Nomor 10 Tahun 1992 Pasal 1 tentang Perkembangan Kependudukan dan Pembangunan Keluarga Sejahtera dalam poin ke-11, keluarga sejahtera adalah keluarga yang dibentuk berdasarkan atas perkawinan yang sah, mampu memenuhi kebutuhan hidup spiritual dan materiil yang layak, bertaqwa kepada Tuhan Yang Maha Esa, memiliki hubungan yang serasi, selaras, dan seimbang antar anggota dan antara keluarga dengan masyarakat dan lingkungan (UU Nomor 10 Tahun 1992 Pasal 1).

Dalam kurun waktu yang tidak singkat data yang terkumpul sangatlah banyak dan di dalam data tersebut tersimpan banyak informasi yang mungkin belum terlalu dimanfaatkan oleh pihak instansi. Dalam penelitian ini peneliti akan melakukan *extracting* pada data yang telah terkumpul menggunakan metode *data mining* dengan algoritma *k-Means*, *k-Means++* dan *k-Medoids*.



*Data mining* merupakan serangkaian proses untuk mendapatkan informasi yang berguna dari gudang data yang besar. *Data mining* juga dapat diartikan sebagai pengekstrakan informasi baru yang diambil dari bongkahan data besar yang membantu dalam pengambilan keputusan. Dalam *data mining* terdapat banyak teknik dalam implementasinya, diantaranya yaitu algoritma Naïve Bayes, decision tree, jaringan saraf tiruan dan masih banyak lainnya.

Dengan dimanfaatkannya data tersebut diharapkan dapat membantu dalam meningkatkan kesejahteraan keluarga. Dengan metode *data mining* diharapkan dapat membantu pihak instansi dalam memantau tingkat kesejahteraan dari setiap periodenya dengan lebih mudah dan cepat. Selain itu juga dapat digunakan sebagai sarana pengambilan keputusan untuk lebih meningkatkan pelayanan keluarga berencana.

Dari permasalahan di atas dapat diambil alternatif dengan cara memanfaatkan teknik *data mining* dengan menggunakan metode *k-Means*, *k-Means++* dan *k-Medoids* untuk identifikasi tingkat kesejahteraan keluarga. Dengan harapan setelah diolah dengan *data mining* dapat membantu menemukan informasi dalam identifikasi tingkat kesejahteraan keluarga.

Metode klasterisasi dengan algoritma *k-Means*, *k-Means++* dan *k-Medoids* dipilih oleh peneliti karena data bersifat numerik. Data yang digunakan dalam penelitian ini juga belum memiliki klasifikasi sebelumnya, jadi metode klasterisasi dipilih peneliti sebagai metode yang dapat dimanfaatkan. Pemanfaatan dari ketiga algoritma tersebut bertujuan untuk meneliti algoritma dengan

performa terbaik dan hasil paling akurat untuk studi kasus yang sedang diteliti.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang di atas dapat dirumuskan permasalahan yang akan diselesaikan dalam penelitian ini adalah:

1. Apakah implementasi *data mining* dapat digunakan untuk menentukan pola dan kluster dari tingkat kesejahteraan penduduk?
2. Bagaimana tahap-tahapan *data mining* terhadap data pendataan keluarga berencana dan sejahtera?
3. Bagaimana mengevaluasi algoritma *k-Means*, *k-Means++* dan *k-Medoids* terkait kluster yang dihasilkan?

## 1.3. Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini adalah:

1. Jumlah data yang diuji coba adalah sebanyak 25 data dengan 14 variabel pada masing-masing data.
2. Data yang digunakan adalah data survei kependudukan tahun 2018.

## 1.4. Tujuan Penelitian

Adapun tujuan penelitian berdasarkan rumusan masalah adalah:

1. Implementasi algoritma *k-Means*, *k-Means++* dan *k-Medoids* untuk menentukan pola pertumbuhan keluarga sejahtera.
2. Menentukan kluster data tingkat kesejahteraan penduduk.
3. Evaluasi hasil penerapan algoritma *k-Means*, *k-Means++* dan *k-Medoids* terhadap hasil kluster.

### **1.5. Manfaat Penelitian**

Dari penelitian ini diharapkan dapat memberikan kontribusi pada upaya peningkatan kesejahteraan penduduk melalui ekstraksi informasi dan menghasilkan klaster tingkat kesejahteraan penduduk sehingga dapat digunakan untuk pengambilan keputusan yang lebih tepat sasaran.

### **1.6. Sistematika Penulisan**

Laporan penelitian ini tersusun atas beberapa bab pembahasan. Sistematika penulisan bab tersebut adalah:

**BAB I** Pendahuluan

**BAB II** Tinjauan Pustaka dan Landasan Teori

**BAB III** Metode Penelitian

**BAB IV** Hasil dan Pembahasan

**BAB V** Kesimpulan dan Saran

## BAB V

### KESIMPULAN DAN SARAN

#### 5.1. Kesimpulan

Dari penelitian yang telah dilakukan dapat ditarik beberapa kesimpulan yaitu:

1. Proses normalisasi data menggunakan *standard scaler* dan *min max scaler* mempengaruhi jumlah iterasi, nilai *inertia* dan waktu dari masing metode.
2. *k-Means++* menghasilkan hasil iterasi dan waktu yang lebih baik dengan menggunakan data hasil standarisasi dengan metode *min max scaler*.
3. Dari ketiga metode yang digunakan hasil *inertia* terbaik ditunjukkan oleh metode *k-Means++* akan tetapi metode *k-Means++* memerlukan waktu rata-rata yang paling lama dalam mencapai konvergensi.
4. Inisialisasi awal terhadap pusat data awal pada *k-Medoids* yang menggunakan data itu sendiri menghasilkan perulangan dan hasil klasterisasi yang konsisten terhadap metode *k-Medoids*.
5. Tingkat persebaran dan densitas data sangat mempengaruhi performa dari masing-masing algoritma.
6. Metode *k-Medoids* lebih sensitive terhadap outlier yang berarti *k-Medoids* lebih cocok untuk model persebaran data yang lebih jauh, ditunjukkan dengan penggunaan standarisasi nilai data pada *min max scaler* menjadikan nilai *inertia* pada kedua metode lain jauh lebih baik.

## 5.2. Saran

Banyaknya data dalam dataset sangat mempengaruhi hasil dari ketiga parameter pengujian. Oleh karena itu, perlu dikembangkan lebih lanjut penelitian menggunakan ketiga algoritma dengan dataset yang lebih besar untuk meningkatkan akurasi dari rata-rata waktu pengujian dan nilai *inertia*.

## REFERENSI

- Arthur, D., & Vassilvitskii, S. (2007). *k-Means++*: The advantages of careful seeding. In *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms* (Vol. 07-09-Janu, pp. 1027–1035).
- API design for machine learning software: experiences from the scikit-learn project, Buitinck et al., 2013.
- Ghahramani, Z. (2004). Unsupervised Learning BT - Advanced Lectures on Machine Learning. *Advanced Lectures on Machine Learning*, 3176(Chapter 5), 72–112. [https://doi.org/10.1007/978-3-540-28650-9\\_5](https://doi.org/10.1007/978-3-540-28650-9_5)
- Han, J., Kamber, M., & Pei, J. (2014). *Data mining: Data mining concepts and techniques. Proceedings - 2013 International Conference on Machine Intelligence Research and Advancement, ICMIRA 2013.* <https://doi.org/10.1109/ICMIRA.2013.45>
- Larose, D. T. (2005). *Discovering Knowledge in Data: An Introduction to Data Mining. Discovering Knowledge in Data: An Introduction to Data Mining.* JOHN WILEY & SONS, INC. <https://doi.org/10.1002/0471687545>
- Lemm, S., Blankertz, B., Dickhaus, T., & Müller, K. R. (2011). Introduction to machine learning for brain imaging. *NeuroImage*, 56(2), 387–399. <https://doi.org/10.1016/j.neuroimage.2010.11.004>

- Mohammed, M., Khan, M. B., & Bashie, E. B. M. (2016). *Machine learning: Algorithms and applications. Machine Learning: Algorithms and Applications.* <https://doi.org/10.1201/9781315371658>
- Shalev-Shwartz, S., & Ben-David, S. (2013). *Understanding machine learning: From theory to algorithms. Understanding Machine Learning: From Theory to Algorithms* (Vol. 9781107057). <https://doi.org/10.1017/CBO9781107298019>
- Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
- Turban, E., Aronson, E. J., & Liang, T.-P. (2005). *Decision Support Systems and Intelligent Systems Edisi 7 Jilid 2* (Jilid 1). Yogyakarta: Andi.
- Undang-Undang Republik Indonesia Nomor 10 Tahun 1992. *Undang-undang (UU) tentang Perkembangan Kependudukan dan Pembangunan Keluarga Sejahtera. 16 April 1992.* Lembaran Negara Republik Indonesia Tahun 1992. Jakarta.
- Vercellis, C. (2009). *Business Intelligence: Data Mining and Optimization for Decision Making. Business Intelligence: Data Mining and Optimization for Decision Making.* <https://doi.org/10.1002/9780470753866>